

METTİN VURAL

İSTANBUL ÜNİVERSİTESİ SAĞ. BİL. ENST.

DOKTORA TEZİ

İSTANBUL-2020



**T.C.
İSTANBUL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
BİYOİSTATİSTİK VE TIP BİLİŞİMİ PROGRAMI**

**ROC EĞRİSİ ALTI ALAN HESAPLAMA
YÖNTEMLERİNİN KARŞILAŞTIRILMASI**

DR.METİN VURAL

(DOKTORA TEZİ)

**DANIŞMAN: PROF.DR. AHMET DİRİCAN
İSTANBUL TIP FAKÜLTESİ
BİYOİSTATİSTİK ANABİLİM DALI**

İSTANBUL-2020

TEZ ONAYI**DOKTORA TEZİ ONAYI**

İstanbul Üniversitesi Sağlık Bilimleri Enstitüsü İstanbul Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Biyoistatistik ve Tıp Bilişimi Programında Doktora öğrencisi Metin Vural tarafından Prof.Dr.Ahmet Dirican'ın danışmanlığında hazırlanan “ROC Eğrisi Altı Alan Hesaplama Yöntemlerinin Karşılaştırılması” başlıklı tez aşağıdaki jüri üyeleri tarafından 23 /12 /2020 tarihinde yapılan Tez Savunma Sınavında başarılı bulunmuş ve Doktora Tezi olarak kabul edilmiştir.

Jüri Başkanı

Prof.Dr. Eray YURTSEVEN
İ.Ü., İstanbul Tıp Fakültesi
Biyoistatistik Anabilim Dalı

Jüri-Danışman

Prof. Dr. Ahmet DİRİCAN
İ.Ü., İstanbul Tıp Fakültesi
Biyoistatistik Anabilim Dalı

Jüri

Prof. Dr. Necdet SÜT
Trakya Üniversitesi, Tıp Fakültesi
Biyoistatistik Anabilim Dalı

Jüri

Doç.Dr. Hülya GÜL
İ.Ü., İstanbul Tıp Fakültesi
Halk Sağlığı Anabilim Dalı

Jüri

Dr.Öğr.Üyesi Ömer UYSAL
Bezmialem V. Üniversitesi, Tıp Fakültesi
Biyoistatistik Anabilim Dalı

BEYAN

Bu tez çalışmasının kendi çalışmam olduğunu, tezin planlanmasından yazımına kadar bütün safhalarda etik dışı davranışımın olmadığını, bu tezdeki bütün bilgileri akademik ve etik kurallar içinde elde ettiğimi, bu tez çalışmayla elde edilmeyen bütün bilgi ve yorumlara kaynak gösterdiğimi ve bu kaynakları da kaynaklar listesine aldığımı, yine bu tezin çalışılması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığı beyan ederim.

Dr. Metin Vural(İmza)



İTHAF

Canım anneme, babama, kardeşlerime ve hayattaki en değerli varlıklarım kızlarım Lara Dilara ve Talya İlayda'ya ithaf ediyorum.



TEŞEKKÜR

Mesleğimizin en önemli basamaklardan biri olan doktora tezimin her aşamasında yanımda olan ve beni destekleyen, manevi desteğini hep hissettiğim, yardımını esirgemeyen çok sevdiğim ve saygı duyduğum hocalarım Prof. Dr. Ahmet Dirican'a, Prof. Dr. Eray Yurtseven'e ve Dr. Öğr. üyesi Ömer Uysal'a teşekkür ederim.

Her zaman yanımda olan, bugünlere gelmemi sağlayan çok değerli annem Kadriye Vural'a, bu tezin yabancı kaynaklarının çevirisinde büyük katkıları olan kızım Lara Dilara ve bana motivasyon kaynağı olan küçük kızım Talya İlayda'ya çok teşekkür ediyorum.



İÇİNDEKİLER

TEZ ONAYI	II
BEYAN.....	II
İTHAF.....	IV
TEŞEKKÜR.....	V
İÇİNDEKİLER	VI
TABLolar LİSTESİ.....	IX
ŞEKİLLER LİSTESİ	X
DENKLEMLER LİSTESİ.....	XI
SEMBOLLER / KISALTMALAR LİSTESİ	XII
ÖZET	XIII
ABSTRACT.....	XIV
1. GİRİŞ VE AMAÇ.....	1
2. GENEL BİLGİLER	3
2.1. Tanı Testleri (Diagnostic Tests).....	3
2.1.1. Doğru Pozitiflik (True Positivity):	3
2.1.2. Yanlış Negatiflik (False Negativity):	4
2.1.3. Yanlış Pozitiflik (False Positivity):	4
2.1.4. Doğru Negatiflik (True Negativity):	5
2.1.5. Duyarlılık (Sensitivity).....	5
2.1.6. Özgüllük (Specificity)	6
2.1.7. Kestirim Değeri (Predictive Value).....	6
2.1.8. Pozitif Kestirim Değeri (Positive Predictive Value)	7
2.1.9. Negatif Kestirim Değeri (Negative Predictive Value)	7
2.1.10. Doğruluk (Accuracy).....	8
2.1.11. Bayes Teoremi (Bayes Theorem).....	9
2.1.12. Olabilirlik Oranı (likelihood Ratio).....	9
2.1.13. Negatif Olabilirlik Oranı (Negative Likelihood Rate)	11
2.1.14. Pozitif Olabilirlik Oranı (Positive Likelihood Rate)	11
2.1.15. Fagan Nomogramlar	11

3. ROC – RECEIVER OPERATING CHARACTERISTICS.....	13
3.1. ROC Eğrisi Altı Alanı (AUROC – Area Under Receiver Operating Characteristic Curve)	13
3.2. ROC Eğrisi Altı Alanı Hesaplama Yöntemleri.....	16
3.2.1. Monte Carlo Simülasyon yöntemi.....	17
3.2.2. Sabit Dağılım.....	23
3.2.3. Sınırsız Destekli Üçgen Dağılım	27
3.2.4. Gauss Yöntemi ile Hesaplama	29
3.2.5. Sınırlı Destekli Üçgen Dağılım	33
3.2.6. Beta Dağılımları	35
3.3. ROC AUC Yorumlama.....	39
3.4. Parametrik Yaklaşımlar – Binormal	41
3.5. Non-Parametrik ROC Eğrisi	43
3.5.1. Ampirik tahmin yöntemleri	43
3.5.2. Çekirdek yoğunluk tahmin yöntemi (Kernel estimator).....	44
3.6. Bayes Yaklaşımları	45
3.6.1. Parametrik Yaklaşım	46
3.6.2. Semiparametrik ; Nonparametrik Yaklaşımlar.....	47
3.6.3. Altın Kural Eksiği	48
4. ROC EĞRİSİ ALTI ALAN HESAPLAMA YÖNTEMLERİNİN KARŞILAŞTIRILMASI:.....	50
4.1. Normal Dağılıma Uygunluk Denetlemesi	50
4.1.1. Merkezi ve dağılım ölçütlerine göre karar verilmesi.	50
4.1.2. Grafikler; histogram, normal olasılık grafiği, Skewness ve kurtosis ile Normale uygunluk değerlendirmesi.	50
4.1.3. Hipotez Testleri ile Denetleme	50
5. GEREÇ VE YÖNTEM.....	51
5.1. Klinik Veriler:	51
5.2. Yöntem:.....	52
6. BULGULAR.....	53
6.1. ROC eğrilerini Karşılaştırılması	54
6.2. ROC EĞRİSİ ALTI ALAN HESAPLAMA YÖNTEMLERİNİN KARŞILAŞTIRILMASI.....	55

7. TARTIŞMA.....	58
KAYNAKÇA.....	60
ÖZGEÇMİŞ	65



TABLolar LİSTESİ

Tablo 1. Doğru Pozitiflik	3
Tablo 2. Yanlış Negatiflik	4
Tablo 3. Yanlış Pozitiflik.....	4
Tablo 4. Doğru Negatiflik.....	5
Tablo 5. Duyarlılık.....	5
Tablo 6. Özgüllük	6
Tablo 7. Pozitif Kestirim Değeri.....	7
Tablo 8. Negatif Kestirim Değeri	8
Tablo 9. Doğruluk.....	9
Tablo 10. Olabilirlik Oranı ve Klinik Tahminler Tablosu.....	11
Tablo 11. Sağlam ve Hasta gruplarının tanımlatıcı değerleri tablosu.....	51
Tablo 12. Sağlam Grup için “One-Sample Kolmogorov-Smirnov Testi” kullanılarak yapılan normallik denetlemesi sonuçları tablosu	51
Tablo 13. Hasta Grup için "One sample Kolgomorov-Smirnov Testi" kullanılarak yapılan normallik denetlemesi sonuçları tablosu	51
Tablo 14. Araştırma değişkenlerinin ROC analizi bulguları (Nonparametrik)	53
Tablo 15. ASRS-Toplam ve Barrat-Toplam ölçeklerinin değerlendirme tablosu.....	54
Tablo 16. ASRS-Toplam ve MKE-Toplam ölçeklerinin değerlendirme tablosu	54
Tablo 17. BARRAT-Toplam ve MKE-Toplam ölçeklerinin değerlendirme tablosu.....	54
Tablo 18. ASRS, AUC değerleri karşılaştırma tablosu	55
Tablo 19. ASRS, AUC değerleri arasındaki farklar tablosu.....	55
Tablo 20. BARRAT, AUC değerleri karşılaştırma tablosu.....	56
Tablo 21. BARRAT, AUC değerleri arasındaki farklar tablosu.....	56
Tablo 22. MKE, AUC değerleri karşılaştırma tablosu	56
Tablo 23. MKE, AUC değerleri arasındaki farklar tablosu	57

ŞEKİLLER LİSTESİ

Şekil 1. Murphy ve Winkler ayırım grafiği	15
Şekil 2. Yüksek oranlı AUC grafiği.....	16
Şekil 3. Normal dağılım eğrisi	18
Şekil 4. Lognormal dağılım	19
Şekil 5. Sabit dağılım.....	20
Şekil 6. Üçgensel dağılım	20
Şekil 7. PERT dağılımı	21
Şekil 8. Sabit dağılımlar.....	23
Şekil 9. ROC eğrisi değerleri.....	24
Şekil 10. Parabolik AUC	26
Şekil 11. Sınırsız destekli üçgen dağılım.....	27
Şekil 12. Asimetrik ROC eğrisi	28
Şekil 13. w_0 ve w_1 e bağlı AUC değeri	29
Şekil 14. Gauss eğrisi.....	30
Şekil 15. İçbükey ROC eğrisi.	31
Şekil 16. δc değişimi ve ROC eğrisi.	32
Şekil 17. Üçgen dağılım.....	33
Şekil 18. ROC eğrisi değişkenliği.....	34
Şekil 19. Doğrusal özelliği olmayan ROC eğrisi.....	35
Şekil 20. Beta dağılım.....	36
Şekil 21. Beta dağılımına bağlı ROC eğrisi.....	37
Şekil 22. Parametre değişikliğine bağlı AUC.	39
Şekil 23. ROC -AUC yorumlama	40
Şekil 24. Binormal dağılım.....	41
Şekil 25. Binormal ROC ve Nonparametrik ROC eğrisi Grafikleri	53

DENKLEMLER LİSTESİ

Denklem 1. Duyarlılık	6
Denklem 2. Özgüllük	6
Denklem 3. Pozitif Kestirim Değeri	7
Denklem 4. Negatif Kestirim Değeri	8
Denklem 5. Doğruluk	9
Denklem 6. Pozitif Kestirim Gücü	9
Denklem 7. Negatif Kestirim Değeri	9
Denklem 8. Olabilirlik Oranı	10
Denklem 9. İsabet Oranı ve yanlış alarm oranı	15
Denklem 10. Yanlış alarm oranı - İsabet oranı	23
Denklem 11.....	24
Denklem 12. Sabit Dağılımlar için AUC hesaplama	25
Denklem 13. Sınırsız destekli üçgen dağılım	27
Denklem 14. Sınırsız destekli üçgen dağılım ROC eğrisi	27
Denklem 15. Sınırsız üçgen dağılım için AUC hesaplama	28
Denklem 16.....	30
Denklem 17.....	30
Denklem 18.....	30
Denklem 19.....	30
Denklem 20. AUC hesaplama denklemi.....	32
Denklem 21.....	33
Denklem 22.....	34
Denklem 23.....	34
Denklem 24.....	35
Denklem 25.....	35
Denklem 26.....	36
Denklem 27.....	36
Denklem 28.....	36
Denklem 29.....	36
Denklem 30.....	37
Denklem 31.....	37
Denklem 32.....	38
Denklem 33.....	38
Denklem 34.....	38

SEMBOLLER / KISALTMALAR LİSTESİ

ROC	:Receiver Operating Characteristics
AUC	:Area Under Curve
EAA	:Eğri Altı Alan
DP	:Doğru Pozitiflik
YP	:Yanlış Pozitiflik
YN	:Yanlış Negatiflik
DN	:Doğru Negatiflik
TPR	:True Positive Rate
FPR	:False Positive Rate
KD	:Kestirim Değeri
PKD	:Pozitif Kestirim Değeri
NKD	:Negatif Kestirim Değeri
LR	:Likelihood Ratio
NLR	:Negative Likelihood Ratio
PLR	:Positive Likelihood Ratio
WBC	:White Blood Cell
PSA	:Prostat Specific Agent
AUROC	:Area Under Receiver Operating Characteristics

ÖZET

Vural, M (2020). ROC (Receiver Operationg Characteristic) Eğrisi Altı Alan Hesaplama Yöntemlerinin Karşılaştırılması.İstanbul Üniversitesi Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı.Doktora. İstanbul

Sağlık alanında, koruyucu tıp veya klinik çalışmalarda doğru ve erken tanı konması, başta ilgili hastalığın toplumsal seyri ve tedavi uygulamaları olmak üzere, olası tüm etkileşimleri ve sonuçlarının kontrolünü sağlayacaktır.

Tanı testi (diagnostik test) şüphelenilen hastalığın varlığı/yokluğunun belirlemesi amacıyla kullanılan değerlendirme kurguları veya özgün gereçlerin kullanımıyla elde edilen ölçüm veya hesaplama sonuçlarıdır.Yeni bir tanı test önerildiğinde, onun insan sağlığı üzerine olan etkisinin araştırılmasının amacı yalın uygulamalar ile risksiz, ekonomik olarak uygun ve yeterince doğruluğa sahip ve karalı sonuçlar verebiliyorsa kullanılacaktır.

Tanı testi performans belirlenmesi ve kıyaslamalarda Receiver Operating Characteristic (ROC) eğrisi çözümlemesi kullanılır. ROC eğrisi altındaki alan (Area Under Curve) tanı testinin doğruluk düzeyinin en iyi göstergesi kabul edilir.

Bu çalışmada,ROC eğrisi analizinin örneklem ve değişkenin sahip olduğu dağılım özelliklerine (parametrik/nonparametrik) bağlı olarak değişebildiği için, eldeki veriye en uygun çözümleme yaklaşımları ve özellikleri ele alınacaktır. Sürekli sonuçlu ölçütlerde altın standarda göre hasta ve sağlamların tanı testi sonuçlarının ve normal dağılımlı (parametrik) olduğu veya olmadığı (nonparametric) durumda kullanılan hesaplama yöntemlerinin sonuç bulgusu olan eğri altı alanların (AUC) birbirleri arasındaki farklılıklar incelenmiştir.Ayrıca, bir tanı testinin sürekli bir ölçütken bunun iki sonuçlu hale dönüştürmenin eğri altı alanların özelliklerine olan etkileri de örneklenmeye çalışılmıştır.

