TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

KUZULARDA BÜYÜME ÖRNEĞİNDE TAHMİN MODELLERİNİN PERFORMANSLARININ İNCELENMESİ

Özlem GÜLLÜ

BİYOİSTATİSTİK ANABİLİM DALI
DOKTORA TEZİ

DANIŞMAN
Prof. Dr. İ. Safa GÜRCAN

2015- ANKARA

TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

KUZULARDA BÜYÜME ÖRNEĞİNDE TAHMİN MODELLERİNİN PERFORMANSLARININ İNCELENMESİ

Özlem GÜLLÜ

BİYOİSTATİSTİK ANABİLİM DALI

DOKTORA TEZİ

DANIŞMAN
Prof. Dr. İ. Safa GÜRCAN

2015- ANKARA

Ankara Üniversitesi Sağlık Bilimleri Enstitüsü
Biyoistatistik Doktora Programı
çerçevesinde yürütülmüş olan bu çalışma, aşağıdaki jüri tarafından
Doktora Tezi olarak kabul edilmiştir.

Tez Savunma Tarihi: 22 / 06 / 2015

Prof. Dr. Necmettin ÜNAL
Ankara Üniversitesi
Jüri Başkanı

Prof. Dr. İ.Safa GÜRCAN
Ankara Üniversitesi
Raportör

Prof. Dr. Yasemin YAVUZ
Ankara Üniversitesi

Doç. Dr. Erdem KARABULUT
Hacettepe Üniversitesi

Doç. Dr. Mehtap AKÇİL OK
Başkent Üniversitesi

İÇİNDEKİLER

KABUL VE ONAY **ii**
İÇİNDEKİLER **iii**
ÖNSÖZ **v**
SİMGELER VE KISALTMALAR DİZİNİ **vi**
ŞEKİLLER DİZİNİ **vii**
ÇİZELGELER DİZİNİ **viii**
1. GİRİŞ **1**
1.1. Türkiye'de Koyun Yetiştiriciliği 6
1.2. Kuzularda Büyüme ve Büyümeyi Etkileyen Faktörler 8
1.3. Tahmin Modelleri 9
1.3.1. Lojistik Regresyon Modeli 10
1.3.1.1. Parametrelerin Önem Testi 13
1.3.1.2. Çoklu Lojistik Regresyon Modeli 15
1.3.1.3. Çoklu Lojistik Regresyon Modelinin Kurulması 15
1.3.1.4. Modelin Önemlilik Testi 18
1.3.1.5. Modelin Uyumunun Belirlenmesi 19
1.3.1.6. Hosmer-Lemeshow İstatistiği 19
1.4. Tahmin Modellerinin Performanslarının Değerlendirilmesi 22
1.4.1. Klasik Performans Değerlendirme Ölçüleri 26
1.4.1.1. Genel Performans 26
1.4.1.2. Ayrımsama (Discrimination) 28
1.4.1.3. Kalibrasyon (Calibration) 32
1.4.2. Literatürde Yeni Önerilen Performans Ölçüleri 33
1.4.2.1. Tekrar Sınıflandırma (Reclassification) Tablosu ve NRI 33
1.4.2.2. IDI 36
1.4.2.3. IDI ve AUC İlişkisi 39
1.4.2.4. NRI ve AUC İlişkisi 39

2. GEREÇ VE YÖNTEM **42**
2.1. Gereç 42
2.2. Yöntem 42
3. BULGULAR **44**
3.1. Tanımlayıcı İstatistikler 44
3.2. Lojistik Regresyon Analiz Sonuçları 44
3.2.1. Model Seçimi 44
3.2.2. Referans ve Geliştirilmiş Modellerin Oluşturulması 45
3.2.3. Modellerin Uyum İyiliği Testi 47
3.3. Klasik Performans Ölçüleri Analiz Sonuçları 49
3.4. Literatürde Yeni Önerilen Performans Ölçüleri Analiz Sonuçları 51
4. TARTIŞMA **55**
5. SONUÇ VE ÖNERİLER **58**
ÖZET **61**
SUMMARY **62**
KAYNAKLAR **63**
EKLER **68**
ÖZGEÇMİŞ **71**

ÖNSÖZ

Bu tez çalışması süresince, çalışmanın yürütülmesi ve sonuçlandırılması aşamalarında ilgi ve desteğini esirgemeyen değerli danışman hocam Prof. Dr. İsmayıl Safa GÜRCAN'a, çalışmanın yöntemi ve uygulanan analizler çerçevesinde rehberlik yaparak görüş ve önerilerini benimle paylaşan değerli hocam Prof. Dr. Yasemin YAVUZ'a, yorum ve yönlendirmeleri ile her zaman destek olan değerli hocam Prof. Dr. Necmettin ÜNAL'a

Ankara Üniversitesi Tıp Fakültesi Biyoistatistik Anabilim Dalı öğretim üyeleri değerli hocalarım; Prof. Dr. Atilla Halil ELHAN'a, Doç. Dr. Serdal Kenan KÖSE'ye, Doç. Dr. Derya ÖZTUNA'ya, Yrd. Doç. Dr. Beyza DOĞANAY ERDOĞAN'a,

Bu çalışmada, ilgi ve desteklerini esirgemeyen sevgili arkadaşlarım Dr. Mustafa Agah TEKİNDAL'a, Dr. Doğukan ÖZEN'e ve Araş. Gör. Can ATEŞ'e,

Beni bu uzun ve meşakkatli yolda hiçbir zaman yalnız bırakmayan maddi manevi desteğini esirgemeyen çok sevgili annem ve babam Sema ve Mehmet GÜLLÜ'ye, her zaman yanımda olan ve yaşadığım zorluklarda bana destek olan tüm dostlarıma, en içten duygularımla teşekkür ederim.

SİMGELER VE KISALTMALAR DİZİNİ

AIC	Akaike Bilgi Kriteri (Akaike Information Criterion)
AUC	Eğri Altında Kalan Alan (The Area Under the Curve)
CRP	C Reaktif Protein
HDL	Yüksek Yoğunluklu Lipoprotein (High-Density Lipoprotein)
IDI	Bütünleşik Ayrımsama İyileştirmesi (Integrated Discrimination Improvement)
KVH	Kardiyovasküler Hastalık
LDH	Laktat Dehidrogenaz
MODS	Çoklu Organ Fonksiyonu Yetmezliği (Multiple Organ DysFunction Syndrome)
NRI	Net Tekrar Sınıflandırma İyileştirmesi (Net Reclassification Improvement)
R^2	Açıklayıcılık Katsayısı
ROC	İşlem Karakteristik Eğrisi (Receiver Operating Characteristic Curve)
SKA	Sütten Kesim Ağırlığı
SKAD	İstenilen Sütten Kesim Ağırlığında Değil
SPI	Yaşam Tahmin İndeksi

ŞEKİLLER DİZİNİ

Şekil 1.1. ROC Eğrisi 29
Şekil 3.1. Referans modelin kalibrasyon grafiği 48
Şekil 3.2. Geliştirilmiş modelin kalibrasyon grafiği 48
Şekil 3.3. Referans model ve geliştirilmiş modele ilişkin ROC Eğrisi 50
Şekil 3.4. Referans ve geliştirilmiş modelin tahmini olasılık değerlerinin saçılım grafiği 52

ÇİZELGELER DİZİNİ

Çizelge 1.1. Türkiye'de çeşitli türlerde hayvan sayısı 7
Çizelge 1.2. Türkiye'de canlı ve kesilen koyun sayısı ile et miktarı 8
Çizelge 1.3. İstenen/istenmeyen olaya sahip ya da sahip olmayan her bir onlu denekler için risk grubunda gözlenen frekanslar 20
Çizelge 1.4. İstenen/istenemeyen olaya sahip ya da sahip olmayan denekler için sabit kesim noktasına dayanarak oluşturulan gözlenen frekanslar 21
Çizelge 1.5. Klasik ve Yeni performans değerlendirme ölçüleri 25
Çizelge 1.6. ROC tahminleri için sınıflandırma (2x2) tablosu 29
Çizelge 3.1. Değişkenlere ait tanımlayıcı istatistikler 44
Çizelge 3.2. Tek değişkenli lojistik regresyon analiz sonuçları 45
Çizelge 3.3. Referans modele ilişkin geriye doğru değişken çıkarma yöntemi ile en son adımda elde edilen final model 46
Çizelge 3.4. Geliştirilmiş modele ilişkin geriye doğru değişken çıkarma yöntemi ile en son adımda elde edilen final model 47
Çizelge 3.5. Hosmer-Lemeshow test istatistiği ve önem kontrolü 48
Çizelge 3.6. Referans modelin sınıflandırma tablosu 49
Çizelge 3.7. Geliştirilmiş modelin sınıflandırma tablosu 49
Çizelge 3.8. Klasik performans ölçülerinin sonuçları 50
Çizelge 3.9. Tahmin modellerine ilişkin ROC analiz sonuçları 51
Çizelge 3.10. Referans ve Geliştirilmiş modelin tahmini olasılıklarının sınıflandırılması 51
Çizelge 3.11. Tekrar sınıflandırma tablosu 52
Çizelge 3.12. İlgilenilen olaya sahip bireylerin sınıflandırma tablosu 52
Çizelge 3.13. İlgilenilen olaya sahip olmayan bireylerin sınıflandırma tablosu 53

1. GİRİŞ

Uygulamalı Sağlık Bilimlerinde, çeşitli teşhis yöntemlerinden, laboratuvar testlerinin sonuçlarından ve tahmin modellerinden yararlanılarak hasta ve sağlıklı bireylerin ayrımının yapılması amaçlanır. Klinik çalışmalarda belirteçler, hastalığın tanısının konması, hastalığın süresi ve gelişimi gibi birçok amaç için kullanılır. Belirteçler, klinik kararların verilmesinde yardımcı olabilmesi ve risk hakkında bilgi verebilmesi açısından önemli araçlardır. Bu açıdan, yeni geliştirilmek istenen bir belirtecin kullanılabilirliğinin değerlendirilmesi için tanı testleri ya da tahmin modelleri sıklıkla kullanılan karar verme yöntemleridir. Son yıllarda bu yöntemlere gittikçe artan bir ilgi gözlenmekte ve bu yöntemlerin uygulamaları sağlık bilimleri literatüründe oldukça geniş yer bulmaktadır. Özellikle tanı testi ve tahmin modelleri ile ilgili yapılan çalışmaların büyük bir bölümü, bu yöntemlerin güvenilirliğinin araştırılmasına ve yöntemlerin karşılaştırılmasına ayrılmıştır.

Tahmin modelleri, ilgilenilen olayın ortaya çıkma olasılığını tahmin eden modellerdir (Zou ve ark., 2007). Bu modeller, sağlık alanında klinik karar verme sürecinde geleceğe yönelik kestirimlerde bulunması açısından çok önemli bir role sahiptir.

Bir tahmin modelinin performansı, modelin tanısal yeterliliği ya da olguları doğru olarak alt gruplara (ilgilenilen/ilgilenilmeyen olay vb.) ayırabilme kapasitesi ile tanımlanabilir. Diğer bir ifade ile yeterliliği yüksek bir model düşük maliyetli, daha az yanlış sonuç veren, daha ucuz ve teknik açıdan uygulanması daha kolay olan modeldir (Tomak ve Bek, 2010).

Özellikle zaman içerisinde risk faktörlerinin değişmesi ile araştırmacılar mevcut modelin risk tahminini geliştirmek için yeni belirteç arayışına girmişlerdir. Dolayısıyla son zamanlarda, mevcut bir modele yeni bir belirteç eklendiğinde tahmin modelinin performansının ve bu belirtecin kullanılabilirliğinin nasıl değerlendirileceğiyle ilgili çalışmalar önem kazanmıştır. Moons ve Harrel (2003), yapmış oldukları çalışmada belirteç ölçümü sonucu modelden elde edilen tahmini riskin bilinmesinin modelleme

açısından daha faydalı olduğunu belirtmişlerdir. Pepe ve ark (2003) ise, daha genel olarak belirtecin, gerçekte ilgilenilen olaya sahip olan ve olmayan bireyleri tanımlaması gerektiğini belirterek sınıflandırmanın daha önemli olduğunu vurgulamışlardır. Öte yandan yeni bir belirtecin kullanılabilirliğinin değerlendirilmesinde, istatistiksel önemliliğin (p-değerinin), belirtecin ne klinik açıdan yararlı olduğunu ne de model performansını geliştirdiği anlamına gelmeyeceği belirtilmiştir (Pencina ve ark., 2008). Çünkü yeterli sayıda örnek kullanıldığında, araştırılan belirteç ile ilgilenilen değişken arasında düşük ya da orta derecede ilişki istatistiksel anlamda önemli olabilmektedir.

Tahmin modelinin performansının değerlendirilmesinde kullanılan en önemli ölçü modelin olguları ayırt edilebilirliğidir. Bunun için birçok ölçü önerilmiş olsa da ROC eğrisi (Receiver Operating Characteristic Curve) altında kalan alan AUC (Area Under the Curve) en sık kullanılanıdır. Bu ölçü, ilgilenilen olaya sahip popülasyondan seçilen bir bireyin kriter değerinin, ilgilenilen olaya sahip olmayan popülasyondan seçilen bir bireyin kriter değerinden daha yüksek olması olasılığını verir. Ancak, yüksek ya da düşük risk grubuna giren bireylerin oranlarına ilişkin gerçek riskler hakkında bilgi vermez. Bir diğer ifade ile, ROC eğrisi risk tahminlerindeki değişime karşı hassas bir yöntem değildir. Aynı zamanda AUC, belirli bir eşik değeri kullanılarak hesaplanmak istenildiğinde de sınırlı kalır. Özellikle modele yeni bir belirteç eklendiğinde AUC da meydana gelen küçük bir artış, bu belirtecin klinik açıdan yararlılığını ya da kullanılabilirliğini açıklamada yeterli olmayabilir (Steyerberg ve ark., 2012). Bu problemlere karşı ise araştırmacılar risk tahmini ile risk sınıflandırılması arasında ayrım yapılması gerektiğini ve AUC ile istatistiksel önemliliğin yanı sıra yeni ölçülerin gerekli olduğunu ileri sürmüşlerdir (Pencina ve ark., 2011).

Pencina ve D'Agostino (2008), tahmin modellerinin performansının değerlendirilmesinde (modelin doğru sınıflandırma gücünün ölçülmesinde), AUC yöntemine alternatif olarak iki yeni performans ölçüsü önermişlerdir. Bunlar, NRI (Net Reclassification Improvement) ve IDI (Integrated Discrimation Improvement) ölçüleridir (Pencina ve ark., 2008). Bu ölçüler, yanıt değişkeninin ikili yapıda olduğu durumlar için geliştirilmiştir ve tahmin modellerinin performanslarının

değerlendirilmesi bölümünde detaylıca açıklanmıştır. Literatürde IDI yöntemi daha çok yer alırken, sağlık bilimleri ile ilgili dergilerde kullanımı daha kolay olması açısından NRI yöntemine daha sık rastlanmaktadır. Literatürde, 2008-2010 yılları arasında tekrar sınıflandırma ile ilgili yayınların 38'inde NRI ve 19'unda ise IDI ölçüleri kullanılmıştır (Tzoulaki ve ark., 2011).

Veteriner hekimlik alanında, Ateca ve ark (2015) yapmış oldukları çalışmada, 2006-2011 yılları arası kayıtlarından elde ettikleri düşük tansiyonlu hiperlaktatemisi olan ve olmayan 67 hasta köpekte prognozu belirlemede etkili olabilecek risk faktörlerini sağkalım analizi ile incelemişlerdir. Hiperlaktatemisi olmayan düşük tansiyonlu hasta köpeklerin hiperlaktatemisi olan köpeklere göre daha iyi bir iyileşme süreci geçirdiklerini ve daha uzun yaşam olasılığına sahip olduklarını belirtmişlerdir. Ayrıca, yaş, kilo, hastalığın ciddiyeti ve hastanede tedavi sürelerinin köpeklerin prognozunu belirlemede etkili risk faktörleri olmadığı sonucunu bulmuşlardır. Kan laktat konsantrasyonu (blood lactate concentration) belirtecinin ise köpeklerin prognozunu belirlemede kullanılabilir bir ölçü olduğunu bildirmişlerdir.

Oğuzoğlu ve ark (2013) yapmış oldukları çalışmada, 2007-2009 yılları arasında Ankara ilinde yaşayan evcil ve yabani 200 kediye ait verileri incelemişlerdir. 138 kedide sindirim sistemi, idrar yolu ve solunum sistemi ile ilgili kronik semptomlara rastlanmıştır. Lojistik regresyon yöntemi ile bu hastalıklara sebep olan virüsleri etkileyen risk faktörleri belirlenmeye çalışılmıştır. Sonuçlara göre yaş, cinsiyet, doğurganlık, bağışıklık sistemi, sağlık durumu ve yaşam biçimi gibi değişkenlerden yaş, sağlık durumu ve yaşam biçimi değişkenlerinin önemli risk faktörleri olduğunu bildirmişlerdir.

Akçay ve ark (2011) yapmış oldukları çalışmada, Türkiye'de broiler coccidiosis'inde etkili risk faktörlerini belirlemek amacıyla, 2006-2007 yılında Türkiye 'nin altı coğrafi bölgesinden seçilen toplam 817 çiftlikte bulunan 1110 kümesten dışkı numuneleri toplamışlardır. Coccidiosis ile ilişkili değişkenlerin belirlenmesinde lojistik regresyon analizi kullanılmıştır. Elde edilen sonuçlara göre, 1110 kümesten alınan dışkı örneklerinin analizi sonucunda kümeslerin %56.2'sinde klinik veya subklinik boyutta coccidiosis saptanmıştır. Ayrıca, mevsim, çiftlikteki

kümes sayısı, etlik pilicin yaşı, havalandırma sistemi, çatı izolasyonu, altlık materyali, iki yetiştirme dönemi arasındaki süre, altlık materyalinin güvenli bir uzaklığa atılması, kümeste giriş odasının olması, kümes çevresi veya içinde kemirgenlerin varlığı, kümes içi havanın durumu ve piliçlerin coccidiosis dışında bir salgın hastalık geçirmesi veya geçirmekte olması gibi değişkenlerin coccidiosis için önemli risk faktörü olduğunu bildirmişlerdir.

Kenney ve ark (2010) yapmış oldukları çalışmada, 2003 ve 2007 yılları arasında sindirim yolunda meydana gelen bir sızıntıdan dolayı septisemi olan ve cerrahi müdahale yapılan 114 köpeğe ait verileri incelemişlerdir. Çoklu organ fonksiyonu yetmezliğinin (MODS) ölüm oranı üzerine etkisini araştırmışlardır. Köpeklerin %78'i bir ya da daha fazla organ fonksiyon yetmezliğine, %50'sinin MODS'a sahip olduğu görülmüştür. Çoklu lojistik regresyon yöntemi ile ölüm oranını etkileyen risk faktörleri arasındaki ilişki incelenmiş ve model oluşturulmuştur. Elde edilen analiz sonuçlarına göre, ölüm oranı MODS olan köpeklerde %70 iken, MODS olmayan köpeklerde %25 olarak hesaplanmıştır. Dolayısıyla organ fonksiyon yetmezliği sayısı arttıkça ölüm oranı arttığı sonucuna ulaşmışlardır.

Çoban ve Tüzemen (2007), subklinik mastitis üzerine etkili faktörlerin belirlenmesi ile ilgili bir çalışma yapmışlardır. Çalışmada, iki farklı işletmede toplam 186 baş Esmer ve Siyah Alaca inekten 5208 gözlem almışlardır. Irk, laktasyon sırası, laktasyon dönemi, buzağılama mevsimi ve işletme gibi faktörlerin subklinik mastitis oluşumu ile ilişkilerinin belirlenmesinde lojistik regresyon analizi kullanılmıştır. Sonuç olarak, ineklerde, laktasyon sırası ve döneminin ilerlemesiyle subklinik mastitise yakalanma riskinin arttığı, subklinik mastitisin Esmer ırk ineklerde, Siyah Alacalara göre önemli derecede düşük oranda olduğu ve dolayısıyla ırkın ineklerde mastitisi belirlemede önemli bir risk faktörü olduğu belirlenmiştir.

Arslan ve ark (2003), Morkaraman ve Corriedale x Morkaraman F_1 kuzularının yarı entansif şartlarda büyüme özelliklerini tespit etmek amacıyla bir çalışma yapmışlardır. Araştırmanın materyalini 103 baş Morkaraman ve 45 baş Corriedale x Morkaraman melezi (F_1) kuzu oluşturmuştur. Genotip, ana yaşı, doğum tipi ve cinsiyet gibi faktörlerin kuzuların doğum ağırlığı ve farklı dönemlerdeki canlı ağırlıkları

üzerine etkileri, regresyon yöntemi ile analiz edilmiştir ve grup ortalamaları arasındaki farklılıkların önem kontrolü Duncan testi ile yapılmıştır. Elde edilen sonuçlara göre, genotip, doğum tipi, ana yaşı ve cinsiyetin doğum ağırlıkları üzerine etkisi ($p<0.01$) önemli; doğum tipinin 15 ve 30. gün canlı ağırlıklarına etkisi $p<0.05$ düzeyinde, 75 ve 90. gün canlı ağırlıklarına etkisi $p<0.01$ seviyesinde önemli, diğer büyüme dönemlerinde ise önemsiz bulunmuştur. Ana yaşı, genotip ve cinsiyetin tüm büyüme dönemlerinde canlı ağırlıklar üzerine etkilerinin ise önemsiz olduğu belirlenmiştir.

Bir diğer çalışma ise King ve ark (2001), ileri derecede hasta olan 624 köpeğin yoğun bakım ünitesinde tedaviye başlandıktan 30 gün sonrasında hayatta kalma olasılıklarını hesaplamışlardır. Yaşam olasılıklarının hesaplanmasında kullanılan yaşam tahmin indeksi (SPI) ve yeni geliştirilen SPI2'nın modele katkıları incelenmiştir. SPI'nın olduğu model için AUC 0.723 ve SPI2'nin dahil olduğu model için ise AUC 0.773 olarak hesaplanmıştır. Daha sonra SPI2'nin olduğu modelde bazı değişkenler çıkarıldıktan sonra modelin AUC değeri 0.76 olarak bulunmuştur. Sonuç olarak ise, SPI2 belirtecinin ileri derecede hasta olan köpeklerde yaşam olasılığının hesaplanmasında kullanılabilecek yeni bir ölçü olabileceği belirtilmiştir.

Akçapınar ve ark (2000) yapmış oldukları bir çalışmada Akkaraman, Sakız x Akkaraman F_1 ve Kıvırcık x Akkaraman F_1 kuzuların 90. güne kadar yaşama gücü ve 180. güne kadar büyüme özellikleri incelenmiştir. Kuzularda büyüme ile kuyruk çevresi ölçüsüne etki eden genotip, yıl, cinsiyet, doğum tipi ve doğum ağırlığı gibi çevresel faktörler en küçük kareler yöntemi ile incelenmiştir. İncelenen çevre faktörlerinin etkilerinin önem kontrolü varyans analizi ile yapılmış ve grupların ortalamaları arasındaki farkın önem kontrolü Duncan testi ile yapılmıştır. Elde edilen sonuçlara göre, doğum ağırlığına doğum yılı, cinsiyet ve doğum tipinin etkisi önemli ($p < 0.001$) olurken genotipin etkisi önemsizdir. 180. güne kadar büyümenin bütün dönemlerindeki canlı ağırlıklar üzerine genotip, doğum yılı, cinsiyet ve doğum ağırlığının etkisi önemli (sırasıyla, $p < 0.05$; $p< 0.01$; $p < 0.001$); doğum tipinin etkisi ise 105. güne kadar önemli ($p < 0.01$; $p < 0.001$), daha sonraki ağırlıklarda ise önemsiz olmuştur. 1997 yılında doğan kuzular 1996 yılında doğanlara, erkekler dişilere, tek doğanlar ikizlere göre büyüme bakımından daha üstün oldukları bildirilmiştir.

Klinik bir çalışmanın, uygun önemlilik derecesinde, yeterli bir güce sahip olması ve mümkün olan en kısa sürede tamamlanması istenilir. Klinik bir deneme ne kadar kısa sürede tamamlanmış olursa olsun, denemenin asıl başarısı, elde edilen verilerin analiz sonuçlarının güvenilirliği ile ölçülür. Veri yapısına uygun olmayan bir veri toplama ve analiz yönteminin seçilmesi, sonuçların güvenilirliğini etkileyebileceği gibi, klinik denemenin uygulanmasında kullanılan zaman ve kaynağın da boşa gitmesine neden olacaktır.

1.1. Türkiye'de Koyun Yetiştiriciliği

Türkiye'de halkın ihtiyacı olan hayvansal ürünlerin önemli bir kısmı koyunculuktan elde edilmektedir. Koyun yetiştiriciliği, et, süt, yapağı, kürk, deri, post ve gübre gibi verimleri ile insanların önemli ihtiyaçlarını karşılamaktadır. Koyun yetiştiriciliğinde et verimi en önemli verim özelliği olmuştur. Et veriminde önemli nokta kuzu eti üretiminin arttırılabilmesidir. Bunun için kuzuların kısa sürede büyüme ve gelişme göstermesi besi kabiliyeti, karkas verimi ve kalitesi yönünden gelişmiş olması istenmektedir (Ceyhan ve ark., 2007).

Koyun yetiştiriciliğinde, bölgenin coğrafi ve ekonomik şartları dikkate alınarak yetiştirme yönünün ve şeklinin iyi tespit edilmesi gerekir. Bu doğrultuda amaca uygun koyun ırkı veya tipinin isabetli seçimi yapılabilir. Yetiştirilecek olan koyun ırkının tespitinde bölgenin iklimi, mera ve yem imkanları ile koyun verimlerinin ekonomik durumu önemli faktörlerdir (Akçapınar ve Özbeyaz, 1999).

Türkiye'de koyun varlığının %95'ni seyrek otlu fakir meralara ve sert iklim koşullarına uyum sağlamış yağlı kuyruklu koyun ırkları oluşturmaktadır. Anadolu platosunun özellikle de Eskişehir'den Sivas'a kadar İç Anadolu Bölge'sinde sayı ve yayılış açısından en önde gelen ırk Akkaraman koyunlarıdır. Daha doğudaki illerde ise Morkaramanlar çoğunluktadır. Diğer yağlı kuyruklu koyun ırkları arasında Göller Bölgesinde Dağlıç, Güneydoğu Bölgesinde İvesi, Kars yöresinde Tuj koyunu sayılabilir. Kıyı kesimlerde ise yağsız ince kuyruklu koyunlar bulunmaktadır. Bu gruba giren ırklardan Kıvırcık Trakya ve Marmara Bölgesinde, Karakaya ise

Karadeniz kıyı şeridinde yaygındır (Günaydin, 2009). Bazı bölgeler de ise, yerel koyun ırkları yetiştirilmektedir. Bunlar, İzmir Çeşme'de ve bazı batı sahillerinde Sakız ırkı, Çanakkale civarında İmroz (Gökçeada) ırkı, Artvin civarında Hemşin, Amasya-Tokat civarında Herik, İzmir civarında Ödemiş koyunu gibi genotiplerdir (Akçapınar, 2000).

Türkiye'de yıllara göre çeşitli türlerin sayısı Çizelge 1.1'de verilmiştir. Türkiye koyun varlığı 2013 yılı verilerine göre 29.3 milyon baştır. Koyun sayısı 2001 yılında yaklaşık 27 milyon iken, 2009 yılına kadar sürekli azalmış, 2010 yılından itibaren tekrar artış göstermeye başlamıştır (Anonim, TUİK, 2015).

Çizelge 1.1. Türkiye'de çeşitli türlerde hayvan sayısı

	Sığır	Koyun	Keçi
2001	10 548 000	26 972 000	7 022 000
2002	9 803 498	25 173 706	6 780 094
2003	9 788 102	25 431 539	6 771 675
2004	10 069 346	25 201 155	6 609 937
2005	10 526 440	25 304 325	6 517 464
2006	10 871 364	25 616 912	6 643 294
2007	11 036 753	25 462 293	6 286 358
2008	10 859 942	23 974 591	5 593 561
2009	10 723 958	21 749 508	5 128 285
2010	11 369 800	23 089 691	6 293 233
2011	12 386 337	25 031 565	7 277 953
2012	13 914 912	27 425 233	8 357 286
2013	14 415 257	29 284 247	9 225 548

Türkiye'de koyun yetiştiriciliğinde hayvansal ürün ihtiyacının karşılanamamasının en büyük sebepleri koyunculuğun büyük miktarda düşük verimli yerli ırklara dayalı olması, erken kuzu kesimlerinin fazla olması ve hayvanların entansif besiye alınmaksızın, kalitesiz mera besisini takiben kesilmesidir (Sezenler ve ark., 2013). Bu nedenle üretimin arttırılması için sadece hayvan sayısını arttırmak çözüm olmamaktadır. Önemli olan hayvan başına verimin arttırılmasıdır. Çizelge 1.2'de kesilen koyun sayıları ve elde edilen et miktarı verilmiştir (Anonim, TÜİK, 2015). Son yıllarda, hayvancılık alanında işletmelerin kurulmasına, hayvan alımına, hayvansal üretimine ve yem bitkileri üretimine devlet tarafından önemli düzeyde teşvikler yapılmıştır. Bu durum hayvan sayısının artmasına katkı sağlamıştır.

Çizelge 1.2. Türkiye'de canlı ve kesilen koyun sayısı ile et miktarı

	Canlı koyun sayısı	Kesilen koyun sayısı (baş)	Et üretim miktarı (ton)
2001	26 972 000	4 747 268	85 661
2002	25 173 706	3 935 393	75 828
2003	25 431 539	3 554 078	63 006
2004	25 201 155	3 933 973	69 715
2005	25 304 325	4 145 343	73 743
2006	25 616 912	4 763 394	81 899
2007	25 462 293	6 428 866	117 524
2008	23 974 591	5 588 906	96 738
2009	21 749 508	3 997 348	74 633
2010	23 089 691	6 873 626	135 687
2011	25 031 565	5 479 546	107 076
2012	27 425 233	4 541 122	97 334
2013	29 284 247	4 958 226	102 943

Yerli koyun ırklarının verim özelliklerinin arttırılması için seleksiyon ve çevre ıslah çalışmalarının yanı sıra melezleme çalışmaları da yapılmaktadır (Arslan ve ark., 2003). Böylece, büyüme, besi ve karkas özellikleri bakımından üstün yeni koyun tipleri elde edilmeye çalışılmaktadır.

1.2. Kuzularda Büyüme ve Büyümeyi Etkileyen Faktörler

Koyun yetiştiriciliğinde ekonomik değeri olan en önemli özelliklerden biri büyümedir. Büyüme, canlının ergin canlı ağırlığına ulaşana kadar gösterdiği ağırlık artışıdır. Bütün hayvanlarda doğumdan sonra mümkün olduğu kadar kısa zaman içinde büyüme ve bu büyümeye göre en iyilerinin damızlığa ayrılması düşünüldüğü gibi, erken büyüyen hayvanlardan kısa zamanda çeşitli ürünlerin alınması da verimli hayvancılık için zorunludur (Mavili, 2012).

Koyunculukta et üretiminin ana kaynağı olan kuzu veriminin ve kuzularda büyüme performansının artırılması istenir (Yılmaz ve Altınel, 2003). Et üretiminin arttırılmasında koyun başına kuzu sayısı ile büyüme performansının arttırılması önemlidir (Demir, 2002).

Kuzuların büyüme ve gelişme döneminde gösterdikleri performansları koyun yetiştiriciliğinde çok önemlidir. Kuzularda büyüme ve gelişmeyi genotip, cinsiyet, yıl,

ana yaşı ve ananın süt verimi ile işletmelerin uyguladığı kuzu büyütme yöntemleri etkilemektedir (Sezenler ve ark., 2013). Kuzuların doğum ağırlığı ise daha sonraki dönemlerde canlı ağırlıklarını etkilemektedir. Dolayısıyla doğum ağırlığının düşük olmaması beklenir. Bu yüzden koyunların özellikle gebeliğin son döneminde beslenmelerine özen gösterilmelidir. Doğumdan sonra da koyunların beslenmesine dikkat edilmelidir. Bu, koyunların süt verimi ve buna bağlı olarak da kuzuların gelişimi bakımından önemlidir (Altın ve ark., 2003).

Bir canlının doğum ağırlığı ile çeşitli dönemlerdeki canlı ağırlıkları kalıtsal yapıya ve çevresel etkenlere bağlı olarak zaman içerisinde şekillenmektedir. Büyüme, canlı türüne, ırkına, çevre şartlarına, hayvanın cinsiyetine, yaşına ve ölçülen karakterin yapısına göre farklılık gösterir (Yıldız ve ark., 2009).

Büyüme ve gelişme yönünden ırklar ve ırk içindeki tipler arasında hatta aynı tip içindeki bireyler arasında da büyük farklılıklar olmaktadır. Bu farklılıkları meydana getiren önemli faktörlerden biri genetik yapıdır. Bu sebeple farklı koyun ırkları aynı çevre şartlarında büyütülseler bile büyüme ve diğer özellikler bakımından farklı sonuçlar vermektedir (Akçapınar ve Özbeyaz, 1999). Yapılan çok sayıda araştırmada (Ünal, 2002 ; Akçapınar ve ark., 2005 ; Ünal ve ark., 2006 ; Ünal ve ark., 2008 ; Yakan ve ark., 2012) çeşitli genotipler arasında kuzularda büyüme bakımından değişkenlikler olduğu, genellikle erkeklerin dişilerden, anası ergin yaşta olanların genç olanlardan, tek doğanların ikiz veya üçüz doğanlardan daha iyi performans gösterdikleri görülmüştür.

1.3. Tahmin Modelleri

Sağlık bilimlerinde istenen/istenmeyen olayın ayrımında, çeşitli teşhis yöntemleri ve laboratuvar testlerinin yanı sıra son yıllarda tahmin modellerinden de yararlanılmaya başlanılmıştır. Tahmin modelleri tanı koymanın yanı sıra, hastalığın izlenmesi, izlemsel verilerin kullanılarak ileriye yönelik risk tahmini yapılması ve hastaların uygun risk gruplarına (düşük, orta, yüksek vs.) göre sınıflandırılarak uygulanacak tedaviye karar verilmesi gibi amaçlar için de kullanılmaktadır.

Bu modeller, birden fazla belirteç bilgisini içeren istatistiksel modellerdir**.** Belirteç ise istenen/istenmeyen olayın riskini tahmin etmede kullanılan bir ölçüdür. Belirteçler ölçülebilir (örn. kolesterol seviyesi) ya da ölçülemeyen (genotip gibi) değişkenler olabilir. Dolayısıyla, belirteçler tahmin modelini oluşturan ve ilgilenilen olay ile ilişkili değişkenlerdir. En yaygın olarak kullanılan tahmin modelleri, lojistik regresyon modeli, Cox regresyon modeli ve sınıflandırma ağaçlarıdır. Bu modelleri oluşturan değişkenlerin bilgisi doğrultusunda her bir birey için risk tahmini elde edilir. Tahmin modelinin amacı, ilişkili risk kategorilerine göre bireyleri doğru sınıflandırmaktır. Buradan elde edilen risk bilgisi, veteriner hekime tedavi yöntemine karar vermede yol gösterici olacaktır. Ayrıca, bu model bireyleri istenmeyen olaya karşı korumaya yönelik önlemler alınmasında, yüksek risk grubuna sahip bireylerin durumunun seyrinin izlenmesinde ve tedaviye dahil edilecek bireylerin seçilmesinde politikalar geliştirmek için de kullanılır (Janes ve ark., 2008).

Yanıt değişkenin ikili yapıda ve zamandan bağımsız olduğu çalışmalarda kullanılan en yaygın tahmin modeli lojistik regresyon modelidir. Bu model biyoloji, tıp, ziraat, ekonomi, veteriner hekimliği, taşıma sahaları gibi çeşitli alanlarda kullanılmaktadır.

1.3.1. Lojistik Regresyon Modeli

Lojistik regresyon modeli, en az değişkeni kullanarak en iyi uyuma sahip olacak şekilde bağımlı değişken (açıklanan ya da yanıt değişken) ile bağımsız değişkenler kümesi (açıklayıcı değişken) arasındaki ilişkiyi tanımlayabilen ve biyolojik olarak kabul edilebilir bir model kurmayı amaçlar.

Basit ya da çoklu doğrusal regresyon problemlerinde, verilen bir bağımsız değişken değerine bağlı olarak bağımlı (yanıt) değişkenin ortalama değeri bulunur. Bu değer koşullu ortalama olarak adlandırılır ve $E(Y \mid x)$ ile gösterilir. Doğrusal regresyon analizinde, koşullu ortalamanın, x'in doğrusal bir denklemi olduğu varsayılır.

$$E(Y \mid x) = \beta_0 + \beta_1 x \qquad \text{(Denklem 1.1)}$$

Denklem 1.1'de x'in aralığı -∞ ile +∞ arasında değişir. Bu durumda $E(Y \mid x)$ mümkün tüm değerleri alır.

İkili lojistik regresyon analizinde, bağımlı değişken ikili yapıda olduğu için, Denklem 1.1'in sol tarafı 0-1 arasında sınırlı olasılık değerleri alır ve bu değerler -∞ ile +∞ sonsuz arasında değerler alabilen açıklayıcı değişkenlerle ilişkilendirilir. Açıklayıcı değişkenlerin sınırsız değerleri alması nedeniyle söz konusu eşitlik her zaman sağlanamamaktadır. Bu durumda, sonuç değeri olarak ifade edilen olasılık değeri çeşitli dönüşümlerle -∞ ile +∞ arasında tanımlı hale getirilir.

Lojistik regresyon modelinde, x bilindiğinde Y'nin koşullu ortalamasını göstermek için $\pi(x) = E(Y \mid x)$ ifadesi kullanılmaktadır. Bu ifadenin açık şekli

$$\pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}} \qquad \text{(Denklem 1.2)}$$

biçimindedir. Denklem 1.2'de $\pi(x)$'in dönüşümü lojit dönüşümüdür. Bu dönüşüm $\pi(x)$ cinsinden Denklem 1.3'de tanımlanır.

$$g(x) = \ln\left[\frac{\pi(x)}{1 - \pi(x)}\right] = \beta_0 + \beta_1 x \qquad \text{(Denklem 1.3)}$$

Bu dönüşümün önemi $g(x)$'in, doğrusal regresyon modelinde istenen çoğu özelliği taşımasıdır. Lojit $g(x)$ parametreleri bakımından doğrusal ve x'in aldığı değerlere bağlı olarak -∞ ile +∞ arasında değişmektedir.

Doğrusal regresyon modelinde, sonuç değişkeninin bir gözlemi $y = E(Y \mid x) + \varepsilon$ şeklinde gösterilir. ε değeri hata terimi olarak adlandırılır ve gözlemin koşullu olasılıktan ne kadar saptığını gösterir. ε'un sıfır ortalama ve bağımsız değişkenin her düzeyinde sabit olacak şekilde σ^2 varyansla normal dağılım göstereceği çok genel bir

varsayımdır. Ancak iki düzeyli sonuç değişkenleri için durum böyle değildir. Bu tür durumlarda, x verildiğinde sonuç değişkenin değeri $y=\pi(x)+\varepsilon$ ile gösterilir. ε'un mümkün olan iki değerden başka bir değer alamayacağı varsayılır. Eğer y=1 ise $\pi(x)$ olasılıkla ε=$1-\pi(x)$ değerini alır ve eğer y=0 ise $1-\pi(x)$ olasılıkla ε= - $\pi(x)$ değerini alır. Böylece ε, sıfır ortalama ve $\pi(x)[1-\pi(x)]$'a eşit varyanslı bir dağılım gösterir. ε'a ilişkin ortalama ve varyans değerlerinin hesaplanması Denklem 1.4'de gösterilmiştir (Elhan, 1997).

$$
\begin{aligned}
&E(\varepsilon)=\left[(1-\pi(x))\pi(x)\right]+\left[(-\pi(x))(1-\pi(x))\right]=0\\
&V(\varepsilon)=E(\varepsilon^2)-\left[E(\varepsilon)\right]^2=E(\varepsilon^2)\\
&=\left[(1-\pi(x))^2\pi(x)\right]+\left[(-\pi(x))^2(1-\pi(x))\right]\\
&=\left[(1-2\pi(x)+\pi(x)^2)\pi(x)\right]\left[\pi(x)^2(1-\pi(x))\right]\\
&=\left[\pi(x)-\pi(x)^2\right]=\pi(x)\left[1-\pi(x)\right]
\end{aligned}
\qquad \text{(Denklem 1.4)}
$$

Lojistik regresyonda bilinmeyen parametreleri kestirmek için en sık kullanılan metot en çok olabilirlik yöntemidir. Bu yöntem genel anlamda, gözlenen veri kümesini elde etmenin olasılığını maksimum yapan bilinmeyen parametrelerin değerini verir. Bu yöntemin uygulanması için ilk olarak olabilirlik fonksiyonunun tanımlanması gerekir. Bu fonksiyon, gözlenen verinin olasılığını, bilinmeyen parametreler cinsinden bir fonksiyon olarak verir. Bu parametrelerin en çok olabilirlik tahmin edicileri bu fonksiyonun değerini en büyük yapan değerleri bulacak şekilde seçilir. Dolayısıyla elde edilen tahmin ediciler, gözlenen verilere çok yakın değerlere sahip olurlar.

Eğer Y, 0 ve 1 olarak kodlandıysa, ifadesi $\pi(x)$ verildiğinde Y'nin 1'e eşit olma koşullu olasılığını verir $[\pi(x)]=P(Y=1|x)]$. $[1-\pi(x)]$ değeri verilen herhangi bir x için Y'nin 0'a eşit olma koşullu olasılığını göstermektedir $[1-\pi(x)=P(Y=0|x)]$.

(x_i, y_i) çiftinin, y_i=1 olduğu zaman olabilirlik fonksiyonuna katkısı $\pi(x_i)$ iken, y_i =0 olduğu zaman olabilirlik fonksiyonuna katkısı $1-\pi(x_i)$ kadardır. (x_i, y_i) çiftinin olabilirlik fonksiyonuna katkısı Denklem 1.5'deki gibi gösterilir.

$$\zeta(x_i) = \pi(x_i)^{y_i}[1-\pi(x_i)]^{1-y_i} \qquad \text{(Denklem 1.5)}$$

Gözlemlerin birbirinden bağımsız oldukları varsayıldığı için, olabilirlik fonksiyonu Denklem 1.5'deki terimlerin çarpılması ile elde edilir.

$$l(\beta) = \prod_{i=1}^{n} \zeta(x_i) \qquad \text{(Denklem 1.6)}$$

En çok olabilirlik ilkesi, Denklem 1.6'yı maksimum yapan tahminin kullanılması gerektiğini vurgular. Ancak, matematiksel olarak Denklem 1.6'daki eşitliğin logaritması ile çalışmak daha kolay olacağından log olabilirlik fonksiyonu ile tanımlanmıştır.

$$L(\beta) = \ln\left[l(\beta)\right] = \sum_{i=1}^{n}\left\{y_i \ln\left[\pi(x_i)\right] + (1-y_i)\ln\left[1-\pi(x_i)\right]\right\} \qquad \text{(Denklem 1.7)}$$

$L(\beta)$ 'yı maksimum yapan β değerini bulabilmek için, $L(\beta)$ 'nın β_0 ve β_1'e göre birinci türevi alındıktan sonra elde edilen ifadeler 0'a eşitlenir ve aşağıdaki denklemler elde edilir.

$$\begin{aligned} &\sum_{i=1}^{n}\left[y_i - \pi(x_i)\right] = 0 \\ &\sum_{i=1}^{n} x_i\left[y_i - \pi(x_i)\right] = 0 \end{aligned} \qquad \text{(Denklem 1.8)}$$

1.3.1.1. Parametrelerin Önem Testi

Lojistik regresyonda bağımsız değişkenlere ait katsayılar tahmin edildikten sonra, olabilirlik fonksiyonları kullanılarak gözlenen değerler ile tahmini değerler Denklem 1.9 ile karşılaştırılır.

$$D = -2\ln\left[\frac{\text{Mevcut modelin olabilirliği}}{\text{Doymuş modelin olabilirliği}}\right] \qquad \text{(Denklem 1.9)}$$

Parantez içindeki ifade olabilirlik oranını göstermektedir. İfadenin başındaki (-2 ln) katsayısı matematikseldir. Dolayısıyla bu değer hipotez testi amacıyla kullanılabilir. Denklem 1.7 kullanılarak Denklem 1.9 aşağıdaki gibi yazılabilir.

$$D = -2\sum_{i=1}^{n}\left[y_i \ln\left(\frac{\hat{\pi}_i}{y_i}\right) + (1-y_i)\ln\left(\frac{1-\hat{\pi}_i}{1-y_i}\right)\right] \qquad \text{(Denklem 1.10)}$$

D istatistiği, bazı yazarlar tarafından sapma (deviance) olarak adlandırılır. Lojistik regresyondaki sapma, doğrusal regresyondaki artık kareler ile aynı role sahiptir. Bağımsız bir değişkenin önemini değerlendirmek için denklemde bağımsız değişkenin olduğu ve olmadığı durumlardaki D değerleri, G istatistiği kullanılarak karşılaştırılır.

$$G = D(\text{değişkensiz model için}) - D(\text{değişkenli model için}) \qquad \text{(Denklem 1.11)}$$

Bu istatistik, doğrusal regresyonda kullanılan F testindeki pay kısmı ile aynı role sahiptir. G istatistiğini hesaplamak için D değerlerinin her ikisi için de doymuş modelin olabilirlikleri ortak olduğundan, G istatistiği Denklem 1.12 gibi ifade edilir.

$$G = -2\ln\left[\frac{\text{Değişkensiz modelin olabilirliği}}{\text{Değişkenli modelin olabilirliği}}\right] \qquad \text{(Denklem 1.12)}$$

Tüm değişkenleri içeren model ile kestirilen modele ilişkin olabilirlik oran değerlerinin farkına dayanan ölçütlerin Ki-Kare dağılımı göstereceği düşünülerek modelin geçerliliği sınanmaktadır. Bu yöntemle modele girecek açıklayıcı değişkenlere karar verilmektedir.