Örnek veri seti üzerinde yapılan değerlendirmeler sonucunda, normal dağılıma sahip tanı testi sonuçlarının parametrik hesaplama yöntemlerinin daha uygun (yansız) olduğu sonucuna varılmıştır.Varyansın geniş olduğu ve/veya çarpık dağılımlı verilerde, parametrik olmayan yönteminlerin (Mann Whitney) kullanılması daha uygun olacaktır.Eldeki sürekli değişkenin iki biçimli (dikotom) hale dönüştürülmesi ve kesim noktası (cut-off point) seçimi de (yüksek/düşük) neden olduğu bilgi kaybı nedeniyle, hesaplama sonuçlarını çok önemli ölçülerde etkileyecektir.

Bu çalışmada Binormal, Binomial exact_De Long ve Nonparametric (Mann-Whitney) çözümleme yöntemleri belli düzeylerde farklılıklar göstermesine rağmen,aralarında gözlenen farkların anlamlı olmadığı gözlenmiştir.

Anahtar Kelimeler: Tanı testi, ROC Eğrisi Analizi, Eğri Altı Alan (AUC).

ABSTRACT

Vural, M (2020). ROC (Receiver Operating Characteristic) Comparison of Methods Used in the Calculation of Area Under Curve. Istanbul University Institute of Health Sciences, Department of Biostatistics. Ph.D. Istanbul

Accurate and early diagnosis in preventive medicine or clinical studies in the field of health will ensure the control of all possible interactions and results, especially the social course and treatment practices of the relevant disease. A diagnostic test is the measurement or calculation results obtained by the use of evaluation setups or original tools used to determine the presence/absence of the suspected disease. When a new diagnostic test is proposed, it will be used in case the purpose of researching its effect on human health is risk-free, economically viable, sufficiently accurate and stable with lean applications. Receiver Operating Characteristic (ROC) curve analysis is used for diagnostic test performance determination and comparisons. The area under the ROC curve (Area Under Curve - AUC) is considered the best indicator of the accuracy level of the diagnostic test.

In this study, since the ROC curve analysis can vary depending on the sample and the distribution properties (parametric/nonparametric) of the variable, the most appropriate analysis approaches and their properties will be discussed according to the data in hand. The differences between the diagnostic test results of patients and healthy individuals were examined according to the gold standard in continuous results criteria, based on the areas under the curve (AUC), which are the result findings of the calculation methods used in the normal distribution (parametric) and otherwise (nonparametric). In addition, as a diagnostic test is a continuous criterion, the effects of converting it into binary results on the properties of the areas under the curve were also tried to be exemplified.

As a result of the evaluations made on the sample data set, it was concluded that the parametric calculation methods of the normal distribution diagnostic test results were more appropriate (unbiased). It would be more appropriate to use a nonparametric method (Mann Whitney) in data with large variance and/or skew distributed data. The transformation of the current continuous variable into a dichotomous form and the selection of the cut-off point (high / low) will also affect the calculation results significantly due to the loss of information caused by this method.

In this study, the differences observed between the Binormal, Binomial exact_De Long and Nonparametric (Mann-Whitney) analysis methods were not significant, although they do differ at certain levels.

Key Words: Diagnostic Test, ROC Curve Analysis, Area Under Curve (AUC).

1. GİRİŞ VE AMAÇ

Herhangi bir doğa gerçeği hakkında bilgi edinmenin de ötesinde, olay veya olgular arasında bağıntıların araştırılması ve karar verme amaçlı çalışmalara bilim ve teknoloji alanlarında sıklıkla rastlanır. Bilindiği gibi ROC (Receiver Operating Characteristic) yöntemi, ilk olarak 2. Dünya Savaşı sırasında İngilizlerin “Chain Home” adı verilen bir radar sisteminde düşman uçakları ve farklı radar sinyallerinin fark analizinde ROC eğrisi kullanılmış ve geliştirilmiş; örneğin bir kaz sürüşü ile düşman hava kuvvetleri uçaklarının yansıttığı radar sinyalleri arasındaki farkı daha doğru şekilde yorumlama noktasında ROC analizi kullanılmıştır. Bu ilk çalışmada da alıcının hassasiyeti arttırıldıkça, yanlış pozitiflerin sayısı da arttığı görülmüştür.

Tıp alanında ve endüstrinin diğer tüm alanlarında benzer şekilde ROC analizinin farklı uygulamalarıyla karşılaşırız. Özellikle tıptaki saha taramalarında düşük risk oranına sahip, hızlı ve ekonomik tanı yöntemlerine gereksinim duyulur. Bu tanıma en uygun analiz yöntemi ROC eğrisi ve analiz yöntemidir. Üzerinde çalışılan tanı testinin değerlendirilmesi genellikle 2 farklı şekilde planlanır: Birincisi, testin gerçek hastalara uygulanması ve gerçekten hasta olmayan kişilere uygulanması sonucunda elde edilen verilerin karşılaştırılmasıdır; ikincisi ise gerçeği yansıttığı öngörülen tanı testinin sonucuna göre yapılan analizdir.

Tanı testlerinin değerlendirilmesinde temel tanımların yapılabilmesi, testin analizinin gerçeğe en yakın sonucun elde edilmesi için önemli bir kriterdir. Bu amaçla, klinikte yapılan (ya da yapılması planlanan) testin sağlam bireyleri hangi oranda ayırt edebildiğinin gösterilebilmesi gerekir. “Özgüllük” olarak tanımlanan bu değerlendirme, tanı testinin temel parametrelerinden biridir. Diğer bir önemli parametre ise, hasta bireylerle hasta olmayan bireylerin yüzde olarak test tarafından ayrılabilme özelliğidir; bu da genel doğruluk yüzdesi olarak değerlendirilir.

Benzer şekilde uygulanan teste uyumlu olarak elde edilen ve hastalığa ait pozitif bulguların ne oranda gerçekten var olan hastalığa işaret ettiği, buna karşın aynı hastalığın negatif bulgularının da hastalığın ne oranda gerçekten var olmadığı durumlara işaret ettiği, testin yorum gücünü önemli ölçüde etkiler.

ROC eğrisinde ise tüm eşik değerlere göre elde edilen duyarlılık (sensitivite) y ekseninde, özgüllük (spesifite) ise (aslen 1- spesifite olarak) x ekseninde noktalar halinde belirtilerek en sonunda birleştirilir ve ortaya ROC eğrisi çıkar. Analiz sonunda AUC

(Area Under Curve) olarak belirtilen deęer ‘Eęri Altında Kalan Alan’ı (EAA) temsil eder ve bu deęer 1’e yaklaştıkça doęru tanı oranı yükselir. Böylece, %100’lük bir tanı gücünde EAA 1’e eşit olur. EAA ve %95 güven aralığı belirlenir. 0.50 deęeri (fark yoktur) güven aralığı dışında ise istatistiksel olarak anlamlı sonuç elde edilir.

Tezin konusu, temel olarak ROC eęrisi altı alanı hesaplamasında kullanılan yöntemlerin tanımlanarak karşılaştırılmalarıdır. ROC eęrisi altındaki alanın hesaplamasında binormal, binomial ve non-parametrik yöntemler kullanılmaktadır. Bu üç yöntemde tanı testinin gücüne olan etkileri bakımından deęerlendirilecektir.



2. GENEL BİLGİLER

2.1. Tanı Testleri (Diagnostic Tests)

Klinik olarak, sayıca fazla birey içeren topluluklardaki hastalıkları henüz belirti vermeden önce sağlıklı bireylerin ayrılmasında kullanılan testlere “Tarama Testi” denir. Tarama testlerinde sağlıklı bireylerin yüksek doğrulukta ayırd edilmesi amaçlanır. Risk taşımayan bu testler sahada uygulanması girişimsiz olması nedeniyle daha kolay olarak değerlendirilir.

Tanı testleri, daha çok sağlık kurumları içerisinde tasarlanarak ve belirli bir düzenek kullanılarak uygulanır. Tanı testlerinde hedef, hastaların ayırd edilmesidir. Daha karmaşık, maliyetli, testin tasarlanış amacı doğrultusunda risk taşıyan, uzun çalışma sürelerine ihtiyacı duyan testlerdir.

Önerilen yeni bir tanı testinin kullanıma uygun olabilmesi için hasta/sağlam birey ayırd edebilme gücü ve karar verme açısından en doğru ve en güvenilir klinik yöntemin sonuçlarına göre hem uyumlu hem de tutarlı olması esastır. Bu tür tanı testleri klinikte “altın standart” olarak tanımlanır.

2.1.1. Doğru Pozitiflik (True Positivity):

Hasta ve hasta olmayan olarak ayrılmış gruplarda analizi yapılan testin hasta bireylerde pozitif çıkması durumudur.

Test Sonucu	Hasta (+)	Sağlam (-)
Pozitif	Doğru Pozitif (DP)	Yanlış Pozitif (YP)
Negatif	Yanlış Negatif (YN)	Doğru Negatif (DN)

Tablo 1. Doğru Pozitiflik

Örnek: Klinik çalışması yapılan bir gebelik testi için gebe olan ve gebe olmayan gruplara test uygulandığında; gebe olan test grubundaki pozitif sonuçlar “Doğru pozitif” değer olarak tablodaki yerini alır.

2.1.2. Yanlış Negatiflik (False Negativity):

Hasta ve hasta olmayan olarak ayrılmış gruplarda analizi yapılan testin hasta bireylerde testin negatif sonuçlanması “Yanlış Negatif” olarak değerlendirilir.

Test Sonucu	Hasta (+)	Sağlam (-)
Pozitif	Doğru Pozitif (DP)	Yanlış Pozitif (YP)
Negatif	Yanlış Negatif (YN)	Doğru Negatif (DN)

Tablo 2. Yanlış Negatiflik

Örnek: Klinik çalışması yapılan bir gebelik testi için gebe olan ve gebe olmayan gruplara test uygulandığında; gerçekte gebe olan test grubundaki testin negatif sonuçlarına “Yanlış Negatif” değer denir.

2.1.3. Yanlış Pozitiflik (False Positivity):

Hasta ve hasta olmayanlar olarak ayrılmış gruplarda analizi yapılan testin hasta bireylerde testin pozitif sonuçlanması “Yanlış Pozitiflik” olarak değerlendirilir

Test Sonucu	Hasta (+)	Sağlam (-)
Pozitif	Doğru Pozitif (DP)	Yanlış Pozitif (YP)
Negatif	Yanlış Negatif (YN)	Doğru Negatif (DN)

Tablo 3. Yanlış Pozitiflik

Örnek: Klinik çalışması yapılan bir gebelik testi için gebe olan ve gebe olmayan gruplara test uygulandığında; gerçekte gebe olmayan test grubundaki testin pozitif sonuçlarına “Yanlış Pozitif” değer denir.

2.1.4. Doğru Negatiflik (True Negativity):

Hasta ve hasta olmayan ayrılmış gruplarda analizi yapılan testin sağlam bireylerde testin negatif sonuçlanması “Doğru Negatif” değer olarak değerlendirilir

Test Sonucu	Hasta (+)	Sağlam (-)
Pozitif	Doğru Pozitif (DP)	Yanlış Pozitif (YP)
Negatif	Yanlış Negatif (YN)	Doğru Negatif (DN)

Tablo 4. Doğru Negatiflik

Örnek: Klinik çalışması yapılan bir gebelik testi için gebe olan ve gebe olmayan gruplara test uygulandığında; gerçekte gebe olmayan grupta testin negatif sonuçlarına “Doğru Negatif” değer denir.

2.1.5. Duyarlılık (Sensitivity)

Duyarlılık, herhangi bir tanı testinin , hasta olduğu varsayılan gruptaki tüm hastaları doğru olarak tanımlayabilme kabiliyetini gösterir. Test sonucu gerçek pozitif hasta sayısının, gerçek pozitif hasta sayısı ile yanlış pozitif hasta sayısının toplamına oranı testin “duyarlılığını” ifade eder.

Test Sonucu	Hasta (+)	Sağlam (-)	Toplam
Pozitif	DP (A)	YP (B)	A + B
Negatif	YN (C)	DN (D)	C + D
Toplam	A + C	B + D	A + B + C + D

Tablo 5. Duyarlılık

$$Duyarlılık = \frac{DP (A)}{DP (A) + YN (C)}$$

Denklem 1. Duyarlılık

2.1.6. Özgüllük (Specificity)

“Gerçek Negatif Oran” olarak da tanımlanır. Uygulanan testin “Gerçek Negatiflik” örneklem sayısının , bu sayı ile gerçek olmayan ve pozitif birey sayısının toplamına oranı “Özgüllük” olarak tanımlanır. Özgüllük, bir testin hastalığı olmayan hastaları ne kadar iyi tanımlayabildiğini gösterir. Duyarlılık ve Özgüllük arasında gözlenen eşzamanlı ve zıt yöndeki ilişki, klinik araştırmacıları eşik değer seçimi konusunda kritik kararlar almaya zorlamaktadır.

Test Sonucu	Hasta (+)	Sağlam (-)	Toplam
Pozitif	DP (A)	YP (B)	A + B
Negatif	YN (C)	DN (D)	C + D
Toplam	A + C	B + D	A + B + C + D

Tablo 6. Özgüllük

$$\text{Özgüllük} = \frac{DN (D)}{YP (B) + DN (D)}$$

Denklem 2. Özgüllük

2.1.7. Kestirim Değeri (Predictive Value)

Kestirim Değeri (KD-PV) testin olasılık durumunu belirleyen değerdir. Özgüllük ve Duyarlılık değerleri bir ölçüm aracının tanı amaçlı kullanıma ne ölçüde uygun olduğunu ifade etmektedir. Ancak, yalnızca testten elde edilen pozitif sonuca bakarak hastalığın gerçekten var olduğunu söylemek için Özgüllük ve Duyarlılık değerlerini

bilmek yeterli değildir. Hastalığın varlığıyla ilgili bilgiye kestirim değerleri hesaplanarak ulaşılmaktadır. Popülasyonda testi yapılan ya da yapılması planlanan hastalığın toplum içindeki görülme sıklığından (prevalansından) etkilenir. Yüksek prevalanslı bir ortamda test yapıldığında , pozitif çıkan hasta grubun gerçekten hasta çıkma ihtimali, düşük prevalanslı bir popülasyonda hastalığa yakalanma olasılığından daha yüksek olacaktır. Özetle Pozitif Kestirim Değeri ve Negatif Kestirim Değerlerinin kullanılabilmesi için çalışmada elde edilen (hasta sayısı / toplam denek sayısı) yüzdesinin, hastalığın toplum içindeki görülme sıklığını (prevalansını) temsil ediyor olması gerekir.

2.1.8. Pozitif Kestirim Değeri (Positive Predictive Value)

Pozitif Kestirim Değeri (PKD-PPV), hasta olupta testten pozitif sonuç olan kişilerin oranı olarak tanımlanmaktadır ve aşağıdaki formülle hesaplanmaktadır.

Test Sonucu	Hasta (+)	Sağlam (-)	Toplam
Pozitif	DP (A)	YP (B)	A + B
Negatif	YN (C)	DN (D)	C + D
Toplam	A + C	B + D	A + B + C + D

Tablo 7. Pozitif Kestirim Değeri

$$\text{Pozitif Kestirim Değeri (PKD - PPV)} = \frac{DP (A)}{DP (A) + YP (B)}$$

Denklem 3. Pozitif Kestirim Değeri

2.1.9. Negatif Kestirim Değeri (Negative Predictive Value)

Negatif Kestirim Değeri (NKD) ise normal olup da testten negatif sonuç olan kişilerin oranı olarak tanımlanmaktadır ve aşağıdaki formülle hesaplanmaktadır.