Değişkenin önem testi için Wald ve Score testleri de kullanılmaktadır. Bu testler için geçerli varsayımlar olabilirlik oran testi ile aynıdır. Ancak, Wald testi değerleri için standart hataların tahminini artırması, Score testi ise hazır paket programlarda sıklıkla bulunmayışı gibi dezavantajlarından dolayı, lojistik regresyonda bir değişkenin önem testi için olabilirlik önem testinin kullanılması önerilir (Hosmer ve Lemeshow, 2000; Akçay, 2009).

1.3.1.2. Çoklu Lojistik Regresyon Modeli

Lojistik regresyon modelinde tek bağımsız değişkenden fazla değişken olması durumuna çok değişkenli durum denir. Çok değişkenli lojistik regresyon modeli için temel düşünce, katsayıların tahmini ve bu katsayıların anlamlılıklarının test edilmesidir. Bu da tek bağımsız değişkenli modele benzer şekilde yapılır.

$x' = (x_1, x_2, ..., x_p)$ vektörü ile gösterilen p tane bağımsız değişken kümesi olsun ve bu değişkenlerin en az aralık ölçekli (sürekli) olduğu varsayılsın. Sonuç değişkeninin Y=1 şeklinde olduğu durumlarda koşullu olasılık $P(Y=1|x) = \pi(x)$ şeklinde olur. Çoklu lojistik regresyon modelinin lojiti aşağıdaki denklem ile verilmiştir.

$$g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_p x_p \qquad \text{(Denklem 1.13)}$$

Denklem 1.13'den sonuç değişkeninin koşullu olasılığı,

$$\pi(x) = \frac{e^{g(x)}}{1 + e^{g(x)}} \qquad \text{(Denklem 1.14)}$$

ile ifade edilir (Hosmer ve Lemeshow, 2000).

1.3.1.3. Çoklu Lojistik Regresyon Modelinin Kurulması

Birbirinden bağımsız *n* tane (x_i, y_i) , i=1,2,....,n gözlem çiftinin olduğu varsayılsın. Çoklu lojistik regresyon modeli kurulurken, tek değişkenli modelde olduğu gibi $\beta' = \left(\beta_1, \beta_2, ..., \beta_p\right)$ tahmin vektörünün elde edilmesi gerekir. Çok değişkenli durumda, tek değişkenli durumda olduğu gibi tahmin metodu en çok olabilirlik olacaktır. Olabilirlik fonksiyonu Denklem 1.6'da verilen tek değişkenli durum ile hemen hemen aynıdır. Tek fark $\pi(x)$ 'in Denklem 1.14'de verildiği gibi çok değişkenli durum için tanımlanmasıdır. Log-olabilirlik fonksiyonun $p+1$ katsayıya göre türevi alınarak, p+1 tane olabilirlik denklemi,

$$\sum_{i=1}^{n}\left[y_i-\pi(x_i)\right]=0 \qquad \text{(Denklem 1.15)}$$

$$\sum_{i=1}^{n}x_{ij}\left[y_i-\pi(x_i)\right]=0 \text{ , j=1,2,...,p} \qquad \text{(Denklem 1.16)}$$

şeklinde elde edilir.

$\hat{\beta}$, bu denklemlerin çözümünü gösteren tahmin vektörü olmak üzere, çoklu lojistik regresyon modeli için tahmin edilen değerler $\hat{\pi}(x_i)$ 'dir. Denklem 1.14'deki ifadenin değeri $\hat{\beta}$ ve x_i'yi kullanarak hesaplanmıştır. Tahmin edilen katsayıların varyans ve kovaryanslarının tahmini için en çok olabilirlik yöntemi kullanılır. Tahminler log olabilirlik fonksiyonlarının ikinci derecede kısmi türevlerinden oluşan matris yardımıyla elde edilir. Kısımi türev $j,u=0,1,2,...p$, $\pi=\pi(\mathrm{x_i})$ olmak üzere aşağıdaki gibidir.

$$\frac{\partial^2 L(\beta)}{\partial\beta_j^2}=-\sum_{i=1}^{n}x_{ij}^2\pi_i(1-\pi_i) \qquad \text{(Denklem 1.17)}$$

$$\frac{\partial^2 L(\beta)}{\partial\beta_j\partial\beta_u}=-\sum_{i=1}^{n}x_{ij}x_{iu}\pi_i(1-\pi_i) \qquad \text{(Denklem 1.18)}$$

Denklem 1.17 ve 1.18'de verilen terimlerin negatiflerini kapsayan $(p+1)x(p+1)$ boyutunda matris $I(\beta)$ ile gösterilir ve "bilgi (information) matrisi" olarak adlandırılır. $\hat{\beta}$'ların varyans ve kovaryansları bu matrisin tersinden elde edilir $\sum\beta=I^{-1}(\beta)$. Çok özel durumlar dışında bu matrisin açık şeklini yazmak mümkün değildir. $\hat{\beta}_j$'nin varyansı $\sigma^2(\beta_j)$ ile gösterilir ki bu, matrisin j'nci köşegen elemanıdır. Bu matrisin köşegen olmayan elemanlarından $\sigma(\beta_j,\beta_u)$ ise, $\hat{\beta}_j$ ve $\hat{\beta}_u$'nun

kovaryansını verir. Varyans ve kovaryansların tahminleri $\sum\left(\hat{\beta}\right)$ ile gösterilmiştir ve $\sum\left(\hat{\beta}\right)$ değeri $\hat{\beta}$ kullanılarak hesaplanmıştır. $\hat{\beta}$ 'ların standart hataları,

$$S\hat{E}(\hat{\beta}_j)=\left[\hat{\sigma}^2(\hat{\beta}_j)\right]^{1/2},\ j=0,1,2,...,p \qquad \text{(Denklem 1.19)}$$

şeklinde elde edilir.

Denklem 1.19 katsayıları test etmek ve güven aralıklarını tahmin etmek için kullanılır. Model kurulurken ve uyum iyiliği tartışırken bilgi (Information) matrisinden yararlanılır.

$$I(\hat{\beta})=X'VX \qquad \text{(Denklem 1.20)}$$

X matrisi $\left[n \text{ x } (p+1)\right]$ boyutunda bir matristir ve her bir denek verisini içermektedir.

$$X=\begin{bmatrix} 1 & x_{11} & ... & x_{1p} \\ 1 & x_{21} & ... & x_{2p} \\ . & . & . & . \\ . & . & . & . \\ 1 & x_{n1} & ... & x_{np} \end{bmatrix}$$

V matrisi $\left[n \text{ x } n\right]$ boyutunda genel elemanı $\hat{\pi}_i(1-\hat{\pi}_i),\ i=1,2,...,n$ olan diyagonal bir matristir (Hosmer ve Lemeshow, 2000; Elhan, 1997).

$$V=\begin{bmatrix} \hat{\pi}_1(1-\hat{\pi}_1) & 0 & ... & 0 \\ 0 & \hat{\pi}_2(1-\hat{\pi}_2) & ... & 0 \\ . & . & . & . \\ . & . & . & . \\ 0 & . & . & \hat{\pi}_n(1-\hat{\pi}_n) \end{bmatrix}$$

1.3.1.4. Modelin Önemlilik Testi

Çok değişkenli lojistik regresyon modeli kurulduktan sonra modelin önemliliğinin test edilmesi gerekmektedir. Bu aşamada yine tek değişkenli lojistik regresyon modelinde olduğu gibi olabilirlik oran testi kullanılmaktadır. Bu test, (1.10) ve (1.12)'nolu denklemlerde verildiği üzere G istatistiği temeline bağlıdır. Modelde bulunan p tane bağımsız değişkenin eğim katsayılarının sıfıra eşit olması hipotezi altında, G istatistiği p serbestlik derecesi ile Ki-Kare dağılımına sahiptir. Eğer, "katsayıların hepsi sıfıra eşittir" H_0 hipotezi reddedilirse, en az bir katsayının sıfırdan farklı olduğu sonucuna varılır. Ancak, bu sonuca varmadan önce değişkenler tek tek Wald test istatistiği ile test edilebilir.

$$W_j = \frac{\hat{\beta}_j}{\hat{SE}(\hat{\beta}_j)} \qquad \text{(Denklem 1.21)}$$

Bu testin amacı, modeldeki değişkenlerin önemli mi yoksa önemsiz mi olduğunu belirlemektir.

Model oluştururken önemli olan nokta, en iyi uyuma sahip modeli en az parametre ile belirlemektir. Bunun için sadece önemli olduğu düşünülen değişkenleri modele dahil ederek yeni bir model kurmak (azaltılmış model) ve bu model ile tüm değişkenlerin dahil olduğu tam modelin olabilirlik oranlarını karşılaştırmak gerekir.

"Modelden çıkarılan değişkenlerin katsayıları sıfıra eşittir" hipotezi altında G istatistiği $(v_1 - v_2)$ serbestlik derecesiyle Ki-Kare dağılımı göstermektedir. Burada v_2=(tam modeldeki değişken sayısı+1), v_1=(azaltılmış modeldeki değişken sayısı+1)'dir.

G = -2[(azaltılmış model için log-olabilirlik)-(tam model için log-olabilirlik)]

G istatistiği için, $(v_1 - v_2)$ serbestlik derecesinde yanılma olasılığı (p) bulunabilir. Eğer p değeri 0.05 'den büyükse değişken sayısı azaltılmış modelin tam

model kadar iyi olduğu söylenir. Bunun yanında modele girecek değişkenlere karar verirken yalnızca istatistiksel önemliliğe bakmak yeterli olmayıp, aynı zamanda birçok etkene de bakmak gerekmektedir. Örneğin, kategorik olarak ölçeklendirilmiş bir bağımsız değişkenin modelden çıkarıldığı ya da modele girdiği zaman, onun bütün tasarım değişkenleri de modelden çıkarılmalıdır ya da modele dahil edilmelidir. Eğer kategorik bir değişkenin k düzeyi varsa, bu değişkenin modelden çıkarılmasının olabilirlik oran testi için serbestlik derecesine katkısı k-1 olacaktır (Elhan, 1997).

1.3.1.5. Modelin Uyumunun Belirlenmesi

Bir modelin uyum iyiliğine karar vermek isteniyorsa, öncelikle model uyumu ile ne denilmek istenildiği hakkında bazı fikirlere sahip olunmalıdır. Sonuç değişkeninin gözlenen değerleri y vektörü ile gösterilsin ve $y' = (y_1, y_2, ..., y_n)$ olsun. Model tarafından tahmin edilen değerler $\hat{y}$ ile gösterilsin ve $\hat{y}' = (\hat{y}_1, \hat{y}_2, ..., \hat{y}_n)$ olsun. Eğer y ve $\hat{y}$ arasındaki uzaklık özet ölçüleri küçükse, ve her bir (y, $\hat{y}$) ikilisinin bu özet ölçülere katkısı sistematik değil ve modelin hata yapısına göre küçükse modelin uyumlu olduğuna karar verilir. Böylece, uygun bir modele tamamen karar verebilmek için hem y ve $\hat{y}$ arasındaki uzaklığın özet ölçüsünün hesaplanması hem de bu ölçülerin her bir parçasının teker teker incelenmesi gerekir (Hosmer ve Lemeshow, 2000; Atabey, 2010).

Uyum iyiliği testinde Hosmer-Lemeshow istatistiği, Ki-Kare istatistiği, -2logL istatistiği, Pearson Ki-Kare istatistiği ve blok Ki-Kare istatistiği gibi kullanılan farklı istatistikler vardır. Bu çalışmada, uygulama kısmında yararlanılacak olan Hosmer-Lemeshow istatistiği üzerinde ayrıntılı olarak durulmuştur.

1.3.1.6. Hosmer-Lemeshow İstatistiği

Hosmer ve Lemeshow, tahmin edilen olasılık değerlerinin gruplandırılmasını önermişlerdir (Hosmer ve Lemeshow, 1982). $J = n$ olduğu durumlarda veri matrisindeki n tane sütun, n tane tahmin edilen olasılık değerine karşılık gelmektedir. Aynı bağımsız değişken değerine sahip olan bireylerin olduğu durumda bunlardan sadece bir tanesi seçilir ve toplam birey sayısı J ile ifade edilir. Burada gruplama

yapmaktaki amaç, mevcut dağılımı Ki-Kare dağılımına yaklaştırarak anlamlı ve yorumlanabilir bir model elde etmektir. Gruplama;
(a) tahmin edilen olasılıkların yüzdesi dikkate alınarak,
(b) tahmin edilen olasılıkların sabit değerleri dikkate alınarak
iki farklı şekilde yapılabilir.

Birinci metot gözlenen ve tahmin edilen beklenen frekansları karşılaştıran istatistik aşağıdaki gibidir.

$$\hat{C}_g^* = \sum_{k=0}^{g} \sum_{l=1}^{10} \frac{(o_{kl} - e_{kl})^2}{e_{kl}} \qquad \text{(Denklem 1.22)}$$

g grup sayısını, o gözlenen değeri, e beklenen değeri, l ise risk grubunu temsil etmektedir. Bu metotta çoğunlukla 10'lu risk grubu kullanılmaktadır. Burada 10'dan daha az ya da daha çok grup kullanılabilir. Hosmer ve Lemeshow tarafından verilen teorik gelişme yalnızca $g > (p+1)$ olmasını gerektirir. Burada g grup sayısı, p ise değişken sayısıdır. $\hat{C}_g^*$ istatistiği $(g-2)$ serbestlik derecesiyle Ki-Kare dağılımına sahiptir (Hosmer ve Lemeshow, 2000; Elhan, 1997).

Çizelge 1.3. İstenen/istenmeyen olaya sahip ya da sahip olmayan denekler için her bir onlu risk grubunda gözlenen frekanslar

	Onlu risk grupları					
Durum	1	2	3	.....	10	Toplam
Olay var (Y=1)	o_{11}	o_{12}	o_{13}	.....	$o_{1,10}$	n_1
Olay yok (Y=0)	o_{01}	o_{02}	o_{03}	.....	$o_{0,10}$	n_0
Toplam	$n/10$	$n/10$	$n/10$		$n/10$	n

Onlu risk metodu epidemiyolojik literatürde çok sıklıkla kullanılmaktadır. Bu metotta, ilk grupta en küçük kestirilen olasılığa sahip n/10 denek bulunurken, en son grupta en büyük kestirilen olasılığa sahip n/10 denek bulunur (Çizelge 1.3). Bu yaklaşımın avantajı her grubun yeterli sayıda (n/10) bireye sahip olması iken, dezavantajı ise kestirilen hastalık olasılığının gerçek değerlerinin gözardı edilmesidir.

Hosmer-Lemeshow, alternatif bir yöntem olarak ikinci metottaki sabit kesim noktaları üzerine kurulu grupları oluşturmayı önermişlerdir (Elhan, 1997). k/10, k=1,2,…..,9 değerleriyle tanımlanan kesim noktalarıyla $g = 10$ grup oluşturulur ve gruplar kestirilen olasılıkları ardı ardına gelen kesim noktaları arasına düşen denekleri kapsar. Örneğin, birinci grup kestirilen olasılığı 0,1'e eşit ya da daha küçük olan bütün bireyleri kapsarken, onuncu grup kestirilen olasılığı 0,9'dan büyük olan denekleri kapsar. Çizelge 1.4'de gösterilen bu kesim noktaları gözlenen değerlerin tablosunu oluşturur.

Çizelge 1.4. İstenen/istenemeyen olaya sahip ya da sahip olmayan denekler için sabit kesim noktasına dayanarak oluşturulan gözlenen frekanslar

Durum	(0-0.10)	(0.11-0.20)	.....	(0.91-1.00)	Toplam
Olay var (Y=1)	o'_{11}	o'_{12}	.....	$o'_{1,10}$	n_1
Olay yok (Y=0)	o'_{01}	o'_{02}	.....	$o'_{0,10}$	n_0
Toplam	m_1	m_2		m_{10}	n

Gözlenen değer o' ve kestirilen beklenen değer ise e' Denklem 1.23'te tanımlanmıştır. Kısaca, H_g^* istatistiği aşağıdaki gibi elde edilir.

$$\hat{H}_g^* = \sum_{k=0}^{1} \sum_{l=1}^{10} \frac{(o'_{kl} - e'_{kl})}{e'_{kl}} \qquad \text{(Denklem 1.23)}$$

Hosmer ve Lemeshow tarafından yapılan bilgisayar benzetimleri sonucunda $\hat{H}_g^*$ 'nin, $\hat{C}_g^*$'den daha güçlü olduğu gösterilmiştir ve böylece $\hat{H}_g^*$ tercih edilen istatistik olabileceği sonucuna varılmıştır (Elhan, 1997).

1.4. Tahmin Modellerinin Performanslarının Değerlendirilmesi

Son yıllarda, özellikle epidemiyolojik ve klinik çalışmalarda tahmin modellerinin performanslarının değerlendirilmesi, özellikle de tahmin modeline yeni eklenen bir belirtecin kullanılabilirliğinin değerlendirilmesi ile ilgili çalışmaların sayısı artmıştır (Cook Nr, 2006; Lauer ve ark., 2007; Van Der Steeg Wa, 2007; Tice ve ark., 2008; Parikh Ni, 2008). Çalışmalarda belirteç, istenen/istenmeyen durumun tanısının konması, izlenmesi ve risklerin değerlendirilmesi gibi birçok amaç için kullanılır. Yeni bir belirteç geliştirilmesi, ilgilenilen duruma ilişkin kararların verilmesinde yardımcı olabilmesi ve risk hakkında bilgi verebilmesi açısından araştırmalar için çok önemlidir. Dolayısıyla, araştırmacılar geliştirilen yeni belirtecin model performansına katkısını ve kullanılabilirliğini değerlendirmek isterler (Janes ve ark., 2008). Yeni belirtecin kullanılabilirliğinin değerlendirilmesinde öncelikle ilgilenilen popülasyonu ve olayı tanımlamak gerektiğini, daha sonra etkisi araştırılmak istenen belirteç ile bilinen değişkenleri birlikte incelenmesi ve uygun tahmin modelinin seçilerek, yeni belirtecin model performansına katkısının değerlendirilmesi gerektiği belirtilmiştir (Pencina ve ark., 2008).

Özellikle yanıt değişkeninin ikili yapıda olduğu durumlarda model performansının ve yeni bir belirtecin modele katkısının değerlendirilmesinde birçok istatistiksel ölçü kullanılmaktadır. İstatistiksel tahmin modellerinde, tahmin değerlerinin gerçek değerlere ne kadar iyi uyduğunu, gözlenen ve tahmini değerlerin ne denli tutarlı olduğunun bilinmesi model performansının değerlendirilmesinde temel oluşturur. Bu amaçla kullanılan genel performans ölçüleri ise, açıklayıcılık katsayısı (R^2) ve Brier Skorudur. Aynı zamanda performans değerlendirilirken gözlenen gerçek olasılıklar ile modelden elde edilen tahmin olasılıkları arasındaki uyum kalibrasyon (calibration) ve modelin istenen veya istenmeyen olaya sahip bireyleri ayırma gücünün ne kadar iyi olduğu ayrımsama (discrimination) da ölçülmesi gereken iki önemli ölçüdür. Kalibrasyon, "uyum iyliği" olarak ta adlandırılan test istatistikleri ile ölçülürken, ayrımsama ise, duyarlılık, seçicilik ve ROC eğrisi altında kalan alan ile ölçülür. Bu ölçülerden, ROC eğrisi altında kalan alan, özellikle bir belirtecin model

performansına katkısının değerlendirilmesinde en yaygın kullanıma sahip olan yöntemdir (Cook, 2007).

Ancak son yıllarda sağlık bilimlerinde yapılan çalışmalarda bu yöntemin bazı kısıtlamaları olduğu ve yeni eklenen bir belirtecin kullanılabilirliğinin değerlendirilmesinde yeteri kadar açıklayıcı olmadığı tartışılmaktadır (Cui, 2009). Özellikle ROC eğrisinin, risk tahminlerindeki değişime karşı hassas bir yöntem olmadığı belirtilmektedir (Cook ve Paynter, 2011). Bilinen değişkenleri içeren ve iyi ayırt edebilen bir modele etkisi araştırılmak istenen yeni bir belirteç eklendiğinde, AUC değerini arttırması ve bu artışın da anlamlı olması beklenir. Fakat bazı çalışmalarda bu artışın çok küçük olduğu ve anlamlı olmadığı görülmüştür. Pencina ve ark (2008) yapmış oldukları bir çalışmada, 1948-1958 yılları arasında, Massachusetts eyaletinin Framingham şehrinde yürütülmüş olan Framingham Kalp çalışmasından toplanan verileri incelemişlerdir. Bağımlı değişken olarak koroner arter hastalığı geçirme durumu alınmıştır. Bireylerin; cinsiyet, yaş, diyabet varlığı, sistolik kan basıncı değişkenleri ile oluşturulan referans modele, HDL (High-density lipoprotein) değişkeni ekleyerek yeni model elde etmişlerdir. Bu iki modelin AUC değerleri karşılaştırıldığında modele HDL değişkenin eklenmesi ile AUC 0.762 ‘den 0.774 değerine küçük bir artış göstermiştir ve bu artış istatistiksel olarak anlamlı bulunmamıştır (p=0.092). Fakat HDL değişkeni, tahmin modeli elde edilirken istatistiksel olarak anlamlı bulunmuş ve modelde yer almıştır. Hatta modele HDL değişkeninin eklenmesi ile AIC (Akaike Information Criterion)’nin 2779 değerinden 2762 değerine düştüğü hesaplanmıştır. Tekrar sınıflandırma tablolarını kullanarak risk kategorileri arasındaki değişim incelendiğinde ise, HDL değişkenin modele eklenmesi ile hasta bireylerin sağlıklı bireylere göre bir üst kategoriye geçme olasılığı %12 daha fazla olduğu bulunmuştur. Yeni eklenen bir belirtecin modele katkısı AUC’ye göre çok küçük ve önemsiz görünürken, tekrar sınıflandırma yapıldığında aynı değişkenin katkısının çok daha önemli olabileceği görülmüştür.

Bu tip problemlerde ROC yöntemine alternatif olarak, Cook ve ark (2006) yeni belirteci içermeyen ve içeren her iki tahmin modelinden elde edilen olasılıkları risk kategorilerine göre karşılaştıran tekrar sınıflandırma tablosunu, Pencina ve ark (2008)

ise bu tabloya ilişkin iki yeni performans ölçüsü önermişlerdir. Bu ölçüler, riskleri klinik kategorilere göre tabakalara ayırır ve modelin bireyleri yüksek ve düşük risk kategorilerine daha doğru ayırt edebilmesini sağlar (Pepe ve ark., 2008). Bunlar, NRI ve IDI ölçüleridir. NRI, istenen veya istenmeyen olaya sahip bireylerin risk kategorileri arasındaki hareketliliği ölçerken, IDI modelin istenen veya istenmeyen olaya sahip bireyleri birbirinden ne kadar ayırt edebildiğini ölçer.

Çizelge 1.5' de klasik ve literatürde yeni önerilen performans değerlendirme ölçüleri ve özellikleri özet haliyle verilmiştir. Bu ölçüler, ilerleyen bölümde ayrıntılı olarak anlatılacaktır.

Çizelge 1.5. Klasik ve Yeni performans değerlendirme ölçüleri

		Ölçüler	Görsellik	Özellikler
Klasik Performans Değerlendirme Ölçüleri	Genel Performans	R^2		Gözlenen ve tahmini değerler arasındaki farkı ölçer.
		Birer Skoru		
	Ayrımsama (Discrimination)	AUC	ROC eğrisi	Modelin ilgilenilen olaya sahip olan ve olmayan kişilerden rastgele seçilmiş iki bireyi doğru tanımlayabilme olasılığı olarak ifade edilir.
		Ayırt Etme Eğimi	Kutu-Çizgi grafiği	İlgilenilen olaya sahip olan ve olmayan bireylerin, tahmini olasılıklarının ortalamaları arasındaki farktır.
	Kalibrasyon (Calibration)	Kalibrasyon kesim noktası (Calibration-in-the-large)	Kalibrasyon ya da doğrulama grafiği	Dışsal doğrulama (external validation) için gözlenen ve tahmini değerleri karşılaştırır.
		Kalibrasyon eğimi		İçsel ve dışsal doğrulama için doğrusal tahmin edicilerin regresyon eğimidir.
		Hosmer-Lemeshow istatistiği		Tahmin edilen olasılıkları gruplandırarak gözlenen değerler ile karşılaştırır.
Yeni Performans Değerlendirme Ölçüleri	Tekrar Sınıflandırma (Reclassification)	Tekrar Sınıflandırma Tablosu	Çapraz tablo ya da Saçılım grafiği	Her iki modelden elde edilen tahmini olasılıkları risk kategorilerine göre çapraz karşılaştırır.
		NRI		Tekrar sınıflandırma tablosunda risk kategorileri arasında meydana gelen değişimi ölçer.
		IDI		Tahmin modelinin ilgilenen olaya sahip olan ve olmayan bireyleri ne kadar doğru ayırt edebildiğini ölçer.

1.4.1. Klasik Performans Değerlendirme Ölçüleri

1.4.1.1. Genel Performans

R^2, bağımlı değişkendeki değişimin bağımsız değişkenlerdeki değişim ile açıklanma oranını belirten, doğrusal regresyon modellerinin karşılaştırılmasında ve performanslarının değerlendirilmesinde en yaygın kullanılan genel performans ölçüsüdür. Bu oran ne denli yüksekse gözlenen ve tahmin değerlerinin o kadar uyumlu olduğu, başka bir deyişle en yüksek R^2 değerine sahip modelin performansının daha iyi olduğu söylenebilir. Bu ölçü, modele ilgisi olmayan bir değişkenin eklenmesi durumunda bile bir miktar artış gösterir. R^2 'nin bu zayıflığını önlemek için, serbestlik derecesine göre düzeltilen *düzeltilmiş* R^2 ($\bar{R}^2$) adlı yeni bir ölçü geliştirilmiştir. R^2 ile $\bar{R}^2$ arasındaki fark çok fazla ise, açıklayıcılık katsayısı olarak $\bar{R}^2$'nin kullanılması önerilmektedir (Alpar, 2011). $\bar{R}^2$,

$$\bar{R}^2 = \left(R^2 - \frac{p}{n-1} \right) \left(\frac{n-1}{n-p-1} \right) \qquad \text{(Denklem 1.24)}$$

ile verilir. Burada n örnek sayısı, p ise değişken sayısıdır.

Bağımlı değişkeninin kategorik yapıda olduğu durumlarda ise, bağımlı değişkenin varyansı bu değişkenin olasılık dağılımına bağlı olarak değişmektedir. Bu durumda lojistik regresyon modelinin performansının değerlendirilmesinde R^2 yerine, olabilirlik oranına göre hesaplanan ve bağımsız değişkendeki bir birimlik değişmenin bağımlı değişkenin koşullu olasılığında yarattığı değişim düzeyini saptamada kullanılan yalancı (pseudo) R^2 istatistikleri geliştirilmiştir (Köksal, 2011). Cox&Snell R^2 ve Nagelkerke R^2 *gibi* ölçüler en sık kullanılanlarıdır. Bu istatistikler küçük çıkma eğiliminde oldukları için model uyumunu değerlendirmekten çok farklı modellerin performanslarını karşılaştırmak için kullanılmaları önerilmektedir (Alpar, 2011).

Cox&Snell R^2: En çok olabilirlik R^2 olarak da bilinen istatistik en küçük 0 değerini alırken en büyük değeri 1 olmamaktadır.

$$R_{cs}^2 = 1 - (\frac{L_1}{L_2})^{\frac{2}{n}} \qquad \text{(Denklem 1.25)}$$

Nagelkerke R^2: Cox&Snell R^2 istatistiğinin alabileceği en büyük değer $1-(L_1)^{\frac{2}{n}}$'dir. Nagelkerke'nin önderdiği düzeltme ile Cox&Snell R^2'nin en büyük değeri 1 olabilmektedir.

$$R_N^2 = 1 - \frac{R_{cs}^2}{1-(L_1)^{\frac{2}{n}}} \qquad \text{(Denklem 1.26)}$$

L_1: Yalnız sabit değer içeren modelin olabilirlik değeri

L_2: Bağımsız değişkenleri içeren modelin olabilirlik değeri

n: Örnek sayısı

Her iki ölçü, [0-1] arasında değer alır. Sonuç 1'e yakın ise gözlem değerleri ile tahmini değerlerin birbirine uyumunun güçlü, 0'a yakın ise zayıf olduğu söylenebilir.

Yanıt değişkeninin ikili yapıda olduğu veriler için model performansının değerlendirilmesinde kullanılan diğer bir ölçü ise Brier Skorudur. Brier skoru Pearson R^2 istatistiğine benzerdir. Gözlenen veriler ile tahmini olasılık değerleri arasındaki farkların karesi alınarak hesaplanan Brier Skoru [0-1] arasında değişen değerler alır. Sonucun 0'a yakın olması modelin gerçek veriye uyumunun iyi olduğunu gösterir.

$$Brier = \frac{1}{\mathrm{n}} \sum_{i=1}^{n} (Y_i - p_i)^2 \qquad \text{(Denklem 1.27)}$$

Bu skorun dezavantajı, yanıt değişkeninin insidansına bağlı olmasıdır. Eğer insidans düşükse, model için maksimum skor da düşük çıkar. Brier skoru Denklem 1.28 ve 1.29'a göre hesaplanarak, %0 ile %100 aralığında değerler alması sağlanabilir (Steyerberg, 2009).

$$Brier_{scaled} = 1 - Brier / Brier_{\max} \quad \text{(Denklem 1.28)}$$

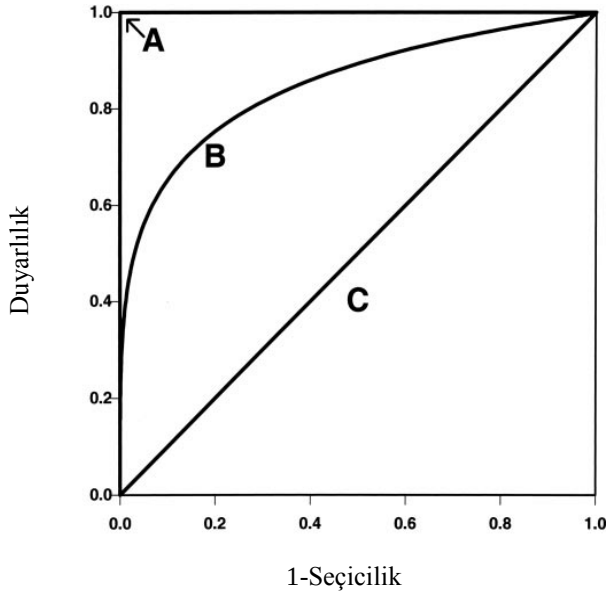
$$Brier_{\max} = (\bar{p}(1-(\bar{p}))^2 + (1-\bar{p})(\bar{p})^2 \quad \text{(Denkelm 1.29)}$$

1.4.1.2. Ayrımsama (Discrimination)

Klinik çalışmalarda, çeşitli teşhis yöntemlerinden, laboratuvar testlerinin sonuçlarından ya da tahmin modellerinden yararlanılarak istenen/istenmeyen olaya sahip bireylerin ayrımının yapılması amaçlanır. Bu durumda tanı için kullanılan bir modelin, hasta bireyleri sağlıklılardan ne kadar doğrulukla ayırt edebildiğinin bilinmesi çok önemlidir. Bu amaçla, ROC analizi tahmin modelinin doğruluğunun değerlendirilmesinde kullanılan en yaygın yöntemdir (Dirican, 2001).

ROC Analizi

Bir tahmin modelinin istenen/istenmeyen olaya sahip bireyleri ne kadar iyi ayırt edebilirliğinin ölçüsü model performansının bir göstergesidir. Bu amaçla, modelden elde edilen tahmini olasılık değerleri seçilen bir eşik değerine göre sınıflandırılır. Sınıfların ayırımı büyük ya da küçük değerler dikkate alınmaksızın rasgele yapıldığından pozitif sınıfında bulunan birimlerin negatif sınıfında da bulunması mümkündür. Yapılan sınıflandırmanın doğruluk düzeyinin değerlendirilmesinde Çizelge 1.6‘da gösterilmiş olan sınıflandırma tablosundan yararlanılır. Burada önemli iki ölçü hesaplanır. Bunlar, doğru pozitiflerin oranı (duyarlılık) ve doğru negatiflerin oranı (seçicilik)’dır. Bu değerler 1’e ne kadar yakın olursa, modelin ayırt edebilirliği de o ölçüde iyidir.



Şekil 1.1. ROC Eğrisi (Zou ve ark.,2007)

Çizelge 1.6. ROC tahminleri için sınıflandırma (2x2) tablosu (Zou ve ark.,2007)

		Gerçek Durum		
		Negatif	Pozitif	Toplam
Test Sonucu	Negatif (D=0)	A=Doğru Negatif	B=Yanlış Negatif	A+B
	Pozitif (D=1)	C=Yanlış Pozitif	D=Doğru Pozitif	C+D
	Toplam	A+C=Hasta değil	B+D=Hasta	A+B+C+D=n

Duyarlılık (Doğru Pozitif Oran = D/(B+D)): Gerçekte pozitif olan tüm olgular içinden, uygulanan test sonucunda pozitif bulunan olguların oranını verir.

Seçicilik (Doğru Negatif Oran = A/(A+C)): Gerçekte negatif olan tüm olgular içinden, test sonucunda negatif bulunanların oranını gösterir.

"1-Seçicilik" (Yanlış Pozitif Oran = C/(A+C)): Gerçekte negatif olan tüm olgular içinden, test sonucunda pozitif bulunanların oranını gösterir.

Doğruluk (Doğru Sınıflandırma Oranı = (A+D)/n): Pozitif ve negatif olan tüm olgular içinden, uygulanan test sonucuna göre pozitif ve negatif olarak kestirilenlerin oranını gösterir.

Bu ölçülerin seçilen eşik değerine göre değişkenlik göstermesi önemli bir dezavantajdır. ROC eğrisi bu olumsuzluğu gidermek amacıyla geliştirilmiştir. Bu eğri, tek bir eşik değeri yerine mümkün tüm eşik değerleri için (prevelanstan bağımsız) dikey eksen üzerinde doğru pozitiflik (duyarlılık) ve yatay eksen üzerinde yanlış pozitiflik (1- seçicilik) oranlarının çakışım noktalarının birleştirilmesi ile elde edilir (Şekil 1.1). Böylece her bir çakışım noktası farklı eşik değerlerine karşılık gelen duyarlılık ve (1-seçicilik) değerlerini ortaya koyar (Kleinbaum, 2010). Genelde düşük yanlış pozitif oranlarını veren eşik değerleri, düşük doğru pozitif oranına da sahiptir. Doğru pozitif oran arttıkça, yanlış pozitif oran da artar (Tomak ve Bek, 2010). Dolayısıyla, doğru negatif oran azalır. Duyarlılık ile seçiciliğin aynı anda artttırılması örneklem büyüklüğünün arttırılması ile sağlanabilir.

Son yıllarda ROC analizi özellikle tıp, veteriner hekimliği, biyoinformatik, radyoloji, psikoloji, makine öğrenme teknikleri ve veri madenciliği alanlarında geniş bir uygulama alanına sahiptir (Zou ve ark., 2007).

ROC Altında Kalan Alan (AUC)

Bu ölçü, ROC eğrisinin performansının değerlendirilmesinde ve model doğruluklarının karşılaştırılmasında yaygın olarak kullanılan bir göstergedir (Köksal, 2011). AUC, tüm mümkün eşik değerleri için duyarlılık ve "1- seçicilik" değerleri ile çizilen grafiğin altında kalan alan olarak tanımlanır. Burada tanımlanan AUC, gerçekte ilgilenilen olaya sahip olan bir bireyin, verilen tanı testi tarafından en yüksek olasılıkla tahmin edebilme olasılığıdır. AUC' un tanım aralığı 0,5 ve 1 arasında olup, bu değerler sırasıyla alt ve üst sınırları oluşturmaktadır. Hiçbir hata olmaksızın hasta ile sağlıklı bireyleri ayırt edebilen model için eğri altında kalan alan 1 br^2(birim kare) iken, hiç ayrımsama yeteneği olmayan bir model için 0.5 br^2(birim kare)'dir. AUC'un 1'e yaklaşması modelin doğru ayrımsama yeteneğinin arttığını gösterir (Öztuna ve ark., 2011).

İki ya da daha çok tahmin modelinden elde edilen ROC eğrilerinin karşılaştırılmasında eğri altında kalan alan ölçüleri karşılaştırılır. Özellikle modele eklenen yeni bir belirtecin modelin ayırt edebilirliğine katkısının değerlendirilmesinde, referans ve geliştirilmiş modelden elde edilen AUC ölçüleri arasındaki farkın anlamlılığı test edilir. DeLong ve ark (1988) yılında yayınladıkları çalışmalarında, iki ve daha fazla ilişkili AUC değerinin karşılaştırılması için genelleştirilmiş U istatistiğine dayanan parametrik olmayan bir yaklaşım sunmuşlardır.

Bu yaklaşıma göre, referans model ile yeni bir belirtecin eklendiği geliştirilmiş modele ait AUC ölçüleri aşağıdaki hipoteze göre test edilir.

$$H_0^{AUC} : AUC_p = AUC_{p-k}$$
$$H_1^{AUC} : AUC_p \neq AUC_{p-k} \qquad \text{(Denklem 1.30)}$$

AUC_p geliştirilmiş modelin, AUC_{p-k} ise referans modelin ölçüleridir. AUC, iki sınıfa ilişkin birim dağılımının karşılaştırılmasında kullanılan Mann-Whitney U istatistiği tarafından tahmin edilir. Bu tahmin edici parametrik olmayan yansız bir tahmin edicidir ve eAUC ile gösterilir.

$$eAUC = \frac{1}{n_0 n_1} = \sum_{x_i \varepsilon D_0} \sum_{x_j \varepsilon D_1} I\left[a'x_i, a'x_j\right] \qquad \text{(Denklem 1.31)}$$

D_0 , ilgilenilen olaya sahip olmayan durumu; D_1 , ise ilgilenilen olaya sahip olan durumu gösterir. n_0 ve n_1 de, ilgili durumların örnek büyüklükleridir.

$$I\left[a'x_i, a'x_j\right] = \begin{cases} 1 & , a'x_i < a'x_j \\ 0.5 & , a'x_i = a'x_j \\ 0 & , \text{dd} \end{cases} \qquad \text{(Denklem 1.32)}$$

DeLong (1988) tarafından geliştirilen ve iki eAUC ölçülerini karşılaştıran test istatistiği,

$$z = \frac{(eAUC_p - eAUC_{p-k} - (AUC_p - AUC_{p-k}))}{((1,-1)S(1,-1)')^{1/2}} \quad \text{(Denklem 1.33)}$$

şeklindedir. S, varyans-kovaryans matrisinin tahmin edicisidir ve z test istatistiği sıfır hipotezi altında normal dağılıma sahiptir (Delong ve ark., 1988).

1.4.1.3. Kalibrasyon (Calibration)

Tahmin modellerinin performanslarının değerlendirilmesinde en önemli adımlardan birisi de modelin kalibrasyonudur. Kalibrasyon, gözlenen gerçek olasılıklar ile model tarafından tahmin edilen olasılıklar arasındaki uyumu ölçer. Bir diğer deyişle, ilgilenilen olaya sahip bir bireyin model tarafından tahmin edilen olasılık değerinin, popülasyonda aynı olayı deneyimlemiş tüm bireylerin olasılık değerlerine çok yakın olmasıdır.

Hosmer-Lemeshow istatistiği, model kalibrasyonun değerlendirilmesinde kullanılan en yaygın testtir (Janes ve ark., 2008). Bu test, 1.3.1.6. Hosmer-Lemeshow İstatistiği başlığı altında ayrıntılı olarak anlatılmıştır. Test istatistiği,

$$C^* = \sum_{t=1}^{k} \frac{(O_k - n_k \bar{p}_k)^2}{n_k \bar{p}_k (1 - \bar{p}_k)} \sim \chi^2_{k-2} \quad \text{(Denklem 1.34)}$$

n_k gözlem sayısı, O_k pozitif gözlem sayısı, $\bar{p}_k$ ise, k. grubun tahmin edilen olasılıklarının ortalamasını göstermektedir. Düşük p değeri ($p<0.05$) gözlenen ve tahmini olasılıklar arasındaki uyumun zayıf olduğunu gösterir.

Bu test, kalibrasyon grafiği ile görsel olarak da değerlendirilebilir. Kalibrasyon grafiğinde yatay eksende tahmin değerleri, düşey eksende gözlem değerleri yer alır. Bu noktalardan geçen eğri, gözlenen değerler ile tahmini değerler arasındaki uyumu gösterir. Mükemmel tahminler 45 derecelik bir doğru oluşturur. Yanıt değişkeninin ikili yapıda olduğu durumlarda, grafiğin düşey ekseninde 0 ve 1 değerleri yer alır. Olasılıklar direkt olarak gözlenmez. Ancak, düzeltme teknikleri kullanılarak tahmin

edilen olasılıklar ile gözlenen olasılıklar arasındaki ilişki düzenlenebilir (Steyerberg ve ark., 2010).

Kalibrasyon grafiği, kesim noktası ve kalibrasyon eğimi yardımı ile yorumlanabilir. Kesim noktası ("calibration-in-large"), ortalama tahmin değeri ile ortalama gözlenen değer arasındaki farktır. Kalibrasyon eğimi ise, tahmin edicilerin etki güçlerinin ortalamasını gösterir. İyi kalibre olmuş bir tahmin modelinde kesim noktası 0, kalibrasyon eğimi ise 1'dir (Steyerberg, 2009).

1.4.2. Literatürde Yeni Önerilen Performans Ölçüleri

1.4.2.1. Tekrar Sınıflandırma (Reclassification) Tablosu ve NRI

İlgilenilen olayın ortaya çıkma olasılıkları, bilinen risk faktörlerini içeren referans model ve bunlara ek olarak yeni belirtecin eklendiği geliştirilmiş model kullanılarak ayrı ayrı tahmin edilir. Bu iki modelden tahmin edilen olasılıklar, anlamlı eşik değerleri kullanılarak (genellikle yüzde ile ifade edilir. Örneğin; <%20-Düşük Riskli, ≥%20-Yüksek Riskli) risk kategorilerine göre sınıflandırılır ve bir çapraz tablo elde edilir. Bu tablo tekrar sınıflandırma tablosu olarak adlandırılır ve saçılım grafiği ile de gösterilebilir. NRI, referans ve geliştirilmiş modelden elde edilen tekrar sınıflandırma tablosundaki (çapraz tablo) istenen / istenmeyen olaya sahip bireylerin risk kategorileri arasındaki değişimini aşağı ve yukarı değişim kavramları kullanarak ölçen bir indekstir. Yukarı doğru hareketlilik, geliştirilmiş modelin gerçekte istenmeyen olaya sahip bireyi daha riskli bir kategoriye sınıflaması iken, aşağı doğru hareketlilik ise, gerçekte istenen olaya sahip bireyin daha düşük riskli bir kategoriye sınıflamasıdır (Gu ve Pepe, 2009).

$$NRI = [P(yukarı \mid D_1) - P(aşağı \mid D_1)] - [P(yukarı \mid D_0) - P(aşağı \mid D_0)]$$

(Denklem 1.35)

$D_1 =$ İlgilenilen olay var

$D_0 =$ İlgilenilen olay yok

olarak tanımlanır (Pencina ve ark., 2008).

$$\hat{P}(yukarı \mid D=1) = \hat{p}_{\text{yukarı},D_1} = \frac{\text{Üst kategoriye geçen ilgilenilen olaya sahip birey sayısı}}{\text{İlgilenilen olaya sahip toplam birey sayısı}}$$

(Denklem 1.36)

$$\hat{P}(aşağı \mid D=1) = \hat{p}_{\text{aşağı},D_1} = \frac{\text{Alt kategoriye geçen ilgilenilen olaya sahip birey sayısı}}{\text{İlgilenilen olaya sahip toplam birey sayısı}}$$

(Denklem 1.37)

$$\hat{P}(yukarı \mid D=0) = \hat{p}_{\text{yukarı},D_0} = \frac{\text{Üst kategoriye geçen ilgilenilen olaya sahip olmayan birey sayısı}}{\text{İlgilenilen olaya sahip olmayan toplam birey sayısı}}$$

(Denklem 1.38)

$$\hat{P}(aşağı \mid D=0) = \hat{p}_{\text{aşağı},D_0} = \frac{\text{Alt kategoriye geçen ilgilenilen olaya sahip olmayan birey sayısı}}{\text{İlgilenilen olaya sahip olmayan toplam birey sayısı}}$$

(Denklem 1.39)

NRI ‘nın tahmin edicisi ise,

$$\hat{NRI} = (\hat{p}_{\text{yukarı},D_1} - \hat{p}_{\text{aşağı},D_1}) - (\hat{p}_{\text{yukarı},D_0} - \hat{p}_{\text{aşağı},D_0}) \qquad \text{(Denklem 1.40)}$$

biçiminde verilir.

NRI, tekrar sınıflandırma tablosunda meydana gelen değişimi ölçen basit bir yüzde hesabı gibi düşünülebilir. Fakat NRI, risk kategorileri arasındaki hareketliliği

ölçmenin yanı sıra, kategoriler arasında doğru yöne hareket edenleri de ayırt eder (Cook ve Ridker, 2009).