Test Sonucu	Hasta (+)	Sağlam (-)	Toplam
Pozitif	DP (A)	YP (B)	A + B
Negatif	YN (C)	DN (D)	C + D
Toplam	A + C	B + D	A + B + C + D

Tablo 8. Negatif Kestirim Değeri

$$\text{Negatif Kestirim Değeri (NKD - NPV)} = \frac{DN (D)}{YN (C) + DN (D)}$$

Denklem 4. Negatif Kestirim Değeri

Aynı evren içinde görülme sıklığı düşük ve görülme sıklığı yüksek olan iki farklı bozukluk için ölçüm araçlarına ait hesaplama yapıldığında , her iki örneklem için Özgüllük ve Duyarlılık değerleri aynı bulunabilir. Ancak bozukluğun görülme sıklığına bağlı olarak kestirim değerlerinde anlamlı farklılıklar hesaplanacaktır. Bir bozukluğun görülme sıklığı azaldıkça PKD'nin azaldığı ve NKD'in arttığı; bozukluğun görülme sıklığı arttıkça PKD'nin arttığı ve NKD'in azaldığı görülmektedir. Muhtelif uygulamalarda sıkça rastlanan sonuçlardan biri; yorumlayıcının, Pozitif Kestirim Değeri yüksek bir ölçüm aracı kullandığını varsaydığı durumlarda, uygulama yaptığı örneklemde bozukluğun görülme sıklığının değişmesi nedeniyle beklediği kesinliğe ulaşamaması durumudur. Bu nedenle, farklı dağılım özelliğine sahip örneklem için görülme sıklığının değişebileceği tahmin ediliyorsa PKD ve NKD'in yeniden hesaplanması daha doğru olacaktır.

2.1.10. Doğruluk (Accuracy)

Bir kümenin ölçümünde doğruluk, elde edilen ölçümlerin belirli bir değere yakınlığını ifade eder. Hassasiyet ise ölçümlerin birbirine yakınlığıdır. Doğruluk, iki şekilde tanımlanabilir: Daha yaygın olarak, sistematik hataların bir tanımı, istatistiksel yanlılığın bir ölçüsüdür. Düşük doğruluk değeri, bir sonuç ile "gerçek" bir değer arasında bir fark yaratır. Diğer taraftan doğruluk, her iki gözlemsel hatanın (rastgele ve sistematik) kombinasyonu olarak tanımlanır. Bu nedenle yüksek doğruluk hem yüksek hassasiyet hem de yüksek doğruluk gerektirir.

Test Sonucu	Hasta (+)	Sağlam (-)	Toplam
Pozitif	DP (A)	YP (B)	A + B
Negatif	YN (C)	DN (D)	C + D
Toplam	A + C	B + D	A + B + C + D

Tablo 9. Doğruluk

$$\text{Doğruluk (Accuracy)} = \frac{DP (A) + DN (D)}{DP (A) + YP (B) + YN (C) + DN (D)}$$

Denklem 5. Doğruluk

2.1.11. Bayes Teoremi (Bayes Theorem)

Yapılan çalışmada Duyarlılık ve Özgüllük dışında bir değere ihtiyaç duyulmadığı durumda, çalışmanın prevalans değerini temsil edip etmemesi gerekliliğine bakılmaz. Ancak çalışma sonucunda prevalans değeri evreni temsil etmiyorsa, buna rağmen Pozitive Kestirim Değeri ve Negative Kestirim değerini kullanmak istiyorsak bu kez prevalans değerini hesap içerisinde uygulamamız gerekir.

$$\text{Pozitive Kestirim Gücü (PKG)} = \frac{\text{Duyarlılık} \times \text{Prevalans}}{(\text{Duy.} \times \text{prev.}) + ((1 - \text{özg}) \times (1 - \text{prev}))}$$

Denklem 6. Pozitif Kestirim Gücü

$$\text{Negative Kestirim Gücü (NKG)} = \frac{\text{Özgüllük} \times (1 - \text{Prev.})}{\text{Özgüllük} \times (1 - \text{Prev.}) + ((1 - \text{duy}) \times \text{prev})}$$

Denklem 7. Negatif Kestirim Değeri

2.1.12. Olabilirlik Oranı (likelihood Ratio)

Olabilirlik oranları, istatistik hesaplarında sıkça başvuru alan yöntemlerden değildir. Buna karşın teşhis doğruluğuna katkı sağlayan en önemli ölçüt olarak kabul edilir. Çünkü hastalığın olasılığı yada olasılıkları arasındaki ilişki için bu hesaplamalara ihtiyaç duyulur. Olabilirlik hesaplamasında basitçe aşağıdaki yöntemler kullanılır ve klinisyenin pratik olarak tanıda ne oranda olabilirlikle karşı karşıya olduğunu gösterir.

Herhangi bir klinik bulgunun Olabilirlik Oranı (LR), hastalığı olan hastalarda bulguların bulunma olasılığının, hastalığı olmayan hastalarda aynı bulgunun olasılığına bölünmesi oranına eşittir.

$$\text{Olabilirlik Oranı (LR)} = \frac{\text{Hastalığı olan bireylerde bulguların bulunma olasılığı}}{\text{Hastalığı olmayan bireylerde bulguların bulunma olasılığı}}$$

Denklem 8. Olabilirlik Oranı

Basit bir örnekle; ultrasonografiye tabi tutulan abdominal distansiyonu olan hastalar arasında, fiziksel olarak “sağ/sol kolik bölge hassasiyeti” , onaylanmış asitli hastaların% 80'inde ve asitsiz hastaların % 40'ında (yani, obezite veya gazdan) görülür. Bu nedenle, asit tespitinde " sağ/sol kolik bölge hassasiyeti " için LR 2.0'dır (yani% 80 bölünme% 40). Benzer şekilde, asitli hastaların% 10'unda ancak diğer nedenlerden uzak olan % 30'unda "lateral timpani” bulgusu mevcutsa, asitlerin saptanmasında "lateral timpani" için LR 0.3'tür (yani% 10'un% 30'a bölünmesi).

Olabilirlik Oranı (LR) ; 0 ile sonsuz arasında olabilir. Bu oranın 1'den büyük olduğu durumlarda bulgular, araştırılan tanıyı savunur; sayı ne kadar büyük olursa, bulgu bu hastalık için o kadar ikna edici olduğunu gösterir. Olasılık Oranları 0 ile 1 arasında olan bulgular, aranan bulgu teşhise karşı olarak değerlendirilir. Olabilirlik Oranı (LR) 0'a ne kadar yakınsa, bulgu ile aranan hastalık ilişkisi o kadar düşük olarak değerlendirilir. Olasılık Oranı (LR) 1 olan bulgular tanı değerine sahip değildir. (Tablo.10)

Olabilirlik Oranı	Olabilirlikteki yaklaşık değişim. (%)
“0” ile “1” arasında hastalığın yada bulgunun olabilirliği azalır.	
0.1	-45
0.2	-30
0.3	-25
0.4	-20
0.5	-15
1	0
“1” den büyük değerler hastalığın yada bulgunun var olabilirliğini artırır.	
2	+15
3	+20
4	+25
5	+30
6	+35
7	

8	+40
9	
10	+45

Tablo 10. Olabilirlik Oranı ve Klinik Tahminler Tablosu

2.1.13. Negatif Olabilirlik Oranı (Negative Likelihood Rate)

Negatif olabilirlik oranı (NLR), negatif testi olan hastalarda tanı alma olasılığındaki değişikliği verir. Değişiklik, genellikle 1'den küçük bir oran şeklindedir. Örneğin, 0.1'lik bir NLR, negatif test sonucuna sahip bir hastada bir rahatsızlığa sahip olma ihtimalinde 10 katlık bir azalma olduğunu gösterecektir.

2.1.14. Pozitif Olabilirlik Oranı (Positive Likelihood Rate)

Olabilirlik oranları sıfırdan sonsuza kadar değişir. Oran “0” ile “1” arasında Negative Olabilirlik Oranı (NLR) tanımlanırken ; Oran “1” e eşitlendiğinde Olabilirlikteki değişim “0” (sıfır) olarak gerçekleşir. Değer, “1” den büyük ve ne kadar yüksek olursa, hastanın yada bulgunun hastalıkta bulunma durumu daha olasıdır. Örnek olarak, pozitif bir test sonucunun LR'sinin 9.2 olduğunu varsayalım. Bu sonuç, durumu olan bir hastada, durumu olmayan bir hastaya göre 9.2 kat daha fazladır.

2.1.15. Fagan Nomogramlar

Tanı testlerinin hasta ve sağlam olanları ayırmadaki başarısını tanımlamak için duyarlılık, özgüllük, pozitif ve negatif kestirim değerleri gibi bazı ölçütler kullanılmaktadır. Duyarlılık ve özgüllük tanı testinin özellikleri olup bir kişinin hasta olma olasılığı hakkında fikir vermez. Bir kişide test sonrası hastalık olasılığını tahmin etmede pozitif ve negatif kestirim değerleri kullanılabilir, ancak bu sonuçları kişiye özel tahmin etmek klinikte uygulanabilir değildir. Bu amaçla olabilirlik oranı ve test öncesi olasılık kullanılır. Pozitif ve negatif olabilirlik oranlarının kullanımı klinik olarak daha pratiktir. Hekime test sonrası hastalık olasılığını her bir hasta için tahmin etme olanağı sağlar.

Bir test sonucu için olabilirlik oranı, gözlemlenen sonucun hasta olanlarda gerçekleşme olasılığının, hasta olmayanlar arasındaki gerçekleşme olasılığına oranıdır. Bir kişide test sonrası olasılığı saptamak için Olabilirlik oranı ve test öncesi olasılık

birlikte kullanılır. Nomogram; test sonrası olasılığı tahmin etmede olabilirlik oranı ve test öncesi olasılığın birlikte kullanımına olanak sağlayan bir araçtır.



3. ROC – RECEIVER OPERATING CHARACTERISTICS

Tanı testlerinin sonuçları eğer var/yok, yüksek/normal veya düşük/normal şeklinde iki sonuçlu ise testlerin başarısı; sensitivite, spesifisite, pozitif prediktif değer, negatif prediktif değer ölçütleriyle değerlendirilir. Ancak test sonuçları sürekli sayısal veriler (örn. hemoglobin değeri, WBC sayısı, plazma glukoz konsantrasyonu, PSA değeri) ya da sıralı sayısal veriler (örn. BI-RADS) şeklinde ise, bu testlerin tanı koymadaki başarılarını değerlendirmek biraz daha karmaşık olabilmektedir.

Tanı testlerinin performansının ölçülmesinde yaygın olarak uygulanan ROC eğrisini modellemek için birçok Direkt ve İndirekt yaklaşım önerilmiştir ve izlenebilirliği nedeniyle tipik olarak her iki popülasyonu modellemek için Gauss dağılımı sıklıkla kullanılır. Monte Carlo simülasyonu ise, eğride kapalı bir form olmaması sorununu önlemek için kullanılır.

3.1. ROC Eğrisi Altı Alanı (AUROC – Area Under Receiver Operating Characteristic Curve)

Tanı sırasında hastalığın sınıflandırılması gereği varsa, bir AUC - ROC Eğrisi ile güvenli sonuçlar elde etmek mümkündür. Birden fazla gruplu sınıflandırma işleminin performansını kontrol etmek veya görselleştirmek gerektiğinde, AUC (Area Under Curve) ROC (Receiver Operating Characteristics) eğrisi kullanılır. AUROC; gruplama modeli ne olursa olsun bu modelin performansını kontrol etmek için en önemli değerlendirme parametrelerinden biridir. İlk olarak 2. Dünya Savaşı sırasında geliştirilmeye başlanmış olan alan hesaplama yöntemleri, tanı testlerinin performans değerlendirmelerinde yaygın olarak kullanılmaya başlanmıştır. Günümüzde tanı testlerinin performansının yanında, doğru ve optimum eşik değerinin saptanmasında da en güvenilir yöntem AUROC olmuştur.

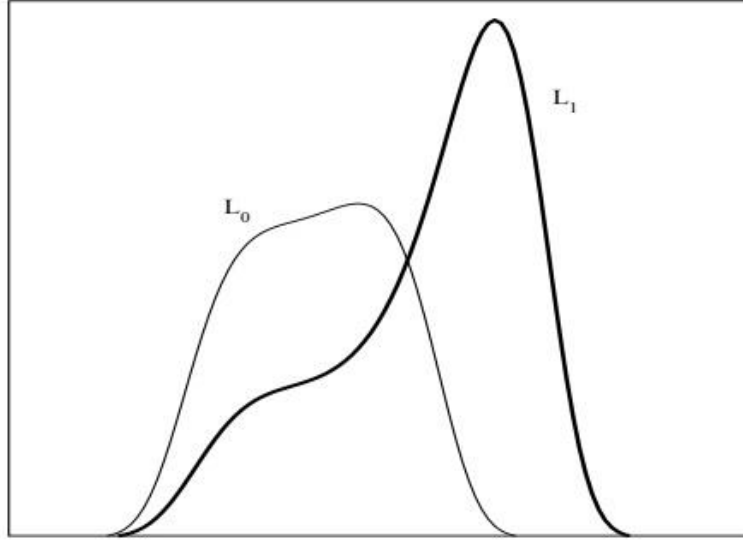
ROC Eğrisinin modellenmesi konusunda birçok araştırma, Direkt ve İndirekt yöntem olmak üzere iki farklı yöntem tarif eder. Çok tercih edilmeyen Direkt yöntemde veriler, herhangi bir dağılım hipotezine bağlı değildir. Bu yaklaşım, herhangi bir varsayıma bağlı olmaksızın popülasyonu “hastalar” ve “sağlıklılar” olarak iki gruba ayırmakla başlar. Ancak bu veriye bağlı ROC eğrisi belirli teorik özellikleri ihlal eder; örneğin monoton olmak zorunda değildir. Araştırmacılar, non-parametrik tahminlerinde bu sorunu çözmek amacıyla non-parametrik olgularda, esnek bir dağılımla, toplumun

temel özelliklerini koruyarak dağılım sağlamayı hedeflemişlerdir. Sonuçta optimal bant genişliğine indirgenmiş olan veri yığımında problem de küçülmüş olacaktır.

İndirekt yaklaşımda ise, her nüfusun belirli bir dağılımı takip ettiği varsayılır ve dolaylı olarak ROC eğrisi için fonksiyonel bir form türetir. Bir eğri oluşturmak amacıyla parametrik ve semi-parametrik yöntemler önerilir. Gauss varsayımı için Goddard ve Hinberg (1990) prostat kanseri gibi bazı senaryolarda her zaman yeterli bir seçenek olmadığını belirtmiştir. Araştırmacılar yöntemin rastgele ve hassasiyetle uygulanmaması durumunda dağılımsal farklılıklara yol açacağından ve böylece sonuçlar güvenilir olmaktan uzaklaşacağı için tavsiye etmemektedirler. Alternatif olarak, dağılım varsaymak yerine ROC eğrisi tanımlanabilir. Örneğin, her iki popülasyon da aynı varyansla bir lojistik dağılımı takip ettiği varsayılmıştır. İngiliz araştırmacılar iki parametrelili üstel bir model önermiştir. Her iki parametrik yöntem de çok benzemektedir, çünkü test puanlarının dağılımı ROC eğrisine tamamen uyumludur. Parametrik bir yöntemin ana avantajları basitlik, ROC eğrisinin beklenen seviyede olması ve az sayıda parametre ile çalışma yeteneği olarak 3 maddede sıralanabilir.