İstenilen olaya sahip olan ve olmayan bireylerin birbirinden bağımsız olduğu varsayılır ve McNemar (1947) testinin mantığı kullanılarak NRI=0 hipotezi Denklem 1.41 ile test edilir (Pencina ve ark., 2008).

$$\frac{\hat{NRI}}{\sqrt{\frac{\hat{p}_{\text{yukarı},D_1}+\hat{p}_{\text{aşağı},D_1}}{\text{İlgilenilen olaya sahip toplam birey sayısı}}+\frac{\hat{p}_{\text{yukarı},D_0}+\hat{p}_{\text{aşağı},D_0}}{\text{İlgilenilen olaya sahip olmayan toplam birey sayısı}}}}$$

(Denklem 1.41)

Ayrıca, ilgilenilen olaya sahip olan ve olmayan bireylerin sınıflamadaki iyileşmesi aşağıdaki eşitlikler kullanılarak test edilebilir.

$$z_{D_1}=\frac{\hat{p}_{\text{yukarı},D_1}-\hat{p}_{\text{aşağı},D_1}}{\sqrt{\frac{\hat{p}_{\text{yukarı},D_1}+\hat{p}_{\text{aşağı},D_1}}{\text{İlgilenilen olaya sahip toplam birey sayısı}}}} \qquad \text{(Denklem 1.42)}$$

$$z_{D_0}=\frac{\hat{p}_{\text{aşağı},D_0}-\hat{p}_{\text{yukarı},D_0}}{\sqrt{\frac{\hat{p}_{\text{aşağı},D_0}+\hat{p}_{\text{yukarı},D_0}}{\text{İlgilenilen olaya sahip olmayan toplam birey sayısı}}}} \qquad \text{(Denklem 1.43)}$$

NRI'nın bir dezavantajı, eşik değeri ve buna bağlı olarak oluşturulan risk kategorilerinin seçimine bağlı olmasıdır. Bu durum NRI'nın geliştirilmesine neden olmuştur. Yukarı doğru hareketlilik 1, aşağı doğru hareketlilik -1 ve değişimin olmadığı durumları 0 olarak tanımlandığında, NRI Denklem 1.44 gibi ifade edilebilir.

$$\frac{\sum_{i_{D_1}} v(i)}{\text{İlgilenilen olaya sahip toplam birey sayısı}}-\frac{\sum_{j_{D_0}} v(j)}{\text{İlgilenilen olaya sahip olmayan toplam birey sayısı}}$$

(Denklem 1.44)

$v(i)$, kategoriler arası hareketi tanımlar. Kategoriler arasındaki hareketlilik, 1,-1 ve 0 değerleri yerine, her bir bireyin tahmini olasılıklarındaki gerçek farklılıklar ile gösterilebilir. Bu durumda geliştirilmiş model $\hat{p}_{\text{geliştirilmiş model}}$ ve referans model $\hat{p}_{\text{referans model}}$ şeklinde ifade edilirse,

$$\frac{\sum_{i_{D_1}} (\hat{p}_{\text{geliştirilmiş model}}(i) - \hat{p}_{\text{referans model}}(i))}{\text{İlgilenilen olaya sahip toplam birey sayısı}} - \frac{\sum_{j_{D_0}} (\hat{p}_{\text{geliştirilmiş model}}(j) - \hat{p}_{\text{referans model}}(j))}{\text{İlgilenilen olaya sahip olmayan toplam birey sayısı}}$$

(Denklem 1.45)

eşitliği elde edilir. Denklem 1.45'te birinci terim duyarlılık değerinde, ikinci negatif terim ise seçicilik değerinde meydana gelen iyileşmeyi ölçer.

Bu ölçü, mümkün bütün eşik değerlerine karşılık gelen duyarlılık ve seçicilik değerlerindeki iyileşmeyi dikkate alır. Bir başka ifade ile duyarlılık eğrisi altında kalan alan ilgilenilen olaya sahip bireylerin tahmini olasılık değerlerinin ortalaması ile tanımlanır. "1- seçicilik" eğrisi altında kalan alan ise, ilgilenilen olaya sahip olmayan bireylerde olayın tahmini görülme olasılığının ortalamasıdır. Ortalama duyarlılık ve "1- seçicilik" değerleri arasındaki fark, model performansının ölçülmesinde kullanılan diğer bir ölçüdür ve bu ölçü IDI 'dır.

1.4.2.2. IDI

NRI'nın sürekli versiyonu olarak bilinen IDI, risk kategorilerinden bağımsız olarak, modelin ilgilenilen/ilgilenmeyen olaya sahip bireyleri birbirinden ne kadar ayırt edebildiğini ölçen bir indekstir. Risk kategorileri belirlenemediği ya da bilinmediği durumlarda kullanılan IDI, [0-1] aralığındaki mümkün tüm eşik değerlerine karşılık gelen referans ve geliştirilmiş modeldeki duyarlılık değerlerinin ortalamasında meydana gelen iyileşme ile "1-seçicilik" değerlerinin ortalamasında meydana gelen artış arasındaki farka eşittir (Pencina ve ark., 2008). IDI,

$$IDI = (IS_{geliştirilmiş\ model} - IS_{referans\ model}) - (IP_{geliştirilmiş\ model} - IP_{referans\ model})$$

(Denklem 1.46)

ile gösterilir.

IS, [0,1] aralığındaki tüm mümkün eşik değerlerine karşılık gelen duyarlılık değerlerinin integrali, *IP* ise (1-seçicilik) değerlerinin integralidir ve aşağıdaki eşitlikler ile hesaplanır.

$$IS = \int_0^1 S(u)du = \int_0^1 \int_u^1 f(x \mid D=1)dxdu \qquad \text{(Denklem 1.47)}$$

$$IP = \int_0^1 P(u)du = \int_0^1 \int_u^1 f(x \mid D=0)dxdu \qquad \text{(Denklem 1.48)}$$

Denklem 1.47 ve Denklem 1.48'de integral sırası değiştirilirse,

$$IS = \int_0^1 \int_0^x f(x \mid D=1)dudx = \int_0^1 xf(x \mid D=1)dx = E(X \mid D=1) \qquad \text{(Denklem 1.49)}$$

$$IP = \int_0^1 \int_0^x f(x \mid D=0)dudx = \int_0^1 xf(x \mid D=0)dx = E(X \mid D=0) \qquad \text{(Denklem 1.50)}$$

ilgilenilen olaya sahip olan ve olmayan duruma ilişkin koşullu beklenen değerleri elde edilir. Bu beklenen değerler, ilgilenilen olaya sahip olan ve olmayan duruma ait örnek ortalamaları ile tahmin edilir. Tahmini örnek ortalamalarına ilişkin hesaplamalar aşağıda gösterilmiştir.

$$\bar{\hat{p}}_{D_1} = \frac{\sum_{i_{D_1}} \hat{p}_i}{\text{İlgilenilen olaya sahip toplam birey sayısı}} \qquad \text{(Denklem 1.51)}$$

$$\bar{\hat{p}}_{D_0} = \frac{\sum_{i_{D_0}} \hat{p}_i}{\text{İlgilenilen olaya sahip olmayan toplam birey sayısı}} \qquad \text{(Denklem 1.52)}$$

Böylece $\hat{IDI}$,

$$\hat{IDI} = (\bar{\hat{p}}_{geliştirilmiş\ \text{model},D_1} - \bar{\hat{p}}_{referans\ \text{model},D_1}) - (\bar{\hat{p}}_{geliştirilmiş\ \text{model},D_0} - \bar{\hat{p}}_{referans\ \text{model},D_0}) \quad \text{(Denklem 1.53)}$$

şeklinde tahmin edilir.

Denklem 1.53, aşağıdaki şekilde de yazılabilir.

$$\hat{IDI} = (\bar{\hat{p}}_{D_1} - \bar{\hat{p}}_{D_0})_{geliştirilmiş\ \text{model}} - (\bar{\hat{p}}_{D_1} - \bar{\hat{p}}_{D_0})_{\text{referans model}} \quad \text{(Denklem 1.54)}$$

Denklem 1.54, geliştirilmiş ve referans modelin ayırt etme eğimlerinin farkına denktir. Yates (1982) tarafından önerilen bu ayırt etme eğimi, ilgilenilen olaya sahip bireylerin ortalama tahmini olasılık değerleri ile ilgilenilen olaya sahip olmayan bireylerin ortalama tahmini olasılık değerleri arasındaki farktır. Bu bağlamda IDI, modellerin ayırt etme eğimleri arasındaki fark kullanılarak da tahmin edilebilir (Pencina ve ark., 2008).

İlgilenilen olaya sahip olan ve olmayan bireylerin birbirinden bağımsız olduğu varsayılır ve McNemar (1947) testinin mantığı kullanılarak IDI=0 hipotezi Denklem 1.55 ile test edilir. $\hat{se}_{D_1}$, ilgilenilen olaya sahip geliştirilmiş ve referans modelden elde edilen tahmini olasılıkları arasındaki farkın standart hatası, $\hat{se}_{D_0}$ ise ilgilenilen olaya sahip olmayan geliştirilmiş ve referans modelden elde edilen tahmini olasılıkları arasındaki farkın standart hatasıdır (Pencina ve ark., 2008).

$$z = \frac{\hat{IDI}}{\sqrt{(\hat{se}_{D_1})^2 + (\hat{se}_{D_0})^2}} \quad \text{(Denklem 1.55)}$$

Eğer yeni eklenen bir belirteç modelin ayırt edebilirliğine bir katkı sağlıyorsa, Denklem 1.53'deki birinci terim pozitif yönde bir artış, ikinci terim ise negatif yönde artış sağlar (Kerr ve ark., 2011).

1.4.2.3. IDI ve AUC İlişkisi

IDI ve AUC ölçüleri, duyarlılık ve "1-seçicilik" değerlerini dikkate alarak hesaplama yapar. Bu bağlamda bu iki ölçü birbirine benzerdir ve her iki ölçüde de duyarlılık değerleri düzeltilmiştir. Fakat, her iki ölçü incelendiğinde IDI'da, duyarlılık değerleri "1-seçicilik" değerinde meydana gelen istenmeyen artış çıkartılarak düzeltilirken, AUC da ise, seçicilik değerlerinin türevleri ile düzeltilir. Bu açıdan IDI ve AUC birbirinden farklı sonuçlar üretir. Eşik değeri u, X 'in olasılık yoğunluk fonksiyonu f ile gösterilirse, duyarlılık ve "1-seçicilik" Denklem 1.56 ve Denklem 1.57'deki gibi ifade edilir.

$$Duyarlılık = S(u) = P(X > u|D = 1) = \int_u^1 f(x|D = 1)\,dx \qquad \text{(Denklem 1.56)}$$

$$1 - Seçicilik = P(u) = P(X > u|D = 0) = \int_u^1 f(x|D = 0)\,dx \qquad \text{(Denklem 1.57)}$$

AUC ise,

$$AUC = \int_0^1 S(u)\frac{d}{du}P(u)\,du = \int_0^1 S(u)\,f(u|D = 0)\,du = P(X_i > X_j \mid D_i = 1, D_j = 0) \qquad \text{(Denklem 1.58)}$$

şeklinde ifade edilir (Pencina ve ark., 2008).

1.4.2.4. NRI ve AUC İlişkisi

İki kategorili yani tek bir eşik değerine sahip olan çalışmalarda NRI, her iki modelin AUC değerlerinin farkları ile ilişkilidir. Sınıflandırma için kullanılan eşik değeri u ile gösterilirse, ilgili eşik değerine ait AUC Denklem 1.59'a eşittir.

$$AUC = \frac{1}{2}(Duyarlılık(u) + Seçicilik(u)) \qquad \text{(Denklem 1.59)}$$

Her iki model için farkları alındığında,

$$\Delta AUC = \frac{1}{2}(\Delta Duyarlılık(u) + \Delta Seçicilik(u)) \qquad \text{(Denklem 1.60)}$$

elde edilir. NRI ise Denklem 1.61'de verildiği şekilde hesaplanır.

$$\begin{aligned} \Delta Duyarlılık &= P(yukarı \mid D_1) - P(aşağı \mid D_1) \\ \Delta Seçicilik &= P(yukarı \mid D_0) - P(aşağı \mid D_0) \end{aligned} \qquad \text{(Denklem 1.61)}$$

Youden indeksi, doğru pozitif oran ile yanlış pozitif oran arasındaki farkı maksimum yapan, ROC eğrisinde sürekli veriler için en iyi ayrım etkinliğini veren eşik değerinin belirlenmesini sağlar.

$$Youden\ indeksi = \max(duyarlılık + seçicilik - 1) \qquad \text{(Denklem 1.62)}$$

Her iki modelden elde edilen Youden indekslerinin farkı alındığında Denklem 1.63 elde edilir.

$$\Delta Youden\ indeksi = \Delta Duyarlılık + \Delta Seçicilik \qquad \text{(Denklem 1.63)}$$

Sonuç olarak, Denklem 1.60'dan aşağıdaki ilişki elde edilir.

$$\Delta AUC(u) = \frac{1}{2}\Delta Youden(u) = \frac{1}{2} NRI(u) \qquad \text{(Denklem 1.64)}$$

Denklem 1.64'e göre, iki kategoriye sahip NRI, her iki modelin AUC değerlerinin farkının iki katına ve Youden indeksteki artışa eşittir (Pencina ve ark., 2011; Steyerberg ve ark., 2011).

Bu çalışmada, tahmin modellerinin performanslarının değerlendirilmesinde kullanılan klasik yöntemler ve bu yöntemlere alternatif olarak sunulan iki yeni yöntem tanıtılarak, avantajları ve dezavantajları incelenecektir. Bu yöntemler, literatürde özellikle yeni bir belirtecin kullanılabilirliğinin belirlenmesi ve var olan bir modelin

geliştirilmesi ile ilgili çalışmalarda kullanılmaktadır. Bu çalışmada, spesifik olarak yetiştiricilik alanındaki uygulamalarda da literatürde önerilen yeni yöntemler NRI ve IDI'nın kullanılması amaçlanmıştır. Bu araştırma NRI ve IDI yöntemlerinden yararlanarak doğum ağırlığı değişkeninin; Akkaraman, Sakız x Akkaraman F_1 ve Kıvırcık x Akkaraman F_1 koyun genotipleri kullanılarak kuzuların büyüme özellikleri üzerine etkisinin incelenmesi amacıyla yapılmıştır.

2. GEREÇ VE YÖNTEM

2.1. Gereç

Çalışmada, Akçapınar ve ark (2000) tarafından 1996-1997 yılları arasında Sivas-Ulaş Tarım İşletmesinde yürütülen bir projenin veri seti kullanılmıştır. Çalışmada kullanılan veriler, 90. günde yaşayan 111 baş Akkaraman, 210 baş Sakız x Akkaraman F_1 ve 121 baş Kıvırcık x Akkaraman F_1 kuzuya ait büyüme verileridir. Çalışmada belirli sebeplerden 90. gün ağırlıkları olmayan 9 kuzuya ait veri analiz dışı tutularak toplam 433 kuzuya ait büyüme verileri kullanılmıştır. Her kuzunun doğum ağırlığı, genotipi, doğum yılı, cinsiyeti ve doğum tipi kaydedilmiştir. Normalden fazla ya da düşük doğum ağırlığındaki kuzuların yaşama gücü düşük olmakta ve bu kuzularda ölüm oranı yüksek olmaktadır. Bu nedenle doğum ağırlığının normal aralıkta olması istenir. Diğer taraftan kuzuların büyüme döneminde ırk standartlarına uygun canlı ağırlık değerleri göstermesi beklenir. Canlı ağırlığın düşük olması verim kaybı olduğu için yetiştiricilikte istenmez. Çalışmada, öncelikle sütten kesim ağırlığı (SKA) kuzuların 90. gün ağırlıklarının ortalaması alınarak tanımlanmıştır. Bu ortalama değerin altında ağırlığa sahip kuzular 1 (yüksek risk grubu), üstünde ağırlığa sahip olan kuzular 0 (düşük risk grubu) olarak kodlanmıştır. Dolayısıyla bağımlı değişken iki kategoriye sahiptir. Her kuzunun 90. gün ağırlığına etki eden risk faktörlerinden genotip, doğum yılı, cinsiyet ve doğum tipi referans modeli oluşturmaktadır. Referans modele etkisi araştırılacak doğum ağırlığı değişkeni, belirteç olarak eklenmiş ve geliştirilmiş model elde edilmiştir.

2.2. Yöntem

Araştırmada kullanılan koçlar kendilerine ayrılan koyunlarla *elde sıfat* yöntemi ile çiftleştirilmiştir. Doğan kuzular kuruduktan sonra 50 g hassas terazi ile tartılmış, kulak küpesi ile numaralandırılmıştır. Kuzular doğumu takiben 2-3 gün süreyle analarıyla birlikte bırakılmışlardır. Daha sonra analarından ayrılarak sabah ve akşam olmak üzere

günde iki kez emzirilmişlerdir. İkinci haftadan itibaren kuru yonca ve kuzu büyütme yemi verilmeye başlanmıştır. Tek ve ikiz doğan kuzular süt emme döneminde ayrı gruplar halinde büyütülmüşlerdir. Kuzular 90 günlükken sütten kesilmişlerdir. Süt kesiminden daha sonra ise erkek ve dişi kuzular ayrı sürüler halinde merada otlatılmışlardır (Akçapınar ve ark.,2000).

Bu çalışmada kuzular, SKA'nın ortalama değerine göre iki gruba ayrılmıştır. Öncelikle aday değişkenlerin belirlenmesi için kuzularda SKA'ya etki eden risk faktörleri tek değişkenli lojistik regresyon analizi uygulanarak belirlenmiştir. Analiz sonuçlarına göre anlamlı değişkenler çoklu lojistik regresyon modeline girmeye aday değişkenler olarak kullanılmıştır. İkinci olarak, lojistik regresyon analizinde geriye doğru değişken çıkarma yöntemi kullanılarak, çoklu lojistik regresyon modelindeki değişkenlerin önemliliği Wald testi ile değerlendirilmiştir. Elde edilen final modelin uyum iyiliği testi yapıldıktan sonra modelden elde edilen tahmini olasılıklar kullanılarak modelin performansı klasik performans ölçülerinden ROC-AUC, Brier Skoru ve Nagelkerke R^2 ölçüleri ve literatürde yeni önerilen performans ölçülerinden NRI ve IDI ölçüleri ile incelenmiştir. Çalışmada farklı eşik değerleri ve risk kategorilerine göre NRI değerleri incelenmiş ve 0.30 eşik değeri istatistiksel olarak anlamlı bulunmuştur. AUC değerindeki artışın anlamlılığı, De-Long testi ile değerlendirilmiştir.

Çalışmada lojistik regresyon analizi ve risk faktörlerinin hesaplanması SPSS for Windows 14.01 (Lisans No:9869264) paket programı ile yapılmıştır. Model performans ölçülerinin hesaplanması ve model performanslarının karşılaştırılması ile ilgili analizler ise R (Version i386 3.1.3) programında PredictABEL, Hmisc, ROCR ve Clinfun paketleri yardımıyla değerlendirilmiştir. Çalışmada kullanılan R kodlarına Ek 1'de yer verilmiştir.

3. BULGULAR

3.1. Tanımlayıcı İstatistikler

Veri setinde kullanılan değişkenlere ait tanımlayıcı istatistikler Çizelge 3.1'de verilmiştir. Veri setindeki doğum ağırlığı (kg) sürekli değişkeninin ortalaması ± standart sapması ve değişim aralığı ise sırasıyla 4.76±0.711 kg, 2.4-6.7 kg olarak hesaplanmıştır.

Çizelge 3.1. Değişkenlere ait tanımlayıcı istatistikler

Değişkenler		n	% n
90.gün canlı ağırlığı (kg)	≤23.37	218	50.3
	>23.37	215	49.7
Genotip	Akkaraman	102	23.6
	Sakız x Akkaraman F_1	210	48.5
	Kıvırcık x Akkaraman F_1	121	27.9
Cinsiyet	Erkek	216	49.9
	Dişi	217	50.1
Doğum Tipi	Tek	234	54.0
	İkiz	199	46.0
Doğum Yılı	1996	141	32.6
	1997	292	67.4

3.2. Lojistik Regresyon Analiz Sonuçları

3.2.1. Model Seçimi

SKA'ya etkisi olabilecek risk faktörlerinin belirlenmesi için öncelikle tek değişkenli lojistik regresyon analizi ile değişkenler incelenmiştir. Analiz sonrasında elde edilen Wald istatistiği ve başlangıç model ile tek değişkenli model arasındaki -2log olabilirlik değerlerinin farkından elde edilen Ki Kare değerleri Çizelge 3.2' de sunulmuştur.

Çizelge 3.2. Tek değişkenli lojistik regresyon analiz sonuçları

Değişken	sd	LR χ^2	p	Wald	p
Genotip	2	9.962	0.002	9.738	0.002
Cinsiyet	1	59.772	<0.001	55.580	<0.001
Doğum Tipi	1	5.20	0.023	5.167	0.023
Doğum Yılı	1	17.034	<0.001	16.486	<0.001
Doğum Ağırlığı	1	42.229	<0.001	36.877	<0.001

Tek değişkenli yaklaşımda değişkenlerin toplam ilişkisi göz ardı edilir. Değişkenlerin her biri sonuç değişkeni ile zayıf bir ilişki içinde olsa bile birlikte alındıkları zaman sonuç değişkeninin önemli bir kestiricisi olabilirler. Bu nedenle Çizelge 3.2'de Wald ve Ki-kare test istatistiklerine ait 'p' değeri 0.20'den küçük olan ($p<0.20$) değişkenler, çok değişkenli model için aday değişken olarak belirlenmiştir (Hoshmer ve Lemeshow 2000). Dolayısıyla bağımsız değişkenlerin tümü çoklu lojistik regresyona aday olmuştur.

3.2.2. Referans ve Geliştirilmiş Modellerin Oluşturulması

Aday değişkenler ile kurulan çoklu lojistik regresyon modelinde hangi değişkenlerin modelde kalacağına karar vermek için çalışmada geriye doğru adımsal çıkarma yöntemi kullanılmıştır. Bu yöntemde ilk adımda genotip, doğum yılı, cinsiyet ve doğum tipi değişkenleri modele dahil edilmiştir. Tüm değişkenler en son adımda anlamlı bulunduğu için referans modelde yer almıştır ve bu değişkenlerin SKA'yı etkileyen önemli risk faktörleri olduğu belirlenmiştir.

Çizelge 3.3. Referans modele ilişkin geriye doğru değişken çıkarma yöntemi ile en son adımda elde edilen final model

Değişken		β	SE(β)	Wald	p	OR	OR'nin %95 Güven Aralığı
Genotip	* Akkaraman						
	Sakız x Akkaraman F_1	0.463	0.283	2.66	0.102	1.59	[0.91-2.77]
	Kıvırcık x Akkaraman F_1	1.173	0.319	13.52	<0.001	3.23	[1.73-6.04]
Cinsiyet	* Erkek						
	Dişi	1.716	0.226	57.74	<0.001	5.56	[3.57-8.65]
Doğum Tipi	*Tek						
	İkiz	0.925	0.239	14.93	<0.001	2.52	[1.58-4.03]
Doğum Yılı	*1997						
	1996	1.336	0.258	26.78	<0.001	3.80	[2.29-6.31]

*Referans kategori

Çizelge 3.3'te yer alan referans modelin odds oranlarına göre, Kıvırcık x Akkaraman F_1 genotipinin, Akkaraman genotipine göre SKA 'nın düşük olma riski 3.23 kat daha yüksektir. Aynı şekilde, dişilerin erkeklere göre 5.56 kat, ikizlerin teklere göre 2.52 kat ve 1996 yılı doğumluların da 1997 yılı doğumlulara göre 3.80 kat SKA'nın düşük olma riski daha yüksek olacaktır.