Non-parametrik ve parametrik bileşenlerin varlığından doğan esneklik, semi-parametrik yöntemi ön plana taşımaktadır ve araştırmalarda daha cazip görünmektedir. Binormal model buna iyi bir örnek oluşturabilir. Çünkü Binormal yöntemde, Gauss dağılımıyla her iki grupta, artış yönünde monoton bir değişim mevcuttur. Eğim ve kesim değerinin daha net görülmesi, parametre tahminini de kolaylaştırır.

Yukarıda belirtilen var/yok, yüksek/normal veya düşük/normal şeklinde iki sonuçlu gözlemler için üretilen tahminlerin kalitesinin nasıl değerlendirilebileceği sorusu ROC eğrisi altında kalan alanı ilgilendirmektedir. Tahmin edilen değer $-\infty$ ile $+\infty$ arasında bir değer olabileceği gibi, 0 ve 1 arasında değişen bir olasılık da olabilir. Aşağıdaki figürde Murphy ve Winkler tahminlerin ve gözlemlerin ortak olasılık dağılımına bakarak bu olasılıkları analiz etmiştir ve bir diyagram çıkarmıştır. Aşağıdaki figürde, L_0 ve L_1 olarak adlandırılan genel özellikte iki sınıf için olasılıklar görülmektedir. Dolayısıyla $L_i(x)$, “x” tahmininin i sayıda sınıf için olasılık dağılımıdır. Bu dağılım grafiğini Murphy ve Winkler “ayırım grafiği” olarak isimlendirmişlerdir.



Şekil 1. Murphy ve Winkler ayırım grafiği

Bu figürde, tahminlerin kalitesinin buna benzer birkaç grafik yardımıyla tam genellikte değerlendirilebileceği gösterilmeye çalışılmıştır. Diğer özellikleri gösteren grafikler de bu kalitenin farklı yönlerini ölçer – çözünürlük, güvenilirlik gibi. ROC eğrisi en basit haliyle, bir karar değeri sürekli olarak tahmin miktarının tüm aralığı boyunca değiştiği için, kararın isabet oranının (veya algılama olasılığının) parametrik bir grafiği olarak tanımlanabilir. Çapraz çizgi rastgele tahminlere karşılık gelir ve içbükeylik performansının bir ölçüsü olarak alınır. ROC eğrisi altındaki alan (AUC) ise genellikle skaler bir ölçü olarak alınır.

0.5 değerinde bir AUC rastgele tahminlere, 1 değerinde bir AUC mükemmel tahminlere karşılık gelir. İsabet oranı ve yanlış alarm oranıysa olasılıklardan hesaplanabilir.

İsabet oranı ve yanlış alarm oranı, olasılıklardan hesaplanabilir:

$$H = \int_t^{\infty} L_1(x) dx \quad , \quad F = \int_t^{\infty} L_0(x) dx$$

Denklem 9. İsabet Oranı ve yanlış alarm oranı

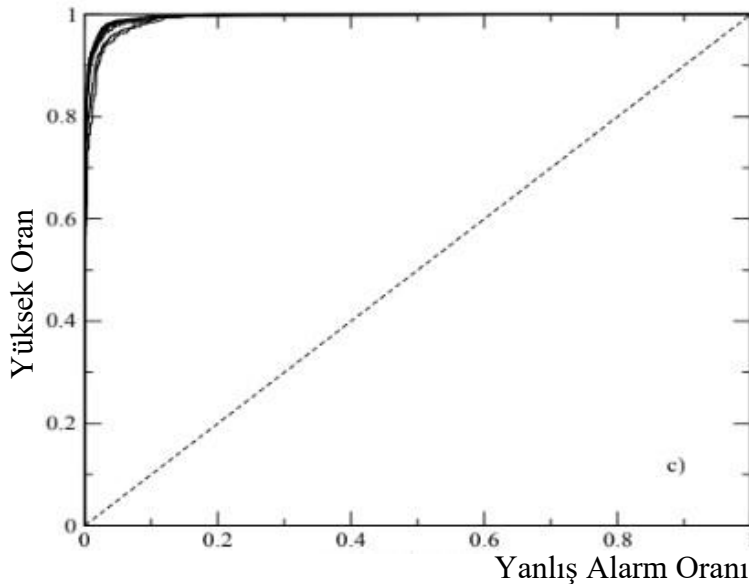
t: karar eşiği

Olasılıklı tahminlerde bu limit 1'dir.

Belirli bir veri kümesi için, normalize edilmiş bir x histogramı, olasılığı görselleştirmenin en iyi yoludur. ROC sistemiyse Murphy-Winkler sisteminden biraz farklılık gösterir. Örneğin, olasılıklı tahminler için Murphy-Winkler sistemi, tahminlerin kategorik olarak azaltılmasını gerektirmez ve aslında tersini bile gerektirir. ROC analizi ise aksine beklenmedik durum tablosuna dayanmaktadır ve bu nedenle, sürekli tahminlerin ikili tahminlere indirgenmesi amacıyla bir karar eşiğinin verilmesini gerektirir. Tabii ki, bir eşiğin sunulması, ROC analizinin hiçbir şekilde Murphy-Winkler çerçevesinden daha az güvenilir olduğu anlamına gelmez; yalnızca performansın farklı yönlerine vurgu yapılarak, performansı değerlendirmenin başka bir yöntemidir. Murphy-Winkler çerçevesi farklı tahmin kümelerini karşılaştırmak için daha uygundur (örneğin, iki tahmincinin tahmin kümelerini); ROC analizinde ise bir karar eşiğinin açıkça varlığı, tahminlere yanıt olarak durum hakkında bir kararın verilmesi veya eylemin yapılmasını gerektirir.

3.2. ROC Eğrisi Altı Alanı Hesaplama Yöntemleri

ROC eğrisini kullanan birçok araştırmacı, farklı durumlarda çok çeşitli tahmin yöntemleri üzerinde çalışırken, tahminlerin çoğunun yüksek derecede içbükey ROC eğrilerine veya eşdeğer derecede yüksek AUC değerlerine yol açtığı gözlemlemiştir. 0.9995 gibi AUC değerleri nadir değildir. Aşağıdaki şekilde, aşırı içbükeyliğe sahip 8 set ROC eğrisi görülmektedir:



Şekil 2. Yüksek oranlı AUC grafiği

Eğrilerin altında yatan tahminler farklı tahmin özelliklerine sahiptir (Murphy-Winkler sistemiyle hesaplanan olasılıklı tahminlerin çeşitli nitelikleri açısından), ancak hepsi oldukça içbükey ROC eğrilerine yol açar. Bu eğriler için AUC değerleri 0.990 ila 0.996 arasında değişmektedir. Bu AUC değerleri neden 1'e yakın diye düşündüğümüzde, tahminler olağanüstü kalitede olduğu için mi, yoksa AUC'nin kendisinin bir ürünü mü olduğunu sorgulamalıyız. Birincisi doğruysa, tüm AUC değerlerinin bir histogramı sağa doğru doruğa ulaşır (veya solda yoğun bir kuyruk gösterir). Bunu test etmek zordur, çünkü gerekli verilerin derlenmesi zor olacaktır. Öte yandan, hata ölçümün kendisindeyse, o zaman açıklamamız elimizin altında olacağı için bu hipotezi test etmek gereksiz olacaktır. Ayrıca nasıl bir durum bire yakın AUC değerlerine yol açabilir? Yukarıda belirtildiği gibi, her iki yaklaşımın da farklı vurguları olmasına rağmen, birbirleriyle ilişkilidir. Sonuçta, bir ROC eğrisinin türetildiği miktarlar - isabet oranı ve yanlış alarm oranı - koşullu dağılımların altında, bazı karar eşiğinin üzerinde olan alanlardır. Ayrıca, ROC eğrilerinin hesaplanması bu dağılımlar hakkında bilgi gerektirmemesine rağmen, ROC eğrilerinin istatistiksel önemini bir değerlendirmesini gerektirir. Örneğin, ROC veya AUC için standart hataları hesaplamak için (parametrik bir yaklaşımla), bu temel dağılımlarla ilgili bazı varsayımlar yapılır. Bu gibi durumda, yukarıdaki soruları cevaplamak için ROC eğrisi ile temel dağılımlar arasındaki bağlantıyı kullanmak doğaldır. Vereceğimiz cevaplarsa, ROC eğrilerini daha temel bir düzeyde yorumlamanın bir yolunu sunar. Bu şekilde, ROC eğrisinin şekli yorumlanabilir veya "açıklanabilir". Temel dağılımlar hakkında bilgi sahibi olmak ise daha iyi tahminlerin geliştirilmesine rehberlik edebilir. Monte Carlo Simülasyonu'nun açıklamasının ardından bu bölümde ele alacağımız beş dağılım, 1) Sabit, 2) Kısıtsız destekli üçgen, 3) Gauss, 4) Kısıtlı destekli üçgen ve 5) Beta dağılımlarıdır. İlk üçü tahmin miktarının gerçek çizgi üzerinde $-\infty$ ila $+\infty$ arasında değiştiği durumlar için uygundur, son ikisiyse olasılıklı tahminler için geçerlidir.

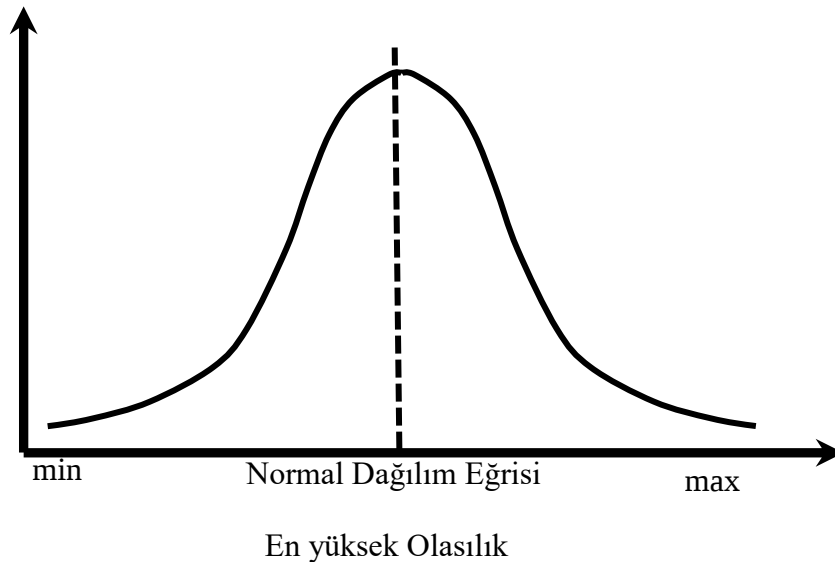
3.2.1. Monte Carlo Simülasyon yöntemi

Monte Carlo simülasyonu, insanların kantitatif analiz ve karar vermede riski hesaba katmalarını sağlayan bilgisayarlı bir matematik tekniğidir. Teknik, finans, proje yönetimi, enerji, imalat, mühendislik, araştırma ve geliştirme, sigorta, petrol, gaz, ulaşım ve çevre gibi çok farklı alanlarda profesyoneller tarafından kullanılır. Monte Carlo simülasyonu, karar vericiye bir dizi olası sonuç ve seçilmiş herhangi bir eylem için ortaya

çıkacak olasılıkları sunar. Süreç devam ederken alınacak kararlar için tüm olası sonuçların yanı sıra, aşırı olasılıkları da gösterir. Teknik ilk olarak atom bombası üzerinde çalışan bilim adamları tarafından kullanılmış; ismi ise Monako'da kumarhaneleriyle ünlü tatil beldesi Monte Carlo'dan gelmiştir. II. Dünya Savaşı'na girişinden bu yana, Monte Carlo simülasyonu çeşitli fiziksel ve kavramsal sistemleri modellemek için kullanılmıştır.

Monte Carlo simülasyonu, doğasında belirsizliği olan herhangi bir faktör için bir dizi değeri - olasılık dağılımını - değiştirerek olası sonuç modellerini oluşturur ve bu şekilde risk analizi yapar. Daha sonra, olasılık işlevlerinden farklı rastgele değerler kümesi kullanarak sonuçları defalarca hesaplar. Belirsizliklerin sayısına ve bunlar için belirtilen aralıklara bağlı olarak, bir Monte Carlo simülasyonu tamamlanmadan önce binlerce veya on binlerce yeniden hesaplama içerebilir. Monte Carlo simülasyonu olası sonuç değerlerinin dağılımlarını üretir.

Olasılık dağılımları kullanılarak, değişkenlerin farklı sonuçlarının meydana gelme olasılıkları farklı olduğu görülebilir. Olasılık dağılımları, bir risk analizinin değişkenlerindeki belirsizliği tanımlamanın çok daha gerçekçi bir yoludur. Ortak olasılık dağılımları aşağıdakileri içerir:

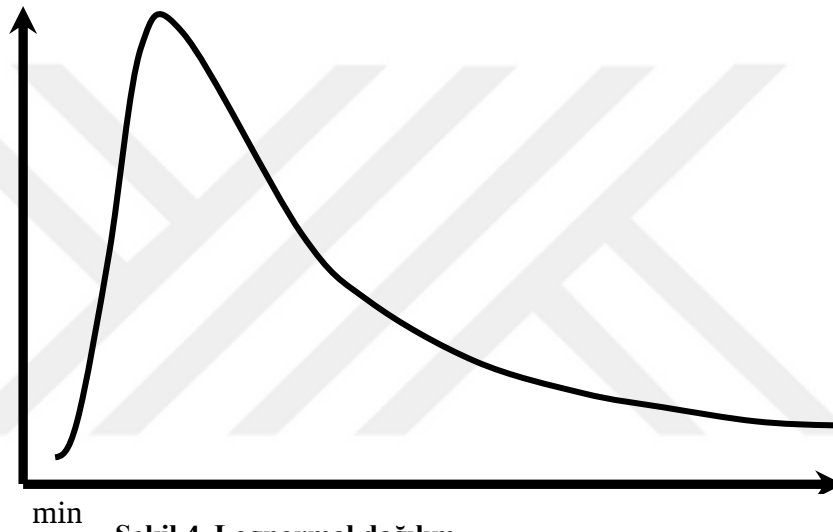


Normal veya “çan eğrisi”. Kullanıcı sadece ortalama veya beklenen değeri ve ortalama ile ilgili varyasyonu tanımlamak için standart sapmayı tanımlar. Ortalamanın

ortadaki değerlerin meydana gelme olasılığı yüksektir. Simetriktir ve insanların boyları gibi birçok doğal fenomeni tanımlayabilir. Normal dağılımlarla açıklanan değişkenlere örnek olarak enflasyon oranları ve enerji fiyatları verilebilir.

Lognormal

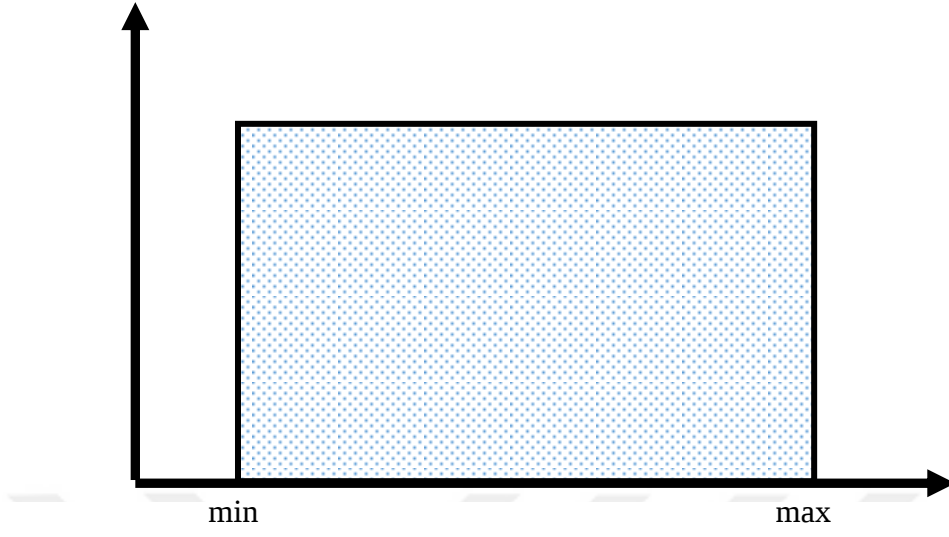
Değerler, normal bir dağılım gibi simetrik değil, pozitif eğimli bir eğridir. Sıfırın altına düşmeyen ancak sınırsız pozitif potansiyele sahip değerleri temsil etmek için kullanılır. Lognormal dağılımlarla tanımlanan değişkenlere örnek olarak gayrimenkul mülk değerleri, hisse senedi fiyatları ve petrol rezervleri gösterilebilir.



Şekil 4. Lognormal dağılım

Sabit

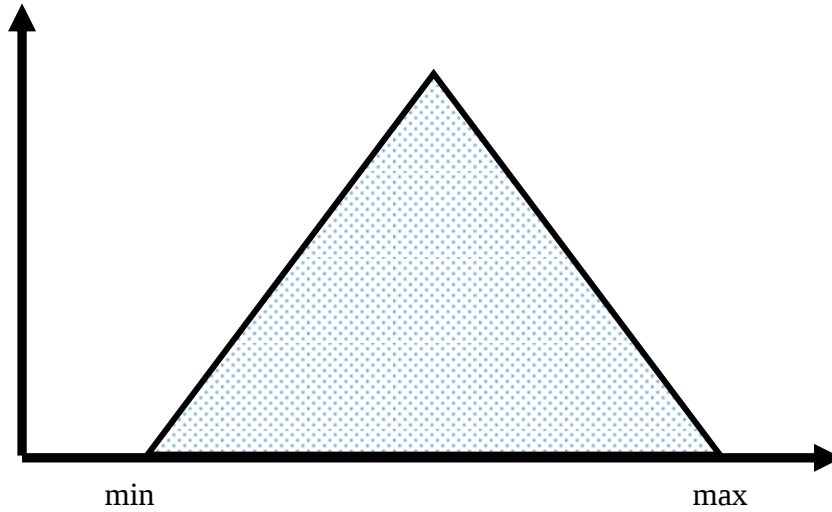
Bu dağılımda tüm değerlerin eşit gerçekleşme şansı vardır ve kullanıcı yalnızca minimum ve maksimum değerleri tanımlar. Aynı şekilde dağıtılabilecek değişkenlere örnek olarak üretim maliyetleri veya yeni bir ürün için gelecekteki satış gelirleri verilebilir.



Şekil 5. Sabit dağılım

Üçgensel

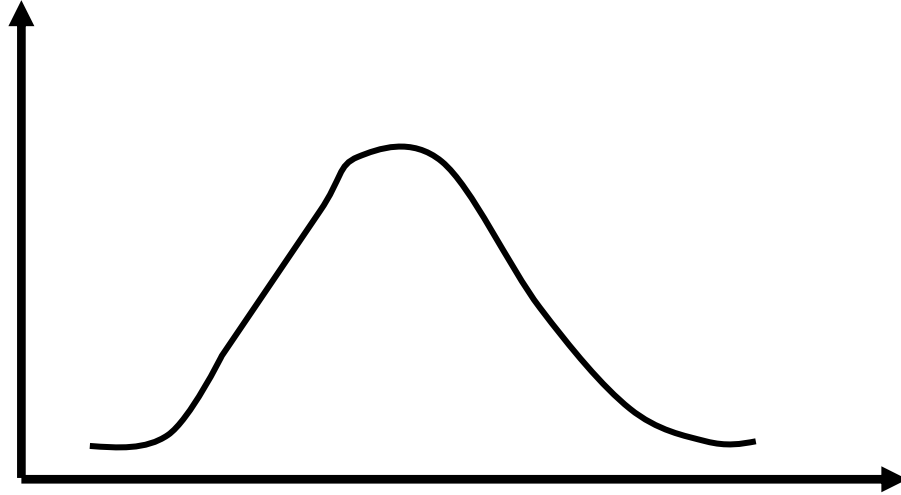
Kullanıcı minimum, maksimum ve yüksek olasılıklı değerleri tanımlar. Büyük olasılıklı olarak işaretlenmiş değerlerin ortaya çıkması daha olasıdır. Üçgen dağılımla tanımlanabilecek değişkenler, birim zaman başına geçmiş satış geçmiş ve stok seviyeleri olabilir.



Şekil 6. Üçgensel dağılım

PERT

Kullanıcı, üçgen dağılımdaki gibi minimum, maksimum ve yüksek olasılıklı değerleri tanımlar. Yüksek olasılıklı değerlerin ortaya çıkması daha olasıdır. Bununla birlikte, en olası ve aşırılıklar arasındaki değerlerin üçgenden daha fazla olması muhtemeldir; yani aşırı uçlar üçgendeki kadar vurgulanmaz. PERT dağılımının kullanımına bir örnek, bir proje yönetim modelindeki bir görevin süresini tanımlamaktır.



Şekil 7. PERT dağılımı

Ayrık

Kullanıcı, meydana gelebilecek belirli değerleri ve her birinin olasılığını tanımlar. Örnek olarak bir davanın sonuçları verilebilir: %20 olumlu karar şansı, %30 olumsuz karar şansı, %40 anlaşma şansı ve %10 dosyanın yanlış değerlendirme şansı gibi.

Bir Monte Carlo simülasyonu sırasında, değerler rastgele olarak girdideki olasılık dağılımlarından örneklenir. Her örnek grubuna yineleme (iteration) denir ve bu örnekten elde edilen sonuç kaydedilir. Monte Carlo simülasyonu bunu yüzlerce veya binlerce kez yapar ve sonuç, olası sonuçların olasılıksal dağılımıdır. Bu şekilde, Monte Carlo simülasyonu ne olabileceğine dair çok kapsamlı bir görüş sağlar. Durumda sadece ne olabileceğini değil, olabileceklerin ne kadar olası olduğunu da söyler. Monte Carlo simülasyonu deterministik veya “tek nokta tahmini” analizine göre bir takım avantajlar sağlar:

Olasılıksal Sonuçlar. Sonuçların sadece ne olabileceğini değil, her sonucun ne kadar olası olduğunu göstermektedir.

Grafik Sonuçlar. Bir Monte Carlo simülasyonunun ürettiği veriler nedeniyle, farklı sonuçlara ve bunların ortaya çıkma şansına sahip grafikler oluşturmak kolaydır. Bu, bulguları diğer paydaşlara iletmek için önemlidir.

Duyarlılık analizi. Sadece birkaç vakada, deterministik analiz hangi değişkenlerin sonucu en çok etkilediğini görmeyi zorlaştırır. Monte Carlo simülasyonunda, hangi girdilerin en alt sonuçlarda en büyük etkiye sahip olduğunu görmek kolaydır.

Senaryo Analizi. Birbirlerinden gerçekten farklı senaryoların etkilerini görmek için farklı girdilere dair değer kombinasyonlarını modellemek deterministik modellerde çok zordur. Monte Carlo simülasyonunu kullanarak analistler, belirli sonuçlar ortaya çıktığında tam olarak hangi girdilerin hangi değerlere sahip olduğunu görebilirler. Bu, daha derin analiz yapılabilmesi için çok değerlidir.

Girdilerin Korelasyonu. Monte Carlo simülasyonunda, giriş değişkenleri arasındaki birbirine bağlı ilişkileri modellemek mümkündür. Modellemenin doğruluğu açısından bazı faktörlerin ne zaman arttığını, diğerlerinin de buna göre nasıl yukarı veya aşağı gittiğini temsil edebilmek önemlidir.

Monte Carlo simülasyonundaki bir gelişme de, tüm dağılım fonksiyonları serisinden daha doğru bir şekilde örnek veren Latin Hypercube örneklemesinin kullanılması olmuştur.

İstatistiklerdeki Monte Carlo deneylerinin standartları Sawilowsky tarafından belirlenmiştir. Uygulamalı istatistiklerde, Monte Carlo yöntemleri en az dört amaç için kullanılabilir:

1. Gerçekçi veri koşulları altında küçük numuneler için rakip istatistikleri karşılaştırmak. Her ne kadar Tip I hata ve istatistiklerin güç özellikleri klasik teorik dağılımlardan çıkarılan verilerle hesaplanabilse de (örn. Normal eğri, Cauchy dağılımı), asimptotik koşullar için (örn. Sonsuz örnek boyutu ve sonsuz küçük etkiler) gerçek veri, çoğunlukla böyle dağılımlar göstermez.

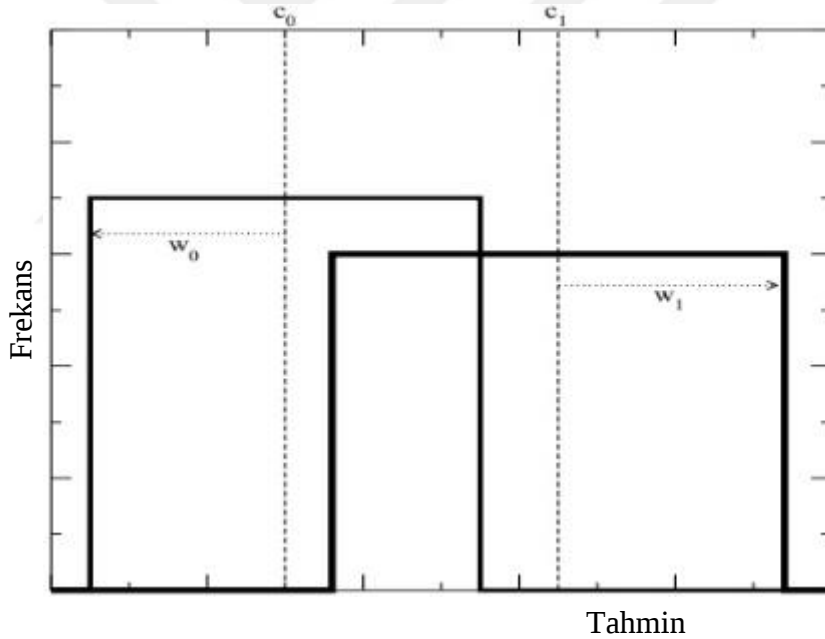
2. Genellikle hesaplanması imkansız permütasyon testleri gibi kesin testlerden daha verimli olan hipotez testlerinin uygulanmasını sağlamak. (Asimptotik dağılımlarda, kritik değerlerden daha doğru değerler verir.)

3. Bayesian çıkarımda posterior dağılımdan rastgele bir örneklem çıkarmak. Bu örneklem, posteriorun tüm temel özelliklerini özetler ve yaklaşımsal bir fikir sunar.

4. Fisher bilgi matrisinin bir tahminini oluşturmak üzere ortalaması alınabilecek negatif logaritmik fonksiyonunun Hessian matrisinin rastgele tahminlerini sağlamak.

Monte Carlo yöntemleri de yaklaşık randomizasyon ve permütasyon testleri arasında orta yolu bulmaya çalışan bir yöntemdir. Yaklaşık bir randomizasyon testi, tüm permütasyonların belirli bir alt kümesine dayanır (bu da haliyle potansiyel olarak tüm permütasyonlarının muazzam bir temizliğini gerektirir). Monte Carlo yaklaşımıysa, belirli sayıda rastgele çizilmiş permütasyona dayanır (hangi permütasyon alt kümesinin seçildiğini takip etmek zorunda olmamanın getirdiği verimlilik için, permütasyon eğer iki kez çekilirse oluşabilecek minör bir kesinlik kaybından fedakarlık yapılır.)

3.2.2. Sabit Dağılım



Şekil 8. Sabit dağılımlar.

Eşit dağılımlı tahminleri içeren genel bir durum şekilde gösterilmiştir. İlgili dört parametre vardır, iki orta (c_0 ve c_1) ve iki yarım genişlik (w_0 ve w_1). Genellikle $c_1 \geq c_0$ olduğu varsayılır. Bu durumda, yanlış alarm oranı ve isabet oranı şu şekilde verilir:

$$F = \frac{c_0 + w_0 - t}{2w_0} \quad , \quad H = \frac{c_1 + w_1 - t}{2w_1}$$

Denklem 10. Yanlış alarm oranı - İsbet oranı

Burada t , bir vakanın sınıf 1 (veya 0) olarak sınıflandırıldığı yukarıdaki (veya aşağıdaki) eşittir.

Hatırlamak gerekir ki (Denklem.10) yukarıdaki figüre özgüdür, yani değişen c_0 ve c_1 'in birbirlerine göre pozisyonları veya farklı genişlikler, farklı sonuçlar verir.

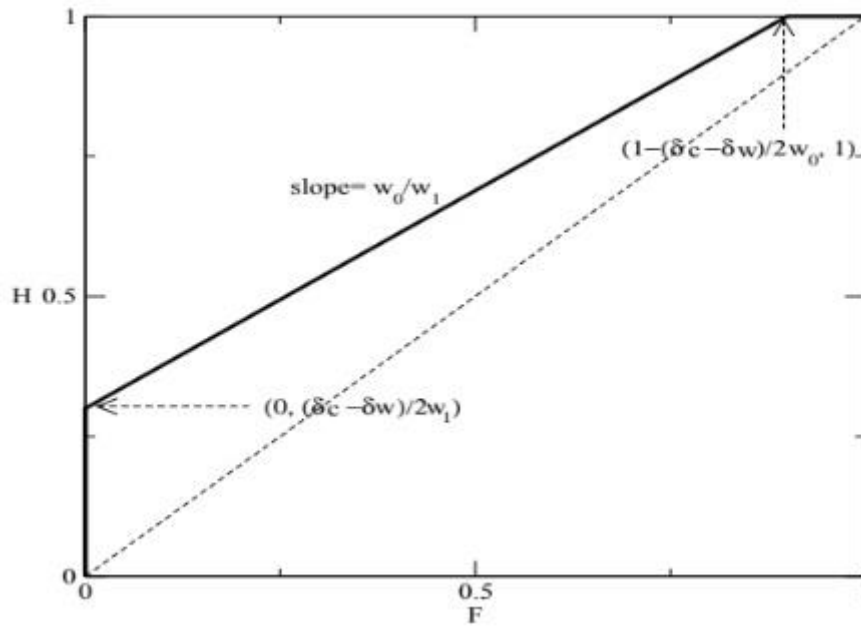
(İstatistikte genellikle c merkezi eğilimi ve w genişliğin yarısını temsil eder ve burada da bu şekilde kullanılmıştır. Gaussian dağılımda bunlar, dağılımın ortalaması ve standart sapması ile çakışır.)

Burada ROC eğrisinin denklemi, direkt olarak Denklem.10'dan gelir:

$$H = \frac{w_0}{w_1} F + \frac{\partial c + \partial w}{2w_1}$$

Denklem 11.

Aşağıdaki şekil de bunun grafiğini göstermektedir:



Şekil 9. ROC eğrisi değerleri

Burada ROC eğrisinin 3 doğru parçasından oluştuğu gözlenmenebilir ve ortadaki doğru parçasının denklemi (Denklem 11.)'dir.