Çalışmada, Çizelge 3.3'te elde edilen referans modele doğum ağırlığı değişkeni eklenerek geriye doğru adımsal çıkarma yöntemi kullanılarak geliştirilmiş model elde edilmiştir. En son adımda, genotip, cinsiyet, doğum tipi ve doğum yılı ve doğum ağırlığı değişkenleri geliştirilmiş modelde yer almıştır ve doğum ağırlığı değişkenin de SKA'yı etkileyen önemli risk faktörü olduğu belirlenmiştir.

Çizelge 3.4. Geliştirilmiş modele ilişkin geriye doğru değişken çıkarma yöntemi ile en son adımda elde edilen final model

Değişken		β	SE(β)	Wald	p	OR	OR'nin %95 Güven Aralığı
Genotip	* Akkaraman						
	Sakız x Akkaraman F_1	0.434	0.289	2.24	0.134	1.54	[0.88-2.72]
	Kıvırcık x Akkaraman F_1	1.194	0.327	13.35	<0.001	3.30	[1.74-6.26]
Cinsiyet	*Erkek						
	Dişi	1.637	0.232	49.95	<0.001	5.14	[3.26-8.09]
Doğum Tipi	*Tek						
	İkiz	0.485	0.264	3.38	0.066	1.62	[0.97-2.72]
Doğum Yılı	*1997						
	1996	1.222	0.266	21.04	<0.001	3.39	[2.01-5.72]
Doğum Ağırlığı		-0.795	0.187	18.00	<0.001	0.45	[0.31-0.65]

*Referans kategori

Çizelge 3.4'te yer alan doğum ağırlığı değişkenin eklendiği geliştirilmiş modelin odds oranlarına göre, Kıvırcık x Akkaraman F_1 genotipinin Akkaraman genotipine göre SKA'nın düşük olma riski 3.30 kat daha yüksektir. Aynı şekilde, dişilerin erkeklere göre 5.14 kat ve 1996 yılı doğumluların da 1997 yılı doğumlulara göre 3.39 kat SKA'nın düşük olma riski daha yüksek olacaktır. Doğum ağırlığındaki bir birimlik artış, SKA'nın düşük olma riskini 0.45 kat azaltacaktır.

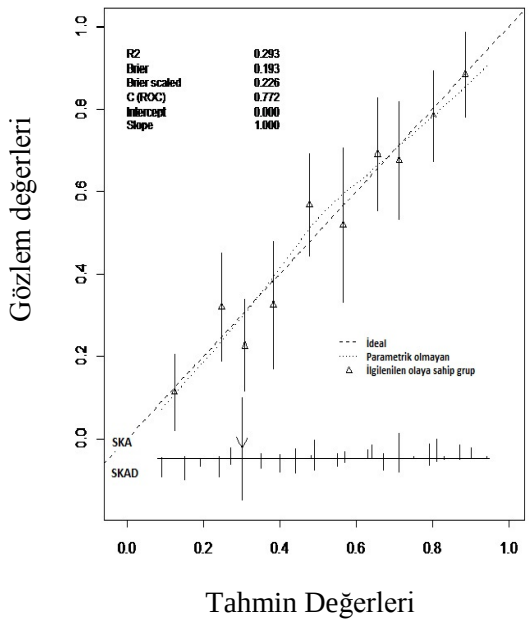
3.2.3. Modellerin Uyum İyiliği Testi

Elde edilen referans ve geliştirilmiş modellerin uyum iyiliği testi Hosmer-Lemeshow'un onlu risk grupları yöntemine göre yapılmıştır. Bu istatistik, bağımsız değişkenlerden hiçbirinin bağımlı değişken ile anlamlı bir ilişki göstermediğini ileri süren sıfır hipotezini test etmektedir.

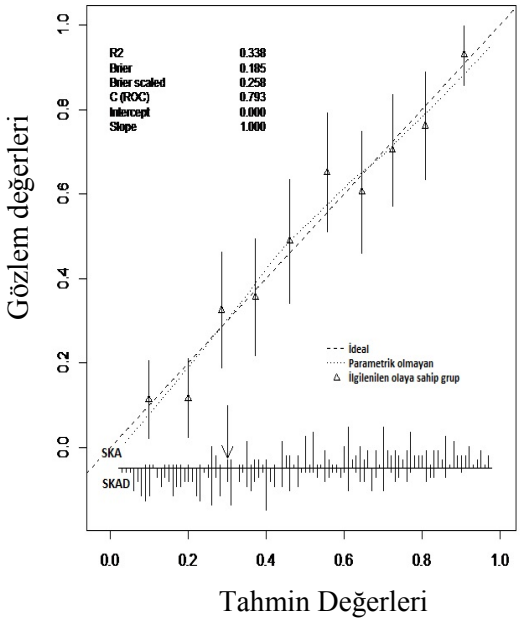
Çizelge 3.5. Hosmer-Lemeshow test istatistiği ve önem kontrolü

Model	Ki-Kare	sd	p
Referans Model	6.075	8	0.639
Geliştirilmiş Model	8.213	8	0.413

Çizelge 3.5'de, değişkenlerin modele uyumunun oldukça iyi olduğu ve lojistik regresyon modellerinin kuzuları SKA'na göre ayırmada yeterli bir model olduğu sonucuna varılmıştır. Bu testi Şekil 3.1 ve Şekil 3.2'de verilen kalibrasyon grafikleri ile görsel olarak değerlendirildiğinde, kalibrasyon kesim noktasının 0 ve eğiminin 1 olmasından dolayı modelin iyi kalibre olmuş bir model olduğu söylenebilir.



Şekil 3.1. Referans modelin kalibrasyon grafiği



Şekil 3.2. Geliştirilmiş modelin kalibrasyon grafiği

Çizelge 3.6'da referans modele ait duyarlılık %66.5, seçicilik ise %74.9 olarak bulunmuştur. Çizelge 3.7'de geliştirilmiş modelde ise, duyarlılık %72.5, seçicilik ise %72.1'dir. Dolayısıyla, modelin doğru sınıflandırma oranı %70.7 iken, modele doğum ağrılığı değişkenin eklenmesi ile elde edilen modelin doğru sınıflandırma oranı %72.3 olarak bulunmuştur.

Çizelge 3.6. Referans modelin sınıflandırma tablosu*

	Beklenen		
Gözlenen	SKA Düşük	SKA Yüksek	Doğru Sınıflandırma Oranı (%)
SKA Düşük	145	73	66.5
SKA Yüksek	54	161	74.9
Genel Oran			70.7

*eşik değeri 0.50

Çizelge 3.7. Geliştirilmiş modelin sınıflandırma tablosu*

	Beklenen		
Gözlenen	SKA Düşük	SKA Yüksek	Doğru Sınıflandırma Oranı (%)
SKA Düşük	158	60	72.5
SKA Yüksek	60	155	72.1
Genel Oran			72.3

*eşik değeri 0.50

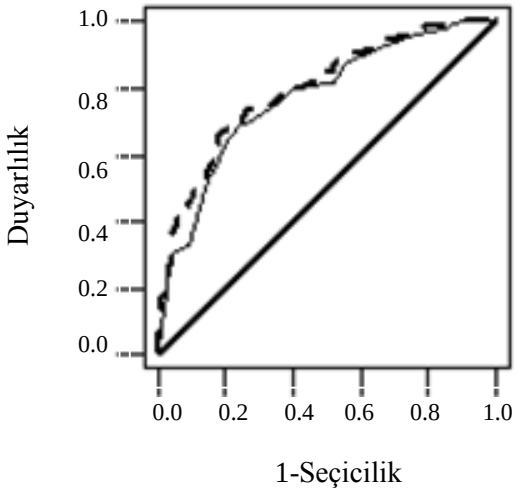
3.3. Klasik Performans Ölçüleri Analiz Sonuçları

Referans model ile doğum ağırlığı değişkeninin eklendiği geliştirilmiş modelin performansları klasik ölçüler ile değerlendirilmiştir ve Çizelge 3.8'de sonuçları verilmiştir. Sonuçlara göre, doğum ağırlığı belirtecinin eklendiği geliştirilmiş modelin Brier scaled değeri %22.6'dan %25.8'e, Nagelkerke R^2 ise %29.3'den %33.8'e yükselmiştir. Aynı zamanda AUC değerinde yaklaşık 0.021'lik bir artış olmuş ve modelin ayırt edebilirliği de 0.033 birim artmıştır.

Çizelge 3.8. Klasik performans ölçülerinin sonuçları

Performans Ölçüleri	Referans Model	Geliştirilmiş Model
Klasik Ölçüler		
Brier	0.193	0.185
Brier Scaled	22.6%	25.8%
Nagelkerke R^2	29.3%	33.8%
ROC-AUC	0.772 [0.729 - 0.816]	0.793 [0.751 - 0.834]
Ayırt Etme Eğimi	0.227	0.260
Kalibrasyon		
H-L Uyum İyiliği Testi	Ki-kare 6.075	Ki-kare 8.213
	p=0.639	p=0.413

Öncelikle her bir tahmin modelinin, SKA'sı düşük ya da yüksek olan kuzuları ayırt edilebilirliği her iki modelin AUC değerleri hesaplanarak değerlendirilmiştir Şekil 3.3'de doğum ağrılığı belirtecinin eklenmesi ile AUC değeri artışmıştır. Tip I hata %5 alındığında her iki modelin AUC değerleri arasındaki fark istatistiksel olarak anlamlı bulunmaktadır. Ancak, çalışmanın duyarlılığının arttırılması ve belirtecin kullanılabilirliğinin daha hassas ölçülebilmesi için Tip I hata %1 olarak alınmıştır. Çalışmada iki modelin AUC değerleri arasındaki fark istatistiksel olarak anlamlı değildir (p=0.035).



Şekil 3.3. Referans model ve geliştirilmiş modele ilişkin ROC Eğrisi

Modellerden elde edilen tahmini olasılık değerlerinin 0.30 eşik değerine göre duyarlılık, seçicilik ve doğruluk değerleri Çizelge 3.9'da verilmiştir. Sonuçlara göre,

geliştirilmiş modelin duyarlılık değerinde bir artış gözlenmezken, seçicilik değeri ve modelin doğru sınıflandırma oranı artmıştır.

Çizelge 3.9. Tahmin modellerine ilişkin ROC analiz sonuçları*

Tahmin Modelleri	Duyarlılık(TPR)	Seçicilik(TNR)	Doğruluk
Referans Model	0.908	0.330	0.621
Geliştirilmiş Model	0.908	0.446	0,713

*eşik değeri 0.30

3.4. Literatürde Yeni Önerilen Performans Ölçüleri Analiz Sonuçları

Çizelge 3.10'da referans ve geliştirilmiş modelden elde edilen tahmini olasılık değerlerinin sınıflandırılması verilmiştir.

Çizelge 3.10. Referans ve Geliştirilmiş modelin tahmini olasılıklarının sınıflandırılması

	Risk Kategorisi	Durum		
Referans Model		0	1	Toplam
	Düşük Risk < 30%	71	20	91
	Yüksek Risk ≥ 30%	144	198	342
Geliştirilmiş Model				
	Düşük Risk < 30%	96	20	116
	Yüksek Risk ≥30%	119	198	317

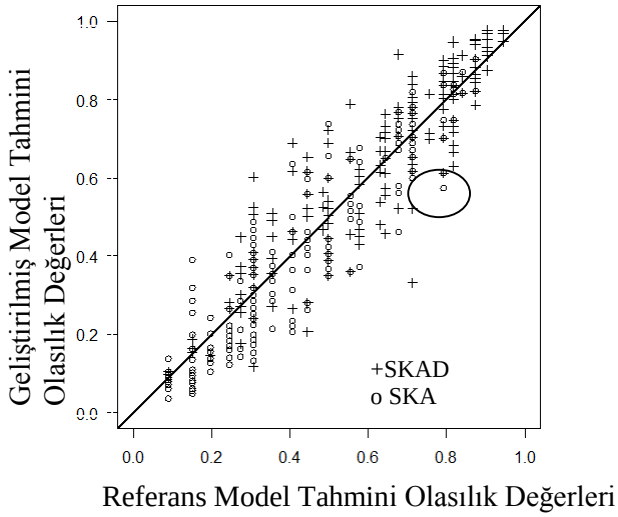
Çizelge 3.10' da elde edilen değerlere göre doğum ağırlığı değişkeninin modele eklenmesi ile gerçekte istenilen SKA'ya sahip olmayan (SKA düşük) yüksek risk grubundaki bireylerin sayısında (198/218= %91-referans model; 198/218= %91-geliştirilmiş model) bir değişiklik meydana gelmezken; gerçekte istenilen SKA'ya sahip olan (SKA yüksek) ama yüksek risk grubunda sınıflandırılmış bireylerin

sayısında (144/215=%67-referans model; 119/215=%55-geliştirilmiş model) ise %12 ‘lik bir azalma sağlanmıştır.

Çizelge 3.11’da tekrar sınıflandırma tablosunda referans modele eklenen doğum ağırlığı değişkeninin risk kategorileri arasında meydana getirdiği değişim gösterilmiştir.

Çizelge 3.11. Tekrar sınıflandırma tablosu

Risk Kategorisi	**Geliştirilmiş Model**		
Referans model	Düşük Risk < 30%	Yüksek Risk ≥30%	Toplam
Düşük Risk < 30%	n=79	n=12	n=91
	13 SKAD (%17)	7 SKAD (%58)	20 SKAD (%22)
Yüksek Risk ≥30%	n=37	n=305	n=342
	7 SKAD (%18)	191 SKAD (%63)	198 SKAD (%58)
Toplam	n=116	n=317	n=433
	20 SKAD(%17)	198 SKAD(%62)	218 SKAD (%50)



Şekil 3.4. Referans ve geliştirilmiş modelin tahmini olasılık değerlerinin saçılım grafiği

Tekrar sınıflandırma tablosunda risk kategorileri arasındaki hareketlilik şekil 3.4'te saçılım grafiği ile de gösterilmiştir. Örneğin; doğum ağırlığı belirtecinin olmadığı referans modelde yüksek risk grubunda gözüken istenilen SKA'ya sahip olan bir birey, doğum ağırlığı belirtecinin dahil olduğu geliştirilmiş modelde çok daha düşük risk grubunda sınıflandırılmıştır.

NRI, SKA'sı düşük ve yüksek olan kuzuların sınıflamadaki iyileşmesini ayrı ayrı hesaplanmasını sağlar. NRI ölçüsüne ait sonuçlar Çizelge 3.12 ve Çizelge 3.13'de verilmiştir.

Çizelge 3.12. İlgilenilen olaya sahip bireylerin sınıflandırma tablosu

	Geliştirilmiş Model		
Referans Model	<30%	≥30%	Toplam
< 30%	13	7	20
≥ 30%	7	191	198
Toplam	20	198	218

Çizelge 3.12'de geliştirilmiş model, 7 bireyi yüksek risk grubunda, 7 bireyi ise düşük risk grubunda sınıflandırmıştır.

Çizelge 3.13. İlenilen olaya sahip olmayan bireylerin sınıflandırma tablosu

	Geliştirilmiş Model		
Referans Model	<30%	≥30%	Toplam
< 30%	66	5	71
≥ 30%	30	114	144
Toplam	96	119	215

Çizelge 3.13'de ise, geliştirilmiş model 5 bireyi yüksek risk grubunda, 30 bireyi ise düşük risk grubunda sınıflandırmıştır. Buna göre NRI,

$$NRI = [P(yukarı \mid D_1) - P(aşağı \mid D_1)] + [P(aşağı \mid D_0) - P(yukarı \mid D_0)]$$
$$NRI = (7-7)/218 + (30-5)/215$$
$$= 0 + 0.116 = 0.116$$

şeklinde hesaplanır. Dolayısıyla, geliştirilmiş model ile istenilen SKA'nın altında (SKA düşük) ağırlığa sahip olan bireylerin sınıflandırılmasında bir iyileşme

sağlanmazken, istenilen SKA'da (SKA yüksek) olan bireylerin yüksek risk grubundan düşük risk grubuna geçme olasılığı %11.6 artmıştır.

Kategoriden bağımsız IDI için modellerin ayırt etme eğimleri hesaplanır. Referans model için, istenilen SKA'da (SKA'sı yüksek) olan kuzuların ortalama tahmini olasılık değeri %61.6 iken, istenilen SKA'nın altında (SKA'sı düşük) ağırlığa sahip kuzuların ortalama tahmini olasılık değeri %38.9'dur. Referans modelin ayırt etme eğimi %22.7'dir. Doğum ağırlığı belirtecini içeren geliştirilmiş modelde ise, istenilen SKA'da olan kuzuların ortalama tahmini olasılık değeri %63.2 iken, istenilen SKA'nın altında ağırlığa sahip bireylerin ortalama tahmini olasılık değeri %37.2'dir. Geliştirilmiş modelin ayırt etme eğimi ise %26.0'dır. Modellerin eğimleri arasındaki fark alındığında IDI, %3.3'tür.

4. TARTIŞMA

Son yıllarda, tahmin modellerinin performanslarının özellikle de tahmin modeline yeni eklenen bir belirtecin kullanılabilirliğinin değerlendirilmesi ile ilgili çalışmalara sağlık alanında sıklıkla rastlanmaktadır (Pencina ve ark., 2008; Cui, 2009; Steyerberg ve ark., 2010; Cook ve Paynter, 2011). Tıp literatüründe, tahmin modeline eklenen yeni bir belirtecin klinik yararlılığının değerlendirilmesinde AUC'nin yanı sıra bu çalışmada bahsedilen literatürde yeni önerilen performans ölçüleri NRI ve IDI da kullanılmaktadır.

Ridker ve ark (2007) yapmış oldukları çalışmada, kadınlarda KVH (Kardiyovasküler Hastalık) tahmini için bilinen klinik risk faktörlerini içeren referans model oluşturmuşlar, ardından bu modele, etkisini araştırmak istedikleri "aile öyküsü" ve "C Reaktif Protein (CRP)" belirteçlerini ekleyerek geliştirilmiş model elde etmişlerdir. Yeni belirteçlerin eklenmesi ile AUC değeri 0,805 'den 0,808 değerine küçük bir artış göstermiştir. Tekrar sınıflandırma tablolarını kullanarak risk kategorileri arasındaki değişim incelendiğinde ise, bireylerin %30'nun kategoriler arasında değişiklik gösterdiği hesaplanmıştır. Steyerberg ve ark (2010) yapmış oldukları bir çalışmada, erkeklerde testis kanseri tahmini için bilinen klinik risk faktörlerini içeren bir referans model oluşturmuşlar ve bu modele etkisini araştırmak istedikleri "laktat dehidrogenaz (LDH)" belirtecini ekleyerek geliştirilmiş model elde etmişlerdir. Yeni belirtecin eklenmesi ile AUC değeri 0.82'den 0.84 değerine küçük bir artış göstermiştir. Ancak, yeni önerilen performans ölçüleri ile incelendiğinde, modele LDH belirteci eklendiğinde testis kanseri olanların sağlıklılara göre bir üst risk kategorisine geçme olasılığı, bir alt kategoriye geçme olasılığından %7.7 daha fazla olduğu ve modelin hasta ve sağlıklıları doğru ayırt edebilirliğinin de %4 arttığı hesaplanmıştır.

Bu çalışmada ise, SKA'sı düşük ve yüksek olan kuzuların sınıflandırılmasında doğum ağırlığı belirtecinin model performansına katkısı değerlendirilmiştir. Genotip, cinsiyet, doğum tipi ve doğum yılı gibi risk faktörlerini içeren referans modele doğum ağırlığı belirtecinin eklenmesi ile geliştirilmiş model elde edilmiştir. Referans modele

doğum ağırlığı belirtecinin eklenmesi ile elde edilen geliştirilmiş modelin Ki-Kare değeri 107.42 den 126.49 değerine yükselmiş ve bu artış istatistiksel açıdan anlamlı bulunmuştur ($p<0.001$).

Elde edilen bu tahmin modellerinin performansları ilk önce klasik performans ölçüleri ile değerlendirilmiştir (Çizelge 3.8). Modele doğum ağırlığı belirtecinin eklenmesi ile modelin Brier scaled ve Nagelkerke R^2 değerleri sırasıyla %25.8 ile %33.8'e yükselmiştir. Modellerin kalibrasyonu Hosmer-Lemeshov uyum iyiliği testi ile değerlendirilmiştir ve elde edilen sonuçlara göre, değişkenlerin modele uyumlu olduğu sonucuna varılmıştır. AUC değerleri referans model ve doğum ağırlığının eklendiği geliştirilmiş model için sırası ile 0.772 ve 0.793 olarak hesaplanmıştır. Buna göre doğum ağırlığı belirtecinin modelin ayırt edebilirliğine katkısı %2.1'dir. AUC değerleri arasındaki bu fark istatistiksel açıdan anlamlı bulunmamıştır ($p=0.035$). Fakat doğum ağırlığı değişkeni, tahmin modeli elde edilirken istatistiksel olarak anlamlı bulunmuş ve modelde yer almıştır.

Ayrıca geliştirilmiş modelin duyarlılık değeri değişmezken, seçicilik değeri 0.330'dan 0.446'ya, doğruluk ise 0.621'den 0.713 değerine yükselmiştir. Dolayısıyla, doğum ağırlığı belirtecinin modele eklenmesi ile SKA'sı düşük ve yüksek olan kuzuların sınıflandırılmasında doğru pozitiflerin oranında bir değişim olmazken, doğru negatiflerin oranı artmıştır. Başka bir ifade ile geliştirilmiş model, gerçekte istenen olaya sahip (SKA'sı yüksek olan) bireyleri ayırt edebilen bir modeldir. Aynı zamanda modelin doğru sınıflandırma oranı da %9.2 artmıştır.

İkinci olarak model performansları özellikle AUC ölçüsüne alternatif olarak literatürde önerilen ölçüler ile değerlendirilmiştir. Tekrar sınıflandırma tablolarını kullanarak SKA'sı düşük ve yüksek olan kuzuların risk kategorileri arasındaki hareketlilik Çizelge 3.11'de daha ayrıntılı olarak gösterilmiştir. Referans model 91 kuzuyu düşük risk grubunda sınıflandırmış ve bunların 20 'sinde SKA'sı gerçekte düşük olan kuzuya rastlanmıştır. Geliştirilmiş model ise, 116 kuzuyu düşük risk grubunda sınıflandırmış ve bunların 20'sinde SKA'sı gerçekte düşük olan kuzuya rastlanmıştır. Böylece, geliştirilmiş modelin hatalı sınıflandırma olasılığı %22'den %17'ye düşmüştür. Benzer şekilde referans model 342 kuzuyu yüksek risk grubunda

sınıflandırmış ve bunların da toplam 198'inde SKA'sı gerçekte düşük olan kuzuya rastlanmıştır. Geliştirilmiş model ise, 317 kuzuyu yüksek risk grubunda sınıflandırmış ve bunların 198'inde SKA'sı gerçekte düşük olan kuzuya rastlanmıştır. Diğer bir ifade ile, geliştirilmiş modelin doğru sınıflandırma olasılığı %58'den %62'ye yükselmiştir.