Birkaç gözlem yapılabilir. İlk olarak (Denklem 11.) denklemi, farklı ortalama ve genişliklere sahip iki modelin, aynı eğime ve kesme noktalarına sahip olmaları durumunda aynı ROC eğrisini verebileceğini gösterir (Şekil.9'da görüldüğü gibi). Bu

nedenle, ROC eğrisi temeldeki parametreleri benzersiz olarak belirtmez. Başka bir deyişle, aynı ROC eğrisine neden olabilen bir dağılımlar ailesi vardır. Bu, daha genel dağılımlar için bile bilinen bir gerçektir.

İkincisi, y eksenineyle çakışan dikey segmentin uzunluğu, w_0 ve w_1 ile belirlenir. Bu mantıklıdır, çünkü altta yatan modelin ne kadar doğru olduğu bu iki miktar tarafından belirlenir. Aksine, orta segmentin eğimi sadece yarım genişliklerin oranına bağlıdır. Bu nedenle, w_0 ve w_1 eşitsizliği kendisini asimetric bir ROC eğrisi olarak yansıtır.

(Denklem.11)' den yola çıkarak, ROC eğrisi altındaki alanı (AUC) hesaplamak mümkündür:

$$AUC = 1 - \frac{1}{8} \left(\frac{\Delta}{\sqrt{w_0 w_1}} \right)^2$$

Denklem 12. Sabit Dağılımlar için AUC hesaplama

burada Δ ;

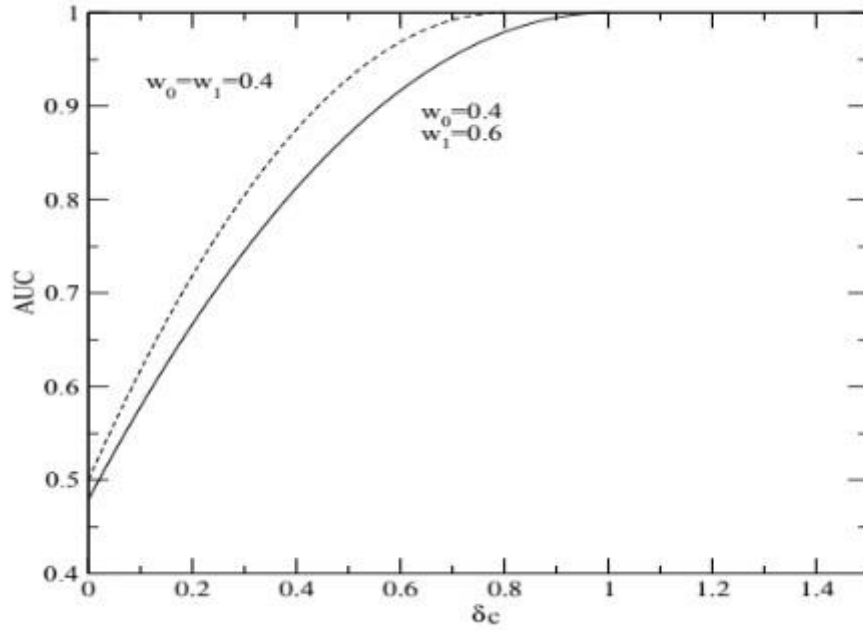
$$\Delta = \delta - (w_0 + w_1)$$

Bu alan formülü de (Denklem 10.)'dan geldiği için, yine Şekil 8.'e özgüdür.

Şekil 8.'de görülen dağılımda $\delta \leq w_0 + w_1$ olduğu için de, δ 'yi arttırmak veya w_0 veya w_1 'i azaltmak, performansı da artırır.

Kısacası, AUC'ye dayanan model seçimi keskin (yani dar genişlikte) ve birbirlerinden iyi ayrılmış sınıf koşullu dağılımları seçer. Altta yatan dağılımlar açısından, δ , w_0 , w_1 miktarlarının her birinin bir performans ölçüsü olarak yorumlanabileceğini unutmamak gerekir.

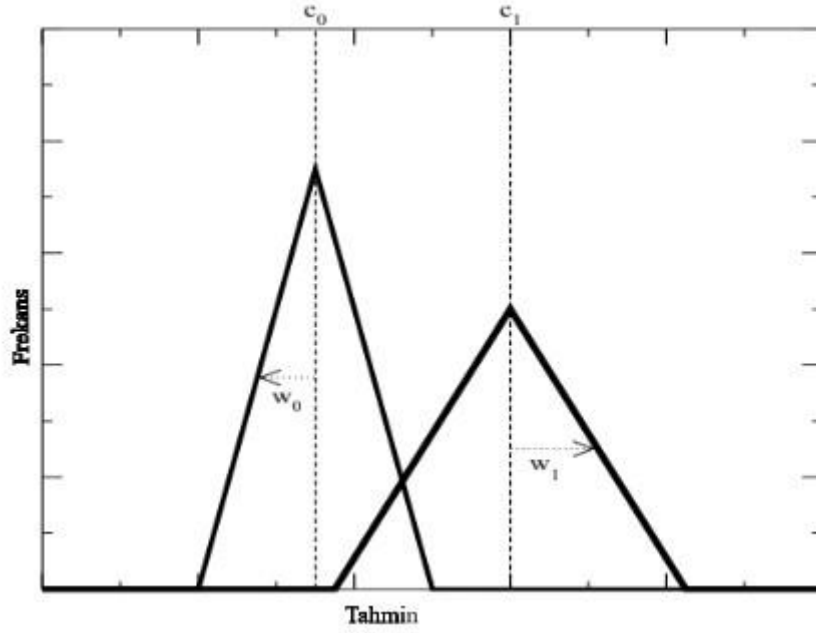
δ ölçüsünün bir fonksiyonu olarak, AUC bir parabolüdür.



Şekil 10. Parabolik AUC

Şekil 10. $w_0 = w_1 = 0.4$ ve $w_0 = 0.4, w_1 = 0.6$ için bir örneği gösterir. AUC eğrisi hızla yükselir ve sonra düzleşir. Uygulamada bire yakın AUC değerlerinin görünümünü açıklayan bu doğrusal olmayan davranıştır. Örneğin, Şekil 10., model δ_c açısından geliştikçe, AUC değeri $\delta_c \sim 0.8$ 'de hızlı bir şekilde 0.99'a yükselir. Ve $\delta_c \geq 0.8$ ile daha iyi modellerin sonsuzluğu, sadece karşılaştırılabilir AUC değerleriyle sonuçlanacaktır, yine de 0.99 civarındadır. Başka bir deyişle, pratikte yüksek AUC değerlerinin sık görülmesi, karşılık gelen modellerin hepsinin AUC eğrisinin “iyi” aralığında olduğunu göstermektedir. AUC'nin altta yatan dağılımlar açısından “iyi” ve “kötü” modeller arasında iyi ayırım yaptığını, ancak “iyi” modellerin kendileri arasında iyi ayrımcılık yapmadığı söylenebilir. w_0 ve w_1 performans ölçümleri için de benzer argümanlar geçerlidir; daha keskin dağılımlar için AUC düzleşir.

3.2.3. Sınırsız Destekli Üçgen Dağılım



Şekil 11. Sınırsız destekli üçgen dağılım

Şekilde görülen yaklaşım, yukarıda anlatılandan daha iyi ama yine de temel bir yaklaşımdır. Bu dağılım için:

$$F = \frac{1}{2} \left(\frac{c_0 + w_0 - t}{w_0} \right)^2, \quad H = 1 - \frac{1}{2} \left(\frac{t - c_1 + w_1}{w_1} \right)^2$$

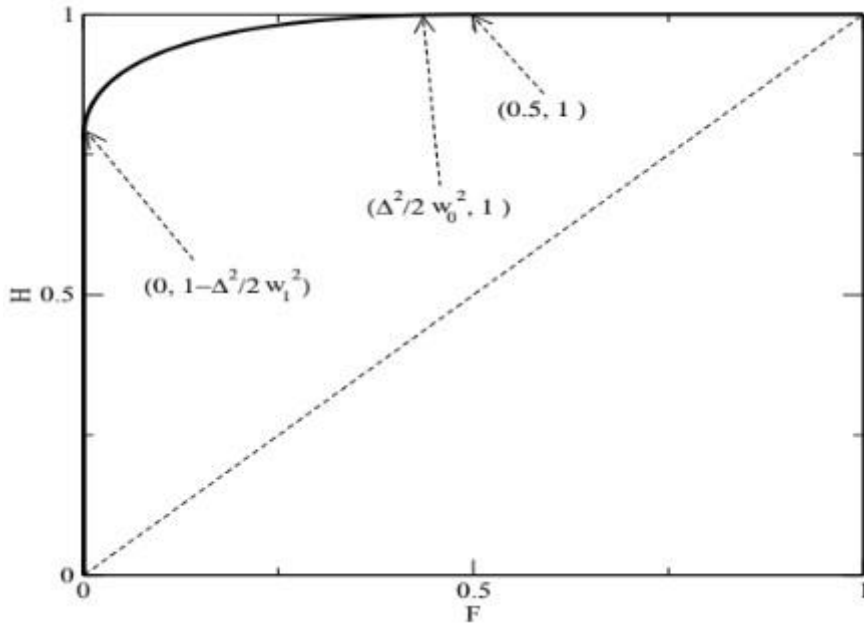
Denklem 13. Sınırsız destekli üçgen dağılım

ROC eğrisi de böylece:

$$H = 1 - \frac{1}{2} \left(\frac{\Delta - w_0 \sqrt{2F}}{w_1} \right)^2$$

Denklem 14. Sınırsız destekli üçgen dağılım ROC eğrisi

Aşağıdaki figür, ROC eğrisini göstermektedir:



Şekil 12. Asimetrik ROC eğrisi

Esasında bu ROC eğrisi, bir önceki bölümde anlatılan eğriden daha gerçekçidir. Ancak ortak özellikleri ise eksenlerle çakışma durumudur. Yukarıdaki şekildeki orta segmentlerin sınırlarından yola çıkarak, ROC eğrisinin yalnızca $w_0 \neq w_1$ olduğunda asimetrik olduğu söylenebilir. Özellikle, eğer içbükeylik çoğunlukla sola doğruysa, $w_0 < w_1$ 'dir. Eğrinin sağa doğru eğilmesiyle $w_0 > w_1$ olduğunu gösterir. Asimetrinin c_i 'den bağımsız olduğunu unutmamak gerekir.

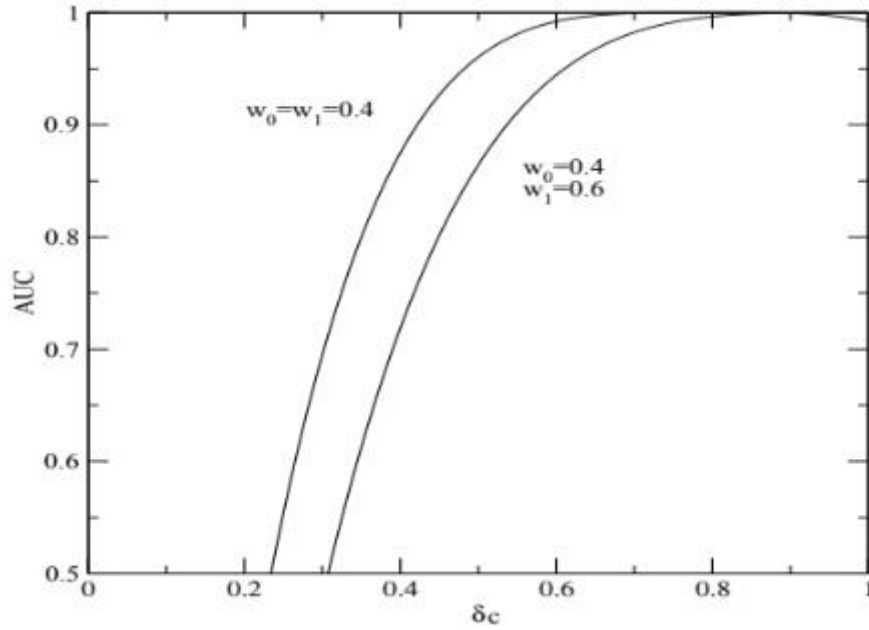
Ayrıca, eğrilerin iki ucu ($F = 0$ ve $H = 1$) bazı yararlı bilgiler de taşır. Eğer $\delta w = \delta c$ ise, eğrinin sağ ucunun $H = 1$ çizgisiyle çakışmadan $(1,1)$ noktasından geçtiğini unutmamak gerekir. Benzer şekilde, $\delta w = -\delta c$ eğrinin sol ucunun $F = 0$ eksenine çakışmadan $(0,0)$ noktasından geçtiğini gösterir. Bu nedenle, eğrinin üst üste binme miktarı ve iki eksen, yarı genişlikler arasındaki farka göre iki orta arasındaki mesafenin bir ölçüsüdür.

Eğri altındaki alan (AUC) şu şekilde hesaplanabilir:

$$AUC = 1 - \frac{1}{8} \left(\frac{\Delta}{\sqrt{w_0 w_1}} \right)^4$$

Denklem 15. Sınırsız üçgen dağılım için AUC hesaplama

Bir önceki bölümdeki AUC ifadesi gibi (Denklem.14) bu ifade de Δ miktarına bir meyil gösterir. Ayrıca, (Denklem.14)'deki ikinci derecede (kuadratik) güce kıyasla Δ 'nın dördüncü dereceden gücü dikkate alındığında, bu AUC'nin daha hızlı yükselmesi ve daha geniş bir platoya sahip olması nedeniyle doğrusal olmama mertebesinin daha yüksek olduğu açıktır. Aşağıdaki şekil, bu dörtlü bağılılığı göstermektedir:



Şekil 13. w_0 ve w_1 e bağlı AUC değeri

AUC eğrisinin daha da düzleştirilmesi AUC'nin iyi modeller arasında ayırım yapamamasını daha da artırmaktadır.

3.2.4. Gauss Yöntemi ile Hesaplama

Gauss yöntemi, ROC eğrisi altı alan (ROC-AUC) hesaplamasında pratik olması bakımından en sık kullanılan yöntemlerdendir. Klinik çalışmalarda baz alınan Gauss yöntemi, klinisyenlerin alışkanlık kazanmış olduğu bir yöntem olmuştur. Özellikle iki farklı grup verinin karşılaştırılması söz konusu ise Gauss yöntemi diğer yöntemlere göre daha efektif olduğu gösterilmiştir. Bu yöntem, esnek uygulama imkanından dolayı atipik verileri daha net göstermekte ve değerlendirmede kolaylık sağlamaktadır.

Sınırsız tahmin miktarlarıyla ilgilenen üç dağıtım arasında Gaussian en gerçekçi yaklaşımı sunar. Bununla birlikte, Gauss yöntemiyle hesaplanan ROC ve AUC ifadeleri, bazı integrallerin ortaya çıkması nedeniyle diğerleri kadar açık değildir.

“i” seviyesindeki tahminlerin olabilesi şu şekilde hesaplanır:

$$L_i(x) = \left(\frac{1}{\sqrt{2\pi w_1^2}} \right)^4 \exp^{-\frac{1}{2} \left(\frac{x-c_1}{w_1} \right)^2}$$

Denklem 16.

$$H = \int_t^\infty L_1(x) dx \quad , \quad F = \int_t^\infty L_0(x) dx \quad ,$$

Denklem 17.

$$F = \phi \left(\frac{c_0 - t}{w_0} \right) \quad , \quad H = \phi \left(\frac{c_1 - t}{w_1} \right)$$

Denklem 18.