Risk kategorileri arasında meydana gelen değişimi yüzde olarak ifade edersek, modele doğum ağırlığı belirtecinin eklenmesi ile SKA'sı gerçekte düşük olanların %3.2'si yüksek risk grubunda, %3.2'si düşük risk grubunda sınıflandırılmıştır. İstenilen SKA'da olan kuzuların, %2.3'ü yüksek risk grubunda sınıflandırılırken, %13.9'u düşük risk grubunda sınıflandırılmıştır. Sonuç olarak doğum ağırlığının modele eklenmesi ile SKA'sı düşük olan kuzuların risk grupları arasında bir iyileşme olmazken, SKA'sı yüksek olan kuzuların bir üst risk grubundan bir alt risk grubuna geçme olasılığı artmıştır. Risk kategorileri arasında meydana gelen değişim indeksi NRI %11.6 ($p<0.001$) olarak bulunmuştur. Böylece doğum ağırlığı belirteci eklendiğinde, istenilen SKA'da olan bireylerin düşük risk kategorisine geçme olasılığı, yüksek risk kategorisine geçme olasılığından %11.6 daha fazladır. Dolayısıyla, doğum ağrılığı belirtecinin modele eklenmesi ile aşağı doğru bir hareketlilik sağlanmıştır. Bu bağlamda, geliştirilmiş modelin referans modele göre performansının daha iyi olduğu söylenebilir.

IDI değeri ise %3.3 ($p<0.001$) olarak bulunmuştur. Modele doğum ağırlığı belirtecinin eklenmesi ile istenilen SKA'ya sahip olan ve olmayan kuzuların ortalama tahmini olasılık değerleri arasındaki ortalama risk farkı %3.3 artmıştır.

5. SONUÇ VE ÖNERİLER

Bu çalışmada, bilinen risk faktörlerini içeren referans modele yeni bir belirteç eklenerek belirtecin modeldeki performansı incelenmek istenmiştir. Bu belirtecin, AUC da meydana getirdiği küçük bir artışın, literatürde yeni önerilen performans ölçüleri ile değerlendirildiğinde çok daha önemli bir iyileşmeye neden olabileceği gösterilmeye çalışılmıştır. Bu nedenle öncelikle tahmin modellerinin performanslarının değerlendirilmesinde kullanılan klasik ölçüler ile literatürde yeni önerilen ölçülerin özellikleri tanıtılmıştır. Daha sonra, literatürde yeni önerilen ölçülerin AUC ile olan ilişkisi üzerinde durulmuştur. Bu çalışmada, kuzularda büyümede SKA'ya etki eden risk faktörlerinin oluşturduğu modele eklenen doğum ağırlığı belirtecinin model performansına katkısı klasik ölçülerden Brier Skoru, Neglarkere R^2, ROC-AUC ve yeni ölçülerden NRI ve IDI ile değerlendirilmiştir.

Koyun yetiştiriciliği ile ilişkilendirilebilecek risk faktörlerinin bilinmesi, yetiştiriciler için çok önemlidir. Bu çalışmada, kuzularda büyümede etkili olan risk faktörlerinin yanı sıra doğum ağırlığı değişkeninin SKA'ı düşük ve yüksek olan kuzuların sınıflandırılmasına etkisi incelenmiştir. Böylece, sürüde istenilen SKA'ya sahip olan ve olmayan kuzuların ayırt edilmesi sağlanabilir ya da bu duruma etkisi olabilecek değişkenler kontrol altına alınabilir.

Sonuç olarak, bu tez kapsamında anlatılan yöntemlerin uygulamalı alanlarda kullanılmaları adına bazı avantajları mevcuttur. Bunlar;

- Literatürde yeni geliştirilen bu performans ölçüleri (NRI ve IDI), klasik yaklaşıma (ROC-AUC) nazaran daha duyarlı sonuçlar üretebilmektedir. Uygulamada, tahmin modellerinin sınıflandırmadaki doğruluğunun ölçülmesinde yaygın olarak kullanılan yöntem ROC eğrisi altında kalan alandır. Fakat, özellikle etkisi araştırılmak istenen yeni bir belirtecin modelde yer alıp almamasına karar verirken AUC yanı sıra, NRI ve IDI yöntemlerinin de dikkate alınması gerektiği literatürde yapılan çalışmalarda da belirtilmiştir. Bu yöntemler, riskleri kategorilere ayırır ve modelin bireyleri yüksek ve düşük risk kategorilerine daha

hassas ayırt edebilmesini sağlar. Aynı zamanda risk kategorileri arasında meydana gelen hareketliliğin yönü ile ilgili bilgi verir.

- Literatürde daha çok klinik çalışmalarda, hastalık tanı ve teşhisinde yeni bir belirteç etkisi araştırılmak istenildiğinde kullanılan bu yöntemlerin, yetiştiricilik alanında da kullanılabileceği bu çalışma ile gösterilmiştir. Özellikle modellerinin geliştirilmesi ile ilgili sonraki çalışmalar için örnek teşkil etmektedir. Etkisi araştırılmak istenen ya da yeni geliştirilen bir belirtecin model performansına katkısının değerlendirilmesinde ve mevcut modelin geliştirilmesinde kullanılan yeni yöntemlerin tanıtılması ileriye yönelik yapılacak çalışmalarda araştırmacılara alternatif değerlendirme araçları sunması açısından araştırmayı orijinal kılmaktadır. Bu bağlamda çalışma, analizlerin güvenilirliklerinin arttırılmasının yanında deneme için ayrılan zaman ve kaynağın da israf edilmesinin önüne geçilebilecek ve aynı zamanda bundan sonraki çalışmalara bir temel oluşturacaktır.

Bununla birlikte, NRI ile ilgili bir dezavantajdan bahsetmek gerekirse, NRI, eşik değerinin ve buna bağlı olarak oluşturulan risk kategorilerinin seçiminden etkilenir. Bu yüzden, model performansının değerlendirilmesinde özellikle NRI'nın hesaplanmasında eşik değerinin seçimi ve buna ilişkin oluşturulan risk kategorileri çok önemli bir role sahiptir. Bu yüzden çalışmalarda öznelliği önlemek için daha önceden önerilmiş olan, anlamlı ve kabul edilmiş eşik değerleri ve risk kategorileri ile değerlendirilme yapılması gerekir. Daha önceden belirli risk kategorileri olmayan çalışmalar için kategoriden bağımsız IDI alternatif bir yol olarak önerilebilir.

Bu yeni ölçüler, daha çok ilgilenilen olaya sahip olan ve olmayan bireylerden hangilerinin doğru olarak sınıflandırıldığının tespit ederek o bireyleri referans model ve yeni belirteç değerini kullanarak elde edilen model ile açıklamaya çalışır. Bu açıdan ileriki çalışmalarda ordinal lojisitik regresyon, Cox regresyon gibi farklı tahmin modelleri için bu yöntemler kullanılarak yeni belirteç etkisi ve modellerin performansları değerlendirilebilir.

Yeni bir belirtecin kullanılabilirliğinin olmasının yanı sıra mali açıdan da uygun olması istenir. İleriye dönük yapılacak çalışmalarda, net fayda eğrileri ya da buna

ilişkin tahminler ile yeni bir belirtecin maliyet etkinliği belirlenerek çalışma genişletilebilir. Buna ilaveten benzetim çalışması yapılarak model performansında meydana gelen çok küçük ve çok büyük değişimler klasik ve yeni önerilen ölçüler ile değerlendirilerek farklı örneklem genişlikleri için sonuçlar karşılaştırılabilir.

ÖZET

Kuzularda Büyüme Örneğinde Tahmin Modellerinin Performanslarının İncelenmesi

Bu çalışmanın amacı, tahmin modellerinin performanslarının değerlendirilmesinde kullanılan klasik yöntemler açıklayıcılık katsayısı, Birer skor, eğri altında kalan alan (AUC) ve bu yöntemlere alternatif olarak sunulan iki yeni yöntem Net Tekrar Sınıflandırma İyileştirmesi (NRI) ve Bütünleşik Ayrımsama İyileştirmesi (IDI) ile doğum ağırlığı değişkeninin model performansına etkisinin incelenmesidir.

Çalışmanın materyalini 1996-1997 yılları arasında Sivas-Ulaş Tarım İşletmesindeki 433 kuzunun büyüme özelliğine ait veriler oluşturmuştur. İstenilen SKA'ya sahip olan ve olmayan kuzuların sınıflandırılmasında doğum ağırlığının model performansına etkisi incelenmiştir. Öncelikle kuzularda büyümeye etki eden risk faktörleri genotip, doğum yılı, cinsiyet ve doğum tipi değişkenlerini içeren referans model oluşturulmuştur. Daha sonra bu modele etkisi araştırılacak doğum ağırlığı belirteci eklenerek geliştirilmiş model elde edilmiştir. Çalışmada lojistik regresyon analizi ve risk faktörlerinin hesaplanması SPSS for Windows 14.01 (Lisans No:9869264) paket programı ile yapılmıştır. Model performans ölçülerinin hesaplanması ve model performanslarının karşılaştırılması ile ilgili analizler ise R (Version i386 3.1.3) programında PredictABEL, Hmisc, ROCR ve Clinfun paketleri yardımıyla değerlendirilmiştir.

Doğum ağırlığı değişkeninin eklenmesi ile AUC değeri 0.772'den 0.793'e yükselmiştir. Buna göre doğum ağırlığı belirtecinin modelin ayırt edebilirliğine katkısı AUC'ye göre %2.1'dir. NRI %11.6 ($p<0.001$) olarak bulunmuştur. Böylece doğum ağırlığı belirteci eklendiğinde, istenilen SKA'da olan bireylerin düşük risk kategorisine geçme olasılığı, yüksek risk kategorisine geçme olasılığından %11.6 daha fazladır. Kategoriden bağımsız IDI ise %3.3 ($p<0.001$) olarak hesaplanmıştır. Yani, istenilen SKA'ya sahip olan ve olmayan kuzuların ortalama tahmini olasılık değerleri arasındaki ortalama risk farkı %3.3 artmıştır.

Sonuç olarak NRI, doğum ağırlığı belirtecinin etkisini risk kategorileri bazında meydana gelen değişimi ölçerek AUC'ye nazaran daha hassas bir şekilde göstermiştir. Literatürde yeni gerliştirilen bu performans ölçüleri (NRI ve IDI) klasik yaklaşıma (ROC-AUC) nazaran daha duyarlı sonuçlar üretmektedirler. Özellikle yeni bir belirtecin kullanılabilirliğinin değerlendirilmesinde AUC 'nin yanı sıra NRI ve IDI da dikkate alınmalıdır.

Anahtar Kelimeler: AUC, belirteç, IDI, model performans ölçüleri, NRI, risk tahmini, tekrar sınıflandırma.

SUMMARY

Evaluating the Performance of Prediction Models in the Case of Growth of Lambs

The purpose of this study is to compare the properties of traditional measures (the coefficient of determination (R^2), the Brier score, AUC (Area Under The Curve)) and novel measures (Net Reclassification Improvement (NRI) and Integrated Discrimination Improvement (IDI)) presented as alternatives to evaluate the predictive performance of birth weight.

The research material consists of data related to the growth properties of 433 lambs of Ulaş Tarım İşletmesi (Ulaş Agricultural Enterprise) located in Sivas province of Turkey between the periods 1996 and 1997. The influence of birth weight on model performance in the classification of the lambs with required weaning weight (WW) and the lambs without required WW was investigated. Firstly, a reference model involving risk factors influential on the growth of lambs such as genotype, birth date, gender and year of birth was formed. Then the marker of birth weight whose effect would be investigated was added to this model. Logistic regression analysis was applied and risk factors were calculated with SPSS for Windows 14.01 (License No: 9869264). Analyses on calculation of model performance measures and comparison of model performance were done with R (Version i386 3.1.3) via PredictABEL, Hmisc, ROCR, and clinfun packages.

When the variable of birth weight was added, the AUC value increased from 0.772 to 0.793. Thus, the contribution of the marker of birth weight to the distinguishability of the model was 2.1% more relative to AUC. When the influence of the marker of birth weight was investigated via NRI, distinguishability was seen to be higher by 11.6% ($p<0.001$). Regardless of categories, IDI was found to be 3.3% ($p<0.001$). In other words, the difference between the average estimated probability values of the lambs with required WW and the lambs without required WW increased by 3.3%.

Consequently, NRI performance measure indicated the influence of the marker of birth weight more precisely than AUC by measuring the improvements on the basis of risk categories. In this regard, it can be said that those newly developed performance indices (NRI and IDI) yielded more sensitive results in comparison to the classical approach (ROC-AUC). In particular the availability of a new marker of evaluation, AUC's as well as the NRI and IDI should be considered.

Key Words: AUC, marker, IDI, model performance measures, NRI, risk prediction, reclassification

KAYNAKLAR

AKÇAPINAR, H., ÖZBEYAZ, C., ÜNAL, N., AVCI, M. (2000). Kuzu Eti Üretimine Uygun Ana ve Baba Hatlarının Geliştirilmesinde Akkaraman, Sakız ve Kıvırcık Koyun Irklarından Yararlanma Imkanları I . Akkaraman Koyunlarda Döl Verimi , Akkaraman , Sakiz X Akkaraman F_1 ve Kıvırcık X Akkaraman F_1 Kuzularda Yaşama Gücü. *Turk J.Vet.Anim.Sci.,* **24:** pp.: 71-79.

AKÇAPINAR, H., ÖZBEYAZ, C. (1999). Hayvan Yetiştiriciliği Temel Kariyer. *2. Baskı. Ankara: Isbn: 975 – 96978 – 0 -7.*

AKÇAPINAR, H., ÜNAL, N., ATASOY F. (2005). The Effects of Early Age Mating on Some Production Traits of Bafra (Chios X Karakaya B_1) Sheep. *TÜBİTAK Türk Veterinerlik ve Hayvancılık Dergisi,***29(3):** pp.:531-536

AKÇAY, A. (2009). Broiler Coccidiosis'inde Risk Faktörlerinin Lojistik Regresyon Analizi ile Belirlenmesi. *Doktora Tezi. Ankara Üniversitesi Sağlik Bilimleri Enstitüsü.*

AKÇAY, A., ERTUĞRUL, O., GÜRCAN, I.S., KARAER, Z. (2011). Quantification of Risk Factors of Coccidiosis in Broilers by Using Lojistic Regression Analysis. *Ankara Üniv Vet Fak Derg*, **58:** pp:.195-202.

ALTIN, T., KARACA, O. & CEMAL, I. (2003). Sütten Kesim Yaşının Koyunlarda Süt Verimi ve Kuzularda Büyüme Üzerine Etkisi. *Yyü Ziraat Fakültesi Tarim Bilimleri Dergisi (J.Agric.Sci),* **13(2)**: pp.: 103-111.

ALPAR, C.R. (2011).Uygulamalı Çok Değişkenli İstatistiksel Yöntemler. *3.Baskı. Detay Yayıncılık, Ankara.*

ANONİM, TUİK (2015). Hayvancılık İstatistikleri(2001-2013). *Erişim: [http://tuikapp.tuik.gov.tr/]. Erişim Tarihi: 30.05.2015.*

ARSLAN, M., YILMAZ, O., ATEŞ, C. T. (2003). Morkaraman ve Corriedale X Morkaraman (F_1) Kuzularında Büyüme. *Yyü. Vet. Fak. Derg,* **14(1)**: pp.: 46-49.

ATABEY, Ö. (2010). Lojistik Regresyon Modeli ve Geriye Doğru Eliminasyon Yöntemiyle Değişken Seçiminin Hipertansiyon Riski Üzerine Uygulamasında Bootsrap Yöntemi.*Yüksek Lisans Tezi. Gazi Üniversitesi Fen Bilimleri Enstitüsü*.

ATECA, L. B., DOMBROWSKI, S. C., SILVERSTEIN, D. C. (2015). Survival Analysis of Critically İll Dogs with Hypotension with or without Hyperlactatemia: 67 Cases (2006–2011). *Javma,* **246(1):** pp.: 100-104.

CEYHAN, A., ERDOĞAN, I., SEZENLER, T. (2007). Gen Kaynaği Olarak Korunan Kıvırcık , Gökçeada Ve Sakız Koyun Irklarının Bazı Verim Özellikleri. *Tekirdağ Ziraat Fakültesi Dergisi,* **4(2)**: pp.: 211-218.

COOK, N.R., BURING, J. E., RIDKER, P. M. (2006). The Effect of Including C-Reactive Protein in Cardiovascular Risk Prediction Models For Women. *Ann Intern Med,* **145:** pp.: 21-9.

COOK, N. R., RIDKER, P. M. (2009). Advances in Measuring the Effect of Individual Predictors of Cardiovascular Risk: The Role of Reclassification Measures. *Annals of Internal Medicine,* **150(11):** pp.: 795-802.

COOK, N. R. (2007). Use And Misuse of The Receiver Operating Characteristic Curve in Risk Prediction. *Circulation,* **115(7):** pp.: 928-35.

COOK, N. R., PAYNTER, N. P. (2011). Performance of Reclassification Statistics in Comparing Risk Prediction Models. *Biometrical Journal. Biometrische Zeitschrift,* **53(2):** pp.: 237-58.

CUI, J. (2009). Overview of Risk Prediction Models in Cardiovascular Disease Research. *Annals Of Epidemiology*, **19(10):** pp.: 711-7.

ÇOBAN, Ö., TÜZEMEN, N. (2007). Siyah Alaca ve Esmer İneklerde Subklinik Mastitis için Risk Faktörleri Giriş Materyal ve Metot Bulgular. *Uludağ Üniversitesi J.Fac.Vet.Med. ,* **26(1-2):** pp.: 27-31.

DELONG, E. R., DELONG, D. M., CLARKE-PEARSON, D. L. (1988). Comparing The Areas Under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics*, **44(3):** pp.: 837-45.

DEMIR, H. (2002). Kıvırcık ve Sakız X Kıvırcık Melezi (F_1) Koyunlarin Döl Verimi ve Kuzularin Yaşama Gücü. *Istanbul Üniversitesi Vet. Fak. Derg,* **28(1):** pp.: 155-161.

DIRICAN, A. (2001). Tanı Testi Performanslarının Değerlendirilmesi ve Kıyaslanması. *Cerrahpaşa J Med ,* **32(1):** pp.: 25-30.

ELHAN, A. (1997). Lojistik Regresyon Analizinin İncelenmesi ve Tıpta Bir Uygulaması. *Yüksek Lisans Tezi. Ankara Üniversitesi Sağlik Bilimleri Enstitüsü. .*

GU, W., PEPE, M. (2009). Measures to Summarize and Compare the Predictive Capacity of Markers, *Uw Biostatistics Working Paper Series University Of Washington.* **p.:**342.

GÜNAYDIN, G. (2009). Koyun Yetiştiriciliğinin Ekonomi Politiği. *U.Ü.Ziraat Fakültesi Dergisi,* **23(2):** pp.: 15-32.

HOSMER, D., LEMESHOW, S. (1982). A Review of Goodness of Fit Statistics For Use in the Develepment of Lojistic Regression. Models. *Am. J. Epidemiol,* **115:** pp.: 92-106.

HOSMER, D., LEMESHOW, S. (2000). Applied Logistic Regression. New York: Wiley And Sons Inc; 2nd Ed.

JANES, H., PEPE, M. S., GU, W. (2008). Assessing yhe Value of Risk Predictions By Using Risk Stratification Tables. *Annals of Internal Medicine,* **149(10)**: pp.: 751.

KENNEY, E. M., ROZANSKI, E. A, RUSH, J. E., DELAFORCADE-BURESS, A. M., BERG, J. R., SILVERSTEIN, D. C.,SHAW, S. P. (2010). Association Between Outcome and Organ System Dysfunction in Dogs With Sepsis: 114 Cases (2003-2007). *Journal Of The American Veterinary Medical Association,* **236(1)**: pp.: 83-87.

KERR, K. F., MCCLELLAND, R. L., BROWN, E. R., LUMLEY, T. (2011). Evaluating the Incremental Value of New Biomarkers with Integrated Discrimination Improvement. *American Journal Of Epidemiology*, **174(3):** pp.:364-374.

KING, L. G., WOHL, J. S., MANNING, A. M., HACKNER, S. G., RAFFE, M. R., MAISLIN, G. (2001). Evaluation of The Survival Prediction Index as a Model of Risk Stratification For Clinical Research in Dogs Admitted to Intensive Care Units at Four Locations. *American Journal Of Veterinary Research,* **62(6):** pp.: 948-954.

KLEINBAUM, D. G., KLEIN, M. (2010) .Logistic Regression: A Self-Learning Text (Statistics For Biology And Health). Third Edition. USA: Springer Sciences Business Media.

KÖKSAL, B. (2011). Regresyon Analizinde Roc Eğrisi Kestirimi ile Model Seçimi. *Yüksek Lisans Tezi. Marmara Üniversitesi Sosyal Bilimleri Enstitüsü.*

LAUER, M., POTHIER, C.E., MAGID, D.J., SMITH, S.S., KATTAN, M.,W. (2007). An Externally Validated Model For Predicting Long-Term Survival After Exercise Treadmill Testing in Patients With Suspected Coronary Artery Disease and a Normal Electrocardiogram.*Ann Intern Med,* **147:** pp.: 821-8.

MAVİLİ, S. (2012). Romanov X Karayaka F_1 Melezi Kuzuların Büyüme Özelliklerinin Belirlenmesi. Istanbul: *Yüksek Lisans Tezi, Gaziosmanpaşa Üniversitesi Fen Bilimleri Enstitüsü.*

MEIGS, J.B., SHRADER, P, SULLIVAN, L.M., MCATEER, J.B., FOX, C.S., DUPUIS, J., MANNING, A.K., FLOREZ, J.C., WILSON, P.W.F., D'AGOSTINO, R.B., CUPPLES, L.A. (2008). Genotype Score n Addition To Common Risk Factors For Prediction of Type 2 Diabetes. *Nejm,* **2219(359):** pp.: 2208.

MOONS, K., HARREL, F. (2003). Sensitivity and Specificity Should Be De-Emphasized in Diagnostic Accuracy Studies. *Acad. Radiol*, **10:** pp.: 670–672.

OĞUZOĞLU, T.Ç., MUZ, D., TİMURKAN, M.Ö., MARAL, N., GURCAN, I.S. (2013). Prevalences of Feline Coronavirus (FCoV), Feline Leukaemia Virus (FeLV), Feline Immunodeficiency Virus (FIV) and Feline Parvovirus (FPV) Among Domestic Cats in Ankara, Turkey, Revue Med. Vet., **164(11):** pp:.511-516.