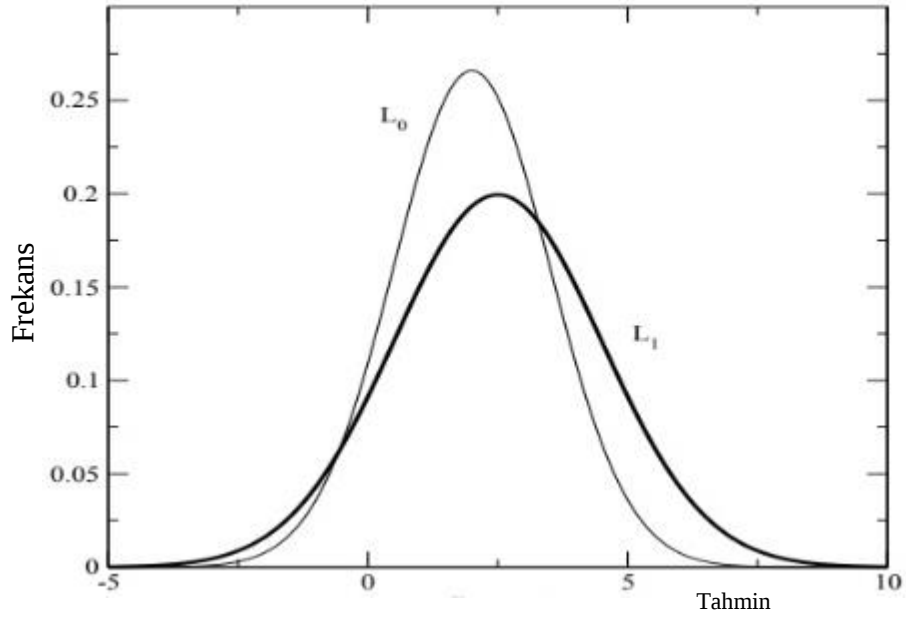
$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp^{-\frac{1}{2}z^2} dz$$

Denklem 19.

($\phi(x)$ standart normal kümülatif dağılımı göstermektedir.)

Bu denklemlerden eşik t 'yi çıkardığımızda, ROC eğrisinin denklemini görürüz:

Bu ifadelerin kastı da aşağıdaki şekilde görülebilir:

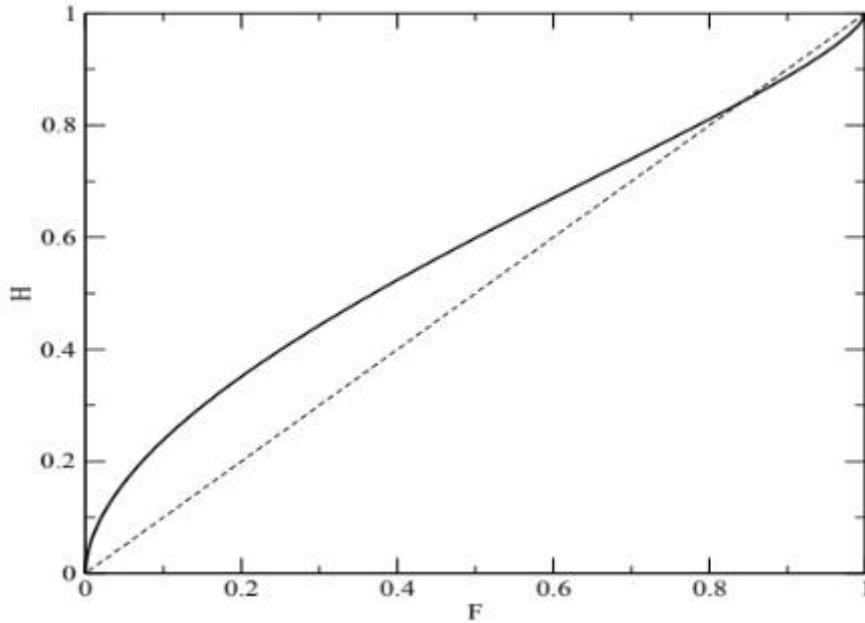


Şekil 14. Gauss eğrisi.

ROC eğrisinin bu eşitliği çok aydınlatıcı olmasa da, bir takım yararlı hesaplamalara izin verir. Örneğin, bir çift olasılık kağıdına çizilirse, ROC eğrisinin $w_0 /$

w_1 eğimine sahip ve $\delta c / w_1$ 'e kesişen düz bir çizgi olacağı anlamına gelir. (Burada bu denklemin, uniform dağılımdaki ROC eğrisi formülüyle olan benzerliği dikkat çekmektedir.) İkinci olarak, ROC eğrisinin eğiminin $L_1(t)/L_0(t)$ olduğunu hesaplamaya izin verir. (Kişinin “optimal” bir karar eşiği aradığı karar teorik uygulamalarında, bu ifade eğimin 1 olduğu eşik değeri tartışmak için verilir. Bununla birlikte, bu seçim iki sınıfın p_1 denen eşit ön olasılıklara sahip olduğunu varsayar. p_1 ve p_0 'a bazen sırasıyla taban oran ve bunun tamamlayıcısı denir. Optimal eşik ise eğim = p_0 / p_1 'e karşılık gelen eşik olmalıdır.) Bu eğime (Denklem 16.) i koyduğumuzda, ROC eğrisinin uçlarındaki eğimin ya 0, ya da ∞ olduğunu bize gösterecektir. Bir diğer deyişle, ROC eğrisi her zaman eksenlere teğettir.

Burada yapılan yaygın bir hata, Gauss dağılımlarına dayanan teorik bir ROC eğrisinin, kanonik ROC davranışına, yani diyagonalin üstünde veya altında içbükey olarak kısıtlanmış olduğunu varsaymaktır. Bu, $w_0 = w_1$ olan simetrik durum için doğru olsa da, genel olarak ROC eğrisi içbükeylikle sınırlı değildir. Bunu göstermenin en kolay yoluysa, $w_0 \neq w_1$ olduğunda, ROC eğrisinin (bitiş noktaları dışında) diyagonalı tam olarak tek bir noktadan kestiğini göstermektir. Bu kesim aşağıdaki şekilde görülmektedir:



Şekil 15. İçbükey ROC eğrisi.

Bu sonuç dikkatle yorumlanmalıdır. Özellikle, görünüşte içbükey bir ampirik ROC eğrisinin $w_0 = w_1$ eşitliğini önerdiği anlamına gelmez. $w_0 \neq w_1$ olduğunda bile, ROC eğrisi hala çoğunlukla içbükey görünebilir. Çünkü $\Phi(x)$, x 'in hızla artan bir

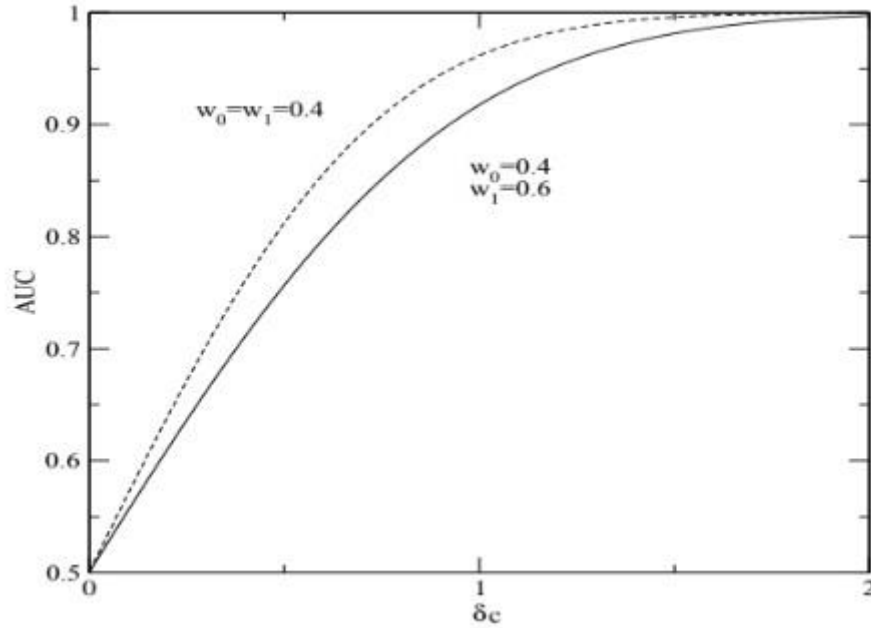
fonksiyonudur. Aslında, x sırasıyla $+2$ veya -2 olduğunda yaklaşık 0 veya 1'dir. Bu nedenle, içbükey ampirik bir ROC eğrisi iki olasılıktan birini önerir: $w_0 = w_1$ veya $w_0 \neq w_1$ ancak hep $|\delta c \delta w| \geq 2$ ile.

Eğim altındaki alan (AUC) böylece hesaplanır:

$$AUC = \Phi\left(\frac{\delta c}{\sqrt{w_0^2 + w_1^2}}\right)$$

Denklem 20. AUC hesaplama denklemi

AUC burada yine, performansı değerlendiren tüm temel parametrelerin (δc , w_0 ve w_1) doğrusal olmayan bir fonksiyonudur. δc 'nin işlevsel bağıllığı aşağıdaki şekilde gösterilmektedir:



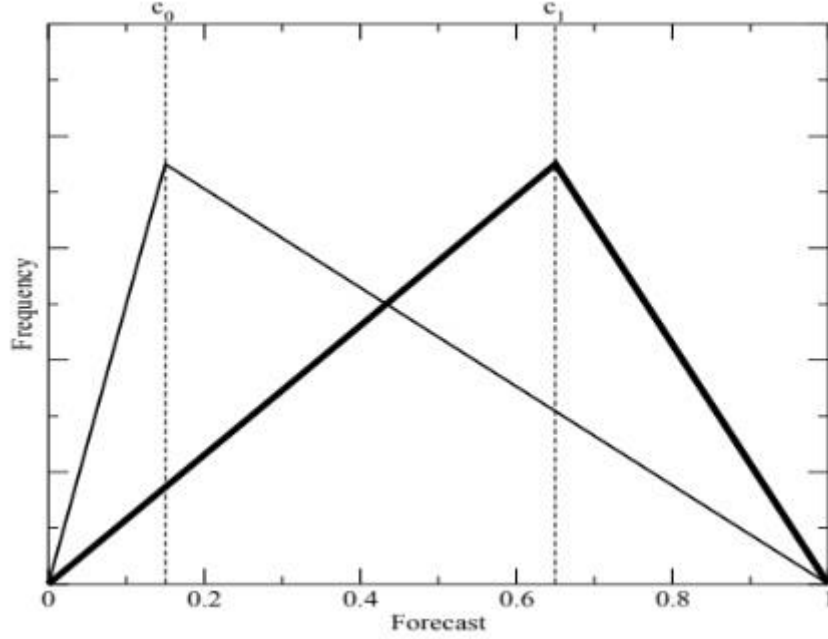
Şekil 16. δc değişimi ve ROC eğrisi.

Açıkçası, bu gerçekçi örnekte bile eğrinin doğrusal olmama durumu mevcuttur. Yine, biri diğerinden belirgin şekilde üstün olan (örneğin, δc 'nin farklı değerlerine sahip) iki “iyi” model, karşılaştırılabilir ve yüksek AUC değerlerine sahip olabilir.

Gösterdiğimiz AUC eşitliği Denklem.20 ise pratikte ampirik AUC değerlerinin neden genellikle 0.9 veya daha yüksek aralıkta olduğunu açıklar. Nedeni, tekrar $\Phi(x)$ 'in davranışına kadar takip edilebilir. Daha önce bahsedildiği gibi, 2 gibi mütevazı derecede büyük x değerleri için Φ , bire yakın değerlere karşılık gelir.

3.2.5. Sınırlı Destekli Üçgen Dağılım

Bazı durumlarda tahmin edilen miktar bir olasılıktır ve bu aralıkla sınırlı dağılımlar gerektirir. Burada ele alınan bu tip iki dağılımdan ilki aşağıdaki şekilde gösterilmiştir:



Şekil 17. Üçgen dağılım.

Bu model, tahminlerin (0 ile 1 arasındaki) tüm olasılıkları kapsadığını varsayar. Murphy ve Winkler'in dilinden, tahminlerin "iyi rafine edildiği" varsayılmaktadır. Ayrıca bu yaklaşımda, parametrelerin yalnızca iki mod (c_0 ve c_1) olduğuna dikkat edilmelidir. (Bu bölümde, c ortalamayı değil, dağılım şeklini temsil eder. Ayrıca dağılımların genişlikleri bağımsız büyüklükler değildir. Ortalama $(1 + c)/3$ ve varyans $1 - c + c^2/18$ olarak verilmiştir.)

Burada üç farklı bölge dikkate alınmalıdır: $t \leq c_0$, $c_0 \leq t \leq c_1$ ve $t \geq c_1$. Önceki örneklerin aksine, burada eksenlerle örtüşen bir bölge yoktur; bu, tahminlerin düzeltilmesi ile ilgili yukarıda belirtilen varsayımın bir sonucudur. Bu alanlar için ROC eğrileri sırasıyla şu şekildedir:

$$H = 1 - \frac{c_0}{c_1}(1 - F)$$

Denklem 21.

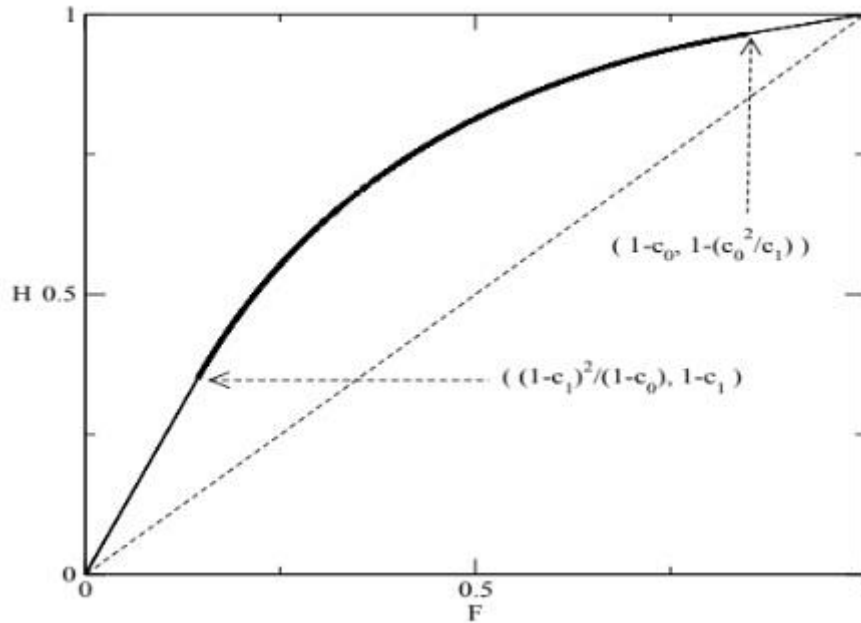
$$H = 1 - \frac{1}{c_1} \left[(1 - \sqrt{(1 - c_0) F}) \right]$$

Denklem 22.

$$H = \left(\frac{1 - c_0}{1 - c_1} \right) F$$

Denklem 23.

Birinci ve üçüncü bölgeler için ROC eğrisinin doğrusal olduğunu, orta bölümün ROC eğrisinin doğrusal olmadığını unutulmamalıdır. ROC eğrisi aşağıdaki şekilde gösterilmektedir:



Şekil 18. ROC eğrisi değişkenliği.

Eğimin tanımından, simetrik bir ROC eğrisinin $c_0 + c_1 = 1$ anlamına geldiği anlaşılmaktadır. c_0 ve c_1 'in herhangi başka bir kombinasyonu asimetrik bir eğriye neden olacaktır. Ayrıca, bir çapraz geçişin olmadığını göstermek de kolaydır; bu şekilde en az bir değişkeni sıfırdan farklı olan bir eğri her zaman diyagonalin ya üstünde ya da altındadır. Ayrıca ROC eğrisinin $c_0 \sim 0,5$ ise sola, $c_1 \sim 0,5$ ise sağa doğru eğileceğini de gösterir.