ÖZTUNA, D., ATEŞ, C., GÜLTEKIN, S., GENÇ, Y. (2011). Kümelenmiş Verilerde Işlem Karakteristiği Eğrisi (Ike) Altında Kalan Alanın Tahmini. *Turkiye Klinikleri J Biostat ,* **3(2):** pp.: 57-62.

PARIKH,N.I., PENCINA, M.J., WANG, T.J., BENJAMIN, E.J., LANIER, K.J., LEVY, D. (2008). Arisk Score For Predicting Near-Term Incidence Of Hypertension: The Framingham Heart Study.*Ann Intern Med. ,* **148**: pp.: 102-10.

PENCINA, M. J., D'AGOSTINO, R. B., D'AGOSTINO, R. B., VASAN, R. S. (2008). Evaluating The Added Predictive Ability of a New Marker: From Area Under The Roc Curve to Reclassification and Beyond. *Statistics in Medicine,* **27(2):** pp.: 157-172.

PENCINA, M. J., D'AGOSTINO, R. B., STEYERBERG, E. W. (2011). Extensions of Net Reclassification Improvement Calculations to Measure Usefulness of New Biomarkers. *Statistics in Medicine,* **30(1):** pp. 11-21.

PEPE, M. S., FENG, Z., HUANG, Y., LONGTON, G., PRENTICE, R., THOMPSON, I. M., ZHENG, Y. (2008). Integrating the Predictiveness of a Marker with Its Performance as a Classifier. *American Journal Of Epidemiology,* **167(3):** pp.: 362-8.

PEPE, M. S., JANES, H., LONGTON, G., LEISENRING, W., NEWCOMB, P. (2003). Limitations of the Odds Ratio in Gauging the Performance of A Diagnostic, Prognostic, or Screening Marker. *American Journal of Epidemiology,* **159:** pp.: 882-890.

RIDKER, P.M., BURING. J.E., RIFAI, N., COOK, N.R. (2007). Development And Validation of Improved Algorithms for the Assessment of Global

Cardiovascular Risk in Women: The Reynolds Risk Score. *Jama,* **297(6):** pp.: 611-619.

SEZENLER, T., SOYSAL, D., YILDRIR M., YÜKSEL M.A., CEYHAN A., YAMAN Y., ERDOĞAN İ., KARADAĞ O. (2013). Karacabey Merinos Koyunların Kuzu Verimi ve Kuzularda Büyüme Performansı Üzerine Bazı Çevre Faktörlerinin Etkisi. *Namik Kemal Üniversitesi Tekirdağ Ziraat Fakültesi Dergisi Journal of Tekirdag Agricultural Faculty,* **10(1):** pp.: 40-47.

STEYERBERG, E. (2009). Clinical Prediction Models A Pratical Aprroach To Development, Validation, And Updating. New York: Springer.

STEYERBERG, E. W., PENCINA, M. J., LINGSMA, H. F., KATTAN, M. W., VICKERS, A. J., VAN CALSTER, B. (2012). Assessing the Incremental Value of Diagnostic and Prognostic Markers: A Review and Illustration. *European Journal of Clinical Investigation,* **42(2):** pp.: 216-28.

STEYERBERG, E. W., VICKERS, A. J., COOK, N. R., GERDS, T., GONEN, M., OBUCHOWSKI, N., PENCINA, M.J.,KATTAN, M. W. (2010). Assessing The Performance of Prediction Models: A Framework for Traditional and Novel Measures. *Epidemiology,* **21(1):** pp.: 128-38.

STEYERBERG, E. W.,CALSTER, B.V., PENCINA, M.J. (2011). Performance Measures for Prediction Models and Markers: Evaluation of Predictions and Classifications, Rev.Esp Cardiol, **248:** pp.:7. doi:10.1016/j.recesp.2011.04.017

TICE, J., CUMMINGS, S.R., SMITH-BINDMAN, R., ICHIKAWA, L., BARLOW, W.E., KERLIKOWSKE, K. (2008). Using Clinical Factors and Mammographic Breast Density to Estimate Breast Cancer Risk: Development and Validation af A New Predictive Model. *Ann Intern Med,* **148:** pp.: 337-347.

TOMAK, L. BEK, Y. (2010). İşlem Karakteristlik Eğrisi Analizi ve Eğri Altında Kalan Alanların Karşılaştırılması. *Journal Of Experimental And Clinical Medicine,* **27:** pp.: 58-60.

TZOULAKI, I., LIBEROPOULOS, G. and IOANNIDIS, J. P. A. (2011). Use of Reclassification for Assessment of Improved Prediction: An Empirical Evaluation. *International Journal of Epidemiology,* **40(4):** pp.: 1094-105.

ÜNAL, N. (2002). Akkaraman ve Sakız x Akkaraman F_1 Kuzularda Yaşama Gücü, Büyüme ve Bazı Vücut Ölçüleri. *TÜBİTAK Türk Veterinerlik ve Hayvancılık Dergisi*, **26(1):** pp. 109-116.

ÜNAL, N., AKÇAPINAR, H., ATASOY, F., AYTAÇ, M. (2006). Some Reproductive And Growth Traits of Crossbred Genotypes Produced by Crossing Local Sheep Breeds of Kıvırcık X White Karaman and Chios X White Karaman in Steppe Conditions. *Archives of Animal Breeding (Archive für Tierzucht),* **49(1):** pp:.55-63

ÜNAL, N., AKÇAPINAR, H., ATASOY, F., YAKAN, A., UĞURLU, M. (2008). Bafra Koyunlarında Bazı Meme Özellikleri ve Kuzularda Büyüme ile Bu Özelliklerin Farklı Süt Kontrol Yöntemleriyle Tespit Edilen Süt Verimi ve Sağım Özellikleriyle Fenotipik Korelasyonları. *Ankara Üniversitesi Veteriner Fakültesi Dergisi,* **55(2):** pp:.117-124

VAN DER STEEG, W.A., BOEKHOLDT, S.M., STEIN, E.A., EL-HARCHAOUI, K., STROES, E.S., SANDHU, M.S. (2007). Role of The Apolipoprotein B-

Apolipoprotein A-I Ratio in Cardiovascular Risk Assessment: A Case-Control Analysis in Epic-Norfolk.. *Ann Intern Med. ,* **146:** pp.: 840-8.

YAKAN, A., ÜNAL, N., DALCI, M.T., (2012). Ankara Şartlarında Akkaraman, İvesi ve Kıvırcık Irklarında Döl Verimi, Büyüme ve Yaşama Gücü. *Lalahan Hayvancılık Araştırma Enstitüsü Dergisi,* **52(1):** pp:.1-10.

YATES, JF. (1982). External Correspondence: Decomposition of the Mean Probability Score. *Organizational Behavior and Human Performance,* **30:** pp.:132-156.

YILDIZ, G., SOYSAL, M. I., GÜRCAN, E. K. (2009). Determination Of Growth Curve With Different Model In The Karacabey Merinos X Kivircik Crossbred Lambs Raised In Tekirdağ Provinces. *Journal Of Tekirdag Agricultural Faculty,* **6(1):** pp.: 11-19.

YILMAZ, A., ALTINEL, A. (2003). Alman Siyah Başli Etçi Irkinin Baba Hatti Olarak Kullanilmasiyla Elde Edilen Üçlü Kullanma Melezi Kuzular Ile Kivircik Ve Türk Merinosu Kuzularin Büyüme Hızı ve Yaşama Gücü Özellikleri. *Istanbul Üniv. Vet. Fak.Derg; ,* **29(2):** pp.: 213-219.

ZOU, K. H., O'MALLEY, A. J., MAURI, L. (2007). Receiver-Operating Characteristic Analysis for Evaluating Diagnostic Tests and Predictive Models. *Circulation,* **115(5):** pp.: 654-7.

EKLER

Ek-1

R KODLARI :

```
library(ROCR)
library(Hmisc)
library(PredictABEL)
library(rms)
library(clinfun)

#### REFERANS MODEL ####

full<-
lrm(doksan_cat1~cn+genotip+yl+dt,data=veridoksan,x=T,y=T,linear.predictors=T)

### Brier Skor (%)          ###

B <- mean((full$y) * (1-plogis(full$linear.predictors))^2 + (1-full$y) *
plogis(full$linear.predictors)^2)
Bmax <- mean(full$y) * (1-mean(full$y))^2 + (1-mean(full$y)) * mean(full$y)^2
Bscaled <- 1 - B/Bmax

### GELİŞTİRİLMİŞ MODEL ###

fulldogum <- update(full, ~ . + dogum) #### Referans modele doğum ağırlığı değişkeni
eklenmiştir.

### Brier Skor (%) ###

B <- mean((fulldogum$y) * (1-plogis(fulldogum$linear.predictors))^2 +(1-
fulldogum$y) * plogis(fulldogum$linear.predictors)^2)
Bmax <- mean(fulldogum$y) * (1-mean(fulldogumy))^2 + (1-mean(fulldogum$y))
*mean(fulldogum$y)^2
Bscaled <- 1 - B/Bmax

### ROC EĞRİSİ ###

library(ROCR)
par(mfrow = c(2,1), pty='s', mar=c(4.2,4,4,1),cex=1, las=1, font =1, col=1)

### REFERANS MODEL İÇİN ROC EĞRİSİ ###

pred.full <- prediction(plogis(full$linear.predictor), full$y)
perf1 <- performance(pred.full,"tpr","fpr") # ROC Eğirisi plot(perf1, xlab="",
cex.lab=1.2)
lines(x=c(0,1), y=c(0,1), lwd=2)
title("A. Development data, n=433", adj=0)
cutoffs <- c(20,17,14)
points(x=perf1@x.values[[1]][cutoffs], y=perf1@y.values[[1]][cutoffs], pch=1, ps=3)
```

```
labROC <- perf1@alpha.values[[1]][cutoffs]
labROC <- paste(as.character(round(100*labROC)),"%", sep="")
text(x=perf1@x.values[[1]][cutoffs]+.01,y=perf1@y.values[[1]][cutoffs]-0.02,
labels=labROC, adj=0)

### GELİŞTİRİLMİŞ MODEL İÇİN ROC EĞRİSİ ###

pred.fulldogum <- prediction(plogis(fulldogum$linear.predictor), full$y)
perf2 <- performance(pred.fulldogum,"tpr","fpr") # ROC curve
plot(perf2, add=T, lty=2, lwd=2)
# perf1@alpha.values cut-off değerleridir#
 # look for 20, 30, 40% points
 cutoffs <- c(219,189,156)
 points(x=perf2@x.values[[1]][cutoffs], y=perf2@y.values[[1]][cutoffs], pch=2, ps=3)

### ROC AREA TEST (DeLong Testi)###

library(clinfun)
roc.area.test(cbind(full$linear.predictor,fulldogum$linear.predictor),full$y)

### BOXPLOTS ###
par(mfrow = c(3,1), pty='s', mar=c(1.2,4,3,1),cex=0.95, font =1, col=1 )
boxplot(plogis(full$linear.predictors)~full$y,ylab="Predicted risk without
dogum",xlab="Normal",ylim=c(0,1),names=Cs(Düşük, Normal), xaxt='n',
cex.lab=1.1)
boxplot(c(mean(plogis(full$linear.predictors[full$y==0])),
mean(plogis(full$linear.predictors[full$y==1])))~ c(0,1),add=T,boxlty=0, staplelty=0,
medlty=0, medlwd=0, medpch=15, xaxt='n')
mtext(side=3, text = paste("A. Development data, slope=",
round(mean(plogis(full$linear.predictors[full$y==1])) -
mean(plogis(full$linear.predictors[full$y==0])),2),sep=""), adj=0, cex=1.1,line=1.2)
boxplot(plogis(fulldogum$linear.predictors)~fulldogum$y,ylab="Predicted risk with
dogum", xlab="Normal",ylim=c(0,1), xaxt='n',cex.lab=1.1)
boxplot(c(mean(plogis(fulldogum$linear.predictors[fulldogum$y==0])),
mean(plogis(fulldogum$linear.predictors[fulldogum$y==1])))~
c(0,1),add=T,boxlty=0, staplelty=0, medlty=0, medlwd=0, medpch=15, xaxt='n')
mtext(side=3, text = paste("B. Development data, slope=",
round(mean(plogis(fulldogum$linear.predictors[fulldogum$y==1]))
mean(plogis(fulldogum$linear.predictors[fulldogum$y==0])),2),sep=""),
adj=0,cex=1.1, line=1.2)

# ## IDI ###

slopeNo <- mean(plogis(full$linear.predictors[full$y==1])) -
mean(plogis(full$linear.predictors[full$y==0]))
slopeYes <- mean(plogis(fulldogum$linear.predictors[fulldogum$y==1])) -
mean(plogis(fulldogum$linear.predictors[fulldogum$y==0]))
slopeYes - slopeNo

### SCATTERPLOT ###

par(mfrow = c(1,1), pty='s', mar=c(4.2,3,4,1),cex=1)
```

```
plot(x=plogis(full$linear.predictors),y=plogis(fulldogum$linear.predictors),xlab="Pre
dicted risk without dogum", ylab="Predicted risk with dogum",cex.lab=1.2,
cex=1,las=1,pty='s',xlim=c(0,1),ylim=c(0,1),
pch=(full$y+.5)*2)
abline(a=0,b=1, lwd=2)
title("Development data, n=433", adj=0)

########################
### TEKRAR SINIFLANDIRMA TABLOSU (RECLASSIFICATION TABLE) ###

Nodogum <- cut(plogis(full$linear.predictors),breaks = c(0,.3,1), labels = c("Nodogum,
<=30%","No dogum, >30%"))
Yesdogum <- cut(plogis(fulldogum$linear.predictors),breaks = c(0,.3,1), labels
=c("With dogum, <=30%","With dogum, >30%"))
table(Nodogum)
table(Yesdogum)
tabReclas <- table(Nodogum, Yesdogum)
tabReclas
tabReclas[2]/(tabReclas[2]+tabReclas[4]) # reclassified among high risk
tabReclas[3]/(tabReclas[1]+tabReclas[3]) # reclassified among low risk

### ÇIKTI TABLOLARI ###

table(Nodogum, Yesdogum, full$y)



### KATEGORİLER ARASI DEĞİŞİM ###

tab <- table(Nodogum, Yesdogum, full$y )
tab[5:8]/(tab[1:4]+tab[5:8])
table(Nodogum, full$y )
tabNo <- table(Nodogum, full$y )
tabNo[3:4]+tabNo[1:2]
tabNo[3:4]/(tabNo[3:4]+tabNo[1:2])
table(Yesdogum, full$y )
tabYes <- table(Yesdogum, full$y )
tabYes[3:4]+tabYes[1:2]
tabYes[3:4]/(tabYes[3:4]+tabYes[1:2])

### NRI AND IDI CALCULATIONS WITH HARRELL'S FUNCTIONS ###

improveProb(x1=as.numeric(Nodogum)-1, x2=as.numeric(Yesdogum)-1, y=full$y)

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

ÖZGEÇMİŞ

I- Bireysel Bilgiler

Adı	Özlem
Soyadı:	GÜLLÜ
Doğum yeri ve tarihi:	Ankara – 06/02/1983
Uyruğu:	T.C.
Medeni durumu:	Bekar
İletişim adresi ve telefonu:	Ankara Üniversitesi Fen Fakültesi B Blok İstatistik Bölümü, Tandoğan Ankara 0312 212 6720/1184

II- Eğitimi

Yüksek Lisans:	Gazi Üniversitesi Bilişim Enstitüsü Sağlık Bilişimi AD. / 2009
Lisans:	Gazi Üniversitesi Fen Edebiyat Fakültesi İstatistik Bölümü / 2006
Önlisans:	Anadolu Üniversitesi Açıköğretim Fakültesi Halkla İlişkiler Bölümü / 2006
Yabancı dili:	İngilizce

III- Unvanları

Uzman – Ankara Üniversitesi İstatistik Bölümü Uygulamalı İstatistik AD. (2010-)

IV- Meslek Deneyimi

Sistem Programcısı- Yükseköğretim Kurulu (YÖK). (2008-2010)
Uzman – Ankara Üniversitesi İstatistik Bölümü Uygulamalı İstatistik AD. (2010-)

V- Üye Olduğu Bilimsel Kuruluşlar

Biyoistatistik Derneği

VI- Bilimsel İlgi Alanları

Yayınları

Gürcan, I.S., Özen, D., Yamaç, A., **Güllü, Ö.** (2014). Investigation of effect of year and season factors on calving difficulty, using Poisson regression model in Simmental x South Anatolian red crossbred cattle, Ankara Üniv Vet Fak Derg, 61: 55-58.

Özlem GÜLLÜ, Mustafa Agah TEKİNDAL, İlker GÖKÇE, Serdal Kenan KÖSE, Kadir TÜRKÖLMEZ, Nomogramlar ile İstatistiksel Modellerin Yaşam Olasılıklarını Belirlemedeki Performanslarının Metot Karşılaştırma Yöntemleri ile Değerlendirilmesi: Mesane Kanseri Örneği, Turkiye Klinikleri J Biostat 2014;6(1):42-52

Bildirileri

Özlem GÜLLÜ, Mustafa Agah TEKİNDAL, İlker GÖKÇE, Serdal Kenan KÖSE, Kadir TÜRKÖLMEZ, Nomogramlar ve Matematiksel Modelleme Arasındaki Uyumun Metot Karşılaştırma Yöntemleri ile Değerlendirilmesi: Mesane Kanseri Örneği. XIII. Ulusal Biyoistatistik Kongresi Bildiri Özetleri Kitabı, 12-14 Eylül 2011 Kızılcahamam/Ankara, Sayfa 25-27

Doğukan ÖZEN , İ.Safa GÜRCAN ,S.Kenan KÖSE, **Özlem GÜLLÜ.** Tekrarlı başarısızlık durumlarında orantılı riskler tipi model yaklaşımları, XIII. Ulusal Biyoistatistik Kongresi 12-14 Eylül 2011 Kızılcahamam/Ankara.

Doğukan ÖZEN, İ.Safa GÜRCAN**, Özlem GÜLLÜ** (2012). Hayvancılıkta verimli yaşam süresi hesaplamalarında Frailty modellerin uygulanabilirliği, XIV. Ulusal Biyoistatistik Kongresi (Uluslar arası Katılımlı). 4-7Eylül 2012, Kadir Has Kongre Merkezi, Kayseri, Bildiri Özetleri Kitabı Sayfa, 90

İ.Safa GÜRCAN,Ayca BABAK, Doğukan ÖZEN, Aytekin YAMAÇ, **Özlem GÜLLÜ**, İneklerde Eko-Yapı Değişkenler Kullanılarak Kornik Endometritis Derecelerinin ROC Eğirisi Yöntemi ile Belirlenmesi, Uluslararası Katılımlı XIV. Ulusal Biyoistatistik Kongresi, 4-7 EYLÜL 2012, Kadir Has Kongre Merkezi, Kayseri, Bildiri Özetleri Kitabı Sayfa, 89

Özlem GÜLLÜ, Mustafa Agah TEKİNDAL, Dogukan ÖZEN, İ.Safa GÜRCAN, Yasemin GENÇ, Tahmin Modellerinin Performanslarının Degerlendirilmesi**,** Uluslararası Katılımlı XIV. Ulusal Biyoistatistik Kongresi, 4-7 EYLÜL 2012, Kadir Has Kongre Merkezi, Kayseri, Bildiri Özetleri Kitabı Sayfa, 92-93

Mustafa Agah TEKİNDAL, Melike TUNÇ, Ayşe Canan YAZICI, Serdal Kenan KÖSE, **Özlem GÜLLÜ**, Can ATEŞ, Beyza Doğanay ERDOĞAN, Box-Behnken Deneme Düzeni ile Engelli Çocuğa Sahip Annelerin Ruhsal Alan Yaşam Kalitesinin Değerlendirilmesi, Uluslararası Katılımlı 15. Ulusal Biyoistatistik Kongresi Bildiri Özetleri Kitabı, Sayfa 79-80, 20-23 Ağustos 2013, Didim, Aydın

Mustafa Agah TEKİNDAL, **Özlem GÜLLÜ**, Ayşe Canan YAZICI, Serdal Kenan KÖSE, Beyza DOĞANAY ERDOĞAN, Can ATEŞ, Yasemin YAVUZ, Melike TEKİNDAL, Sağdan Sansürlü Verilerin Parametrik, Yarı Parametrik ve Parametrik Olmayan Yöntemlerle Değerlendirilmesi, XVI. Ulusal Biyoistatistik Kongresi, 10-12 Eylül 2014, Side ANTALYA, S. 93

İ.Safa GÜRCAN, Pınar AMBARCIOĞLU, Ufuk Kaya, Doğukan ÖZEN, **Özlem GÜLLÜ,** Comparisons of the Common Used Growth Models on the Pekin Ducks, The 8th Conference of Eastern Mediterranean Region of the International Biometric Society Abstract Book s.26, 11-15 May 2015, Cappadocia, Nevşehir,

Mustafa Agah TEKİNDAL, Ayşe Canan Yazıcı, **Özlem GÜLLÜ**, Beyza Doğanay ERDOĞAN, Yasemin YAVUZ, Ersin ÖĞÜŞ, Evaluatıng Rıght-Censored Data Through Substitution, Parametric, Semi-Parametric, and Non-Parametrıc Methods: A Simulation Study, The 8th Conference of Eastern Mediterranean Region of the International Biometric Society Abstract Book s.131, 11-15 May 2015, Cappadocia, Nevşehir.

Özlem GÜLLÜ, Can Ateş, Mustafa Agah TEKİNDAL, Beyza Doğanay ERDOĞAN, Yasemin YAVUZ, İ. Safa GÜRCAN, Berkay EKİCİ, Evaluating the Improvement Of Model

Performance with The Addition of Mean Platelet Volume Marker in A Cardiovascular Risk Prediction Model, The 8th Conference of Eastern Mediterranean Region of the International Biometric Society Abstract Book s.135, 11-15 May 2015, Cappadocia, Nevşehir.

Özlem GÜLLÜ, Melike TEKİNDAL, Mustafa Agah TEKİNDAL, Ayşe Canan YAZICI, Measuring The Quality of Service in The Health Sector Through Servqual Method: The Case of A Physical Therapy and Rehabilitation Center, The 8th Conference of Eastern Mediterranean Region of the International Biometric Society Abstract Book s.164, 11-15 May 2015, Cappadocia, Nevşehir.

VII- Bilimsel Etkinlikleri

Verdiği Konferans ya da Seminerleri

Kovaryans Analizi ile Varyans Analizinin Karşılaştırılması ve Bir Uygulama
Logaritmik Doğrusal Modeller ve Uygulaması

Eğitim Programı Dışında Aldığı Kurslar

Mixed Models with Medical Applications, Uluslararası Yaz Okulu Kursu. University of Southampton,University of Reading, University of Ege, İzmir, 2013

Bilgisayar Uygulamalı Yapısal Eşitlik Modellemesi Kursu, Eskişehir Osmangazi Üniversitesi, Eskişehir, 2015

Organizasyonunda Katkıda Bulunduğu Bilimsel Toplantılar

XVI. Ulusal Biyoistatistik Kongresi, 10-12 Eylül 2014, Side ANTALYA (Düzenleme Kurulu Üyesi).