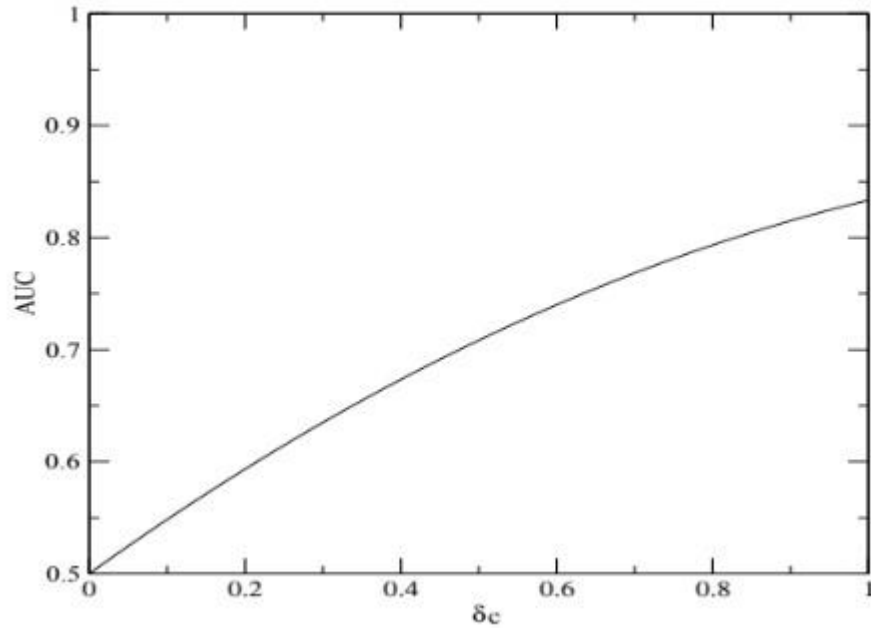
Eğri altındaki alan (AUC) da aşağıda verildiği gibi olur:

$$AUC = \frac{1}{2} + \frac{1}{2}(c_1 - c_0) - \frac{(c_1 - c_0)^3}{6c_1(1 - c_0)}$$

Denklem 24.

İlk olarak, AUC'nin $(c_1 - c_0)$ ve $c_1 (1 - c_0)$ olarak iki bağımsız değere bağlı olduğuna dikkat edin. Birincinin küçük değerleri, yani düşük performans için, Denklem.16'daki ilk iki terim eşitliğe hakim olur ve $c_1 - c_0$ üzerinde doğrusal bir bağımlılığa yol açar. Ancak, daha yüksek performans değerleri için, eşitlikteki son değer (negatif işaret nedeniyle) AUC'yi doğrusal olmayan bir şekle büründürür. Bu doğrusal olmayan eğilim de yine daha iyi modeller için AUC eğrisinin düzleşmesine yol açar.

Aşağıdaki şekil, AUC'yi $c_1 - c_0$ ölçüsünün bir fonksiyonu olarak göstermektedir:



Şekil 19. Doğrusal özelliği olmayan ROC eğrisi.

Düzleştirmenin şekilde belirgin olmamasının nedeni, modelin sadeliğinin yüksek AUC değerlerine izin vermemesidir. Aslında, Denklem 24.'e göre izin verilen en yüksek AUC değeri sadece 5/6 veya 0.83'tür.

3.2.6. Beta Dağılımları

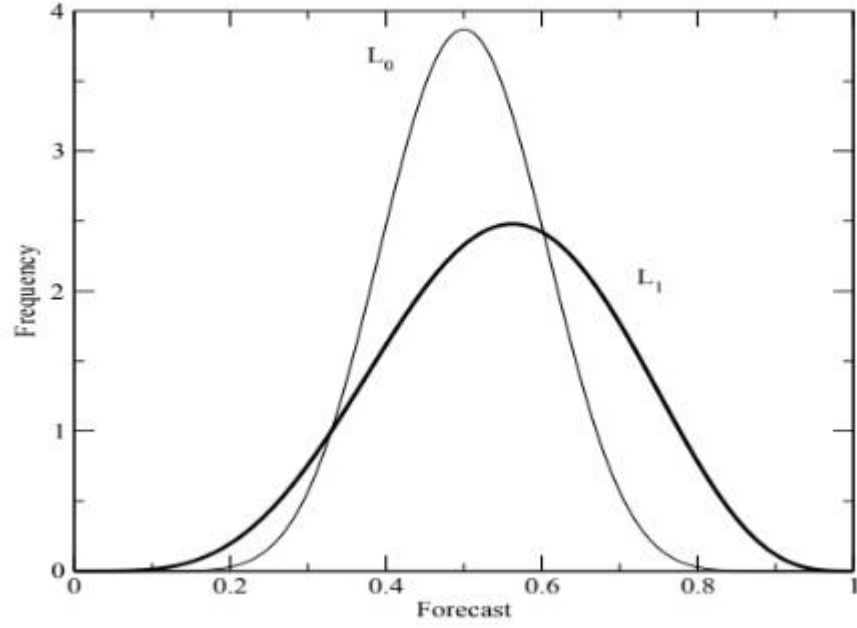
İstatistiksel tahminlerin daha gerçekçi bir gösterimi beta dağılımlardır:

$$L_i(x) = \frac{1}{B(a_i, b_i)} x^{a_i-1} (1-x)^{b_i-1},$$

Denklem 25.

Burada $B(a_i, b_i) = \int_0^1 x^{a_i-1} (1-x)^{b_i-1}$ olarak tanımlanabilir.

Aşağıda bir örnek gösterilmiştir:



Şekil 20. Beta dağılımı.

Bu örnekte, dağılımların kendilerinin muhtemelen asimetrik olduğuna dikkat edilmelidir. Eğer a_i ve b_i tamsayıysa, şu ifade doğrudur:

$$B(a_i, b_i) = \frac{(a_i-1)!(b_i-1)!}{(a_i+b_i-1)!}$$

Denklem 26.

Ortalama, mod ve varyans sırasıyla şu şekilde gösterilebilir:

$$c_i = \frac{a_i}{a_i+b_i}$$

Denklem 27.

$$m_i = \frac{a_i-1}{a_i+b_i-2}$$

Denklem 28.

$$w_i^2 = \frac{a_i \cdot b_i}{(a_i+b_i)^2(a_i+b_i+1)}$$

Denklem 29.

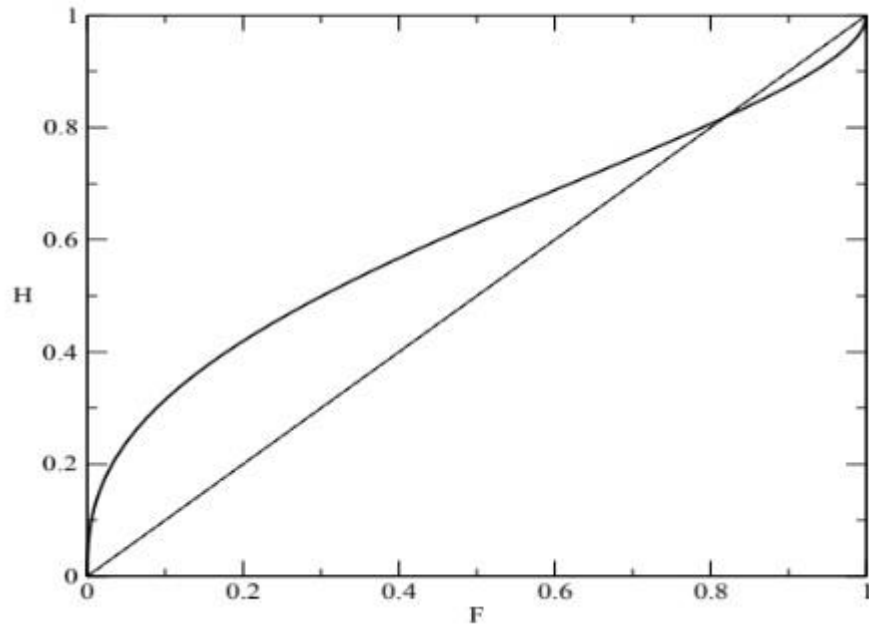
Bu durumda, olasılıkların a_i ve b_i cinsinden yazıldığı göz önüne alındığında, bu miktarların hangi kombinasyonunun performansın bir ölçüsü olduğunu sormak doğaldır.

Teorik karara dayalı bakış açısından, doğal miktar $L_1(x) / L_0(x)$ 'dir ve bu oran $(a_1 - a_0)$ ve $(b_1 - b_0)$ ' ın bir fonksiyonudur. (Teknik olarak, bu ifade ilgili önceki olasılıkların oranı ile çarpılmalıdır. x 'in fonksiyonları olmadıkları içinse burada ihmal edilmektedir.) Bu nedenle, bu iki fark doğal performans ölçütleridir. Bu önlemlerin her birinin hem c hem de w 'ye bağlı olduğuna da dikkat edilmelidir. Örneğin:

$$a_1 - a_0 = (c_1 - c_0) + \left[\frac{c_1^2(1-c_1)}{w_1^2} - \frac{c_1^2(1-c_1)}{w_0^2} \right]$$

Denklem 30.

Bu eşitliğe tekabül eden ROC eğrisi aşağıda gösterilmiştir:



Şekil 21. Beta dağılımına bağlı ROC eğrisi

F ve H için analitik ifadeler aydınlatıcı olmasa da ROC eğrisinin eğimi:

$$\text{slope}(t) = \frac{B(a_0, b_0)}{B(a_1, b_1)} t^{a_1-a_0} (1-t)^{b_1-b_0}$$

Denklem 31.

Simetrik bir ROC eğrisi, eğrinin uç noktalarındaki eğimlerin çarpımının ters orantılı olmasını gerektirir. Ve bunun gerçekleşmesi için $(a_1 + b_1) = (a_0 + b_0)$ olmalıdır. Yani $(a_0 + b_0) = (a_1 + b_1)$ ise, ROC eğrisi simetriktir, bu da ortalar ve varyanslar açısından aşağıdakine tekabül eder:

$$\frac{c_1(1-c_1)}{w_1^2} = \frac{c_0(1-c_0)}{w_0^2}$$

Denklem 32.

Ampirik bir ROC eğrisinde görünen bir asimetri, bu denklemin ihlal edildiğini ima eder. Simetrik ROC durumunda, iki performans ölçümü olan $a_1 - a_0$ ve $b_1 - b_0$ 'ın yalnızca işaret bakımından farklılık gösterdiğine dikkat edilmelidir.

Ayrıca, $a_1 > a_0$ ve $b_1 > b_0$ olduğunda bir çaprazlama meydana gelir, çünkü iki uçtaki eğimler daha sonra 1'den daha azdır. Bu iki eşitsizlik birlikte, şunu gösterir:

$$\frac{c_1(1-c_1)}{w_1^2} > \frac{c_0(1-c_0)}{w_0^2}$$

Denklem 33.

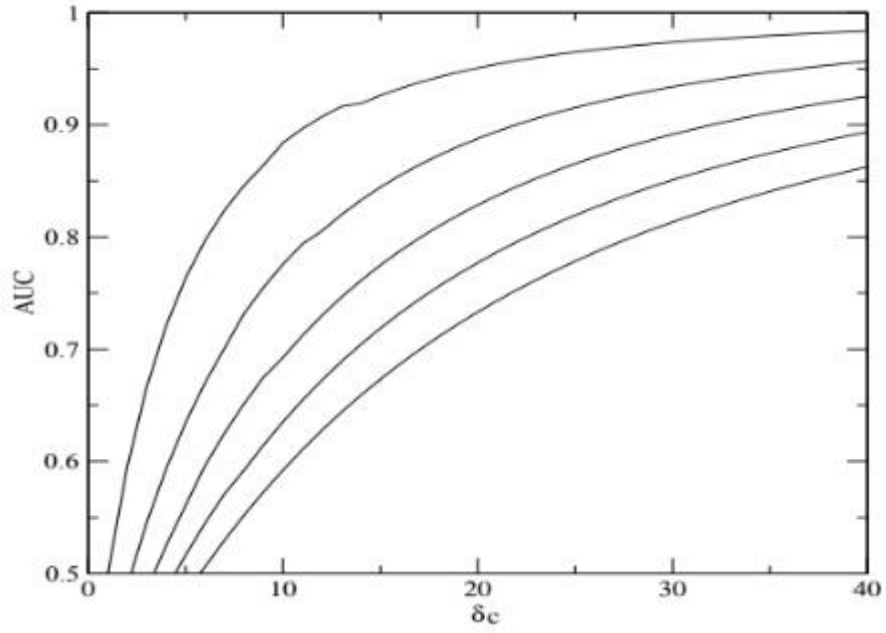
Simetrik bir ROC eğrisinin koşulu olan Denklem 32. ile karşılaştırıldığında, $c(1-c)/w^2$ değeri, ROC eğrisinin hem simetrisini hem de çapraz geçişini belirler. Çapraz geçiş ise Şekil 21.'de gösterilmektedir.

AUC eşitliğini türetmek biraz yorucudur, ancak tamsayı parametreleri için şu şekilde hesaplanabilir:

$$AUC = \frac{1}{(a_1+b_1)} \frac{1}{B(a_0,b_0)} \sum_{k=1}^{a_1} \frac{B(a_0+a_1-k, b_0+b_1+k-1)}{B(a_1-k+1, b_1+k)}$$

Denklem 34.

Aşağıdaki şekil, parametrelerin birkaç farklı değeri için $(a_1 - a_0)$ ölçümünün bir fonksiyonu olarak bir AUC grafiğini göstermektedir. AUC 1'e yakın değerlere ulaştığında doğrusal olmama durumu açıkça görülmektedir:

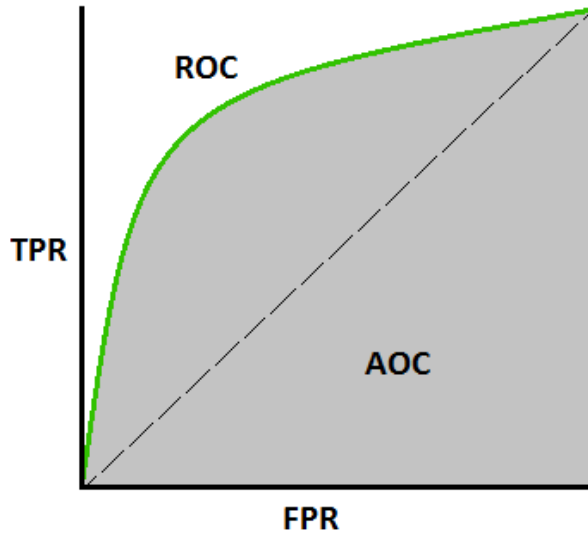


Şekil 22. Parametre değişikliğine bağlı AUC.

3.3. ROC AUC Yorumlama

AUC - ROC eğrisi, esas olarak eşik değerine bağlı bir performans ölçüm tekniğidir. ROC bir olasılık eğrisidir ve AUC ayrılabilirliğin derecesini veya ölçüsünü temsil eder. Modelleri sınıflandırarak, grupları kendi aralarında ne ölçüde ayırt edebildiğini anlatır. AUC yükseldikçe, gruplama modellerindeki tahmin başarısı da artma eğilimindedir. Özetle, AUC ne kadar yüksek olursa, kullanılan yöntemin hastalığı olan ve olmayan durumları ayırt etmekte gücü yükselir.

ROC eğrisi, TPR'nin y ekseninde ve FPR'nin x ekseninde olduğu FPR'ye karşı TPR ile çizilir.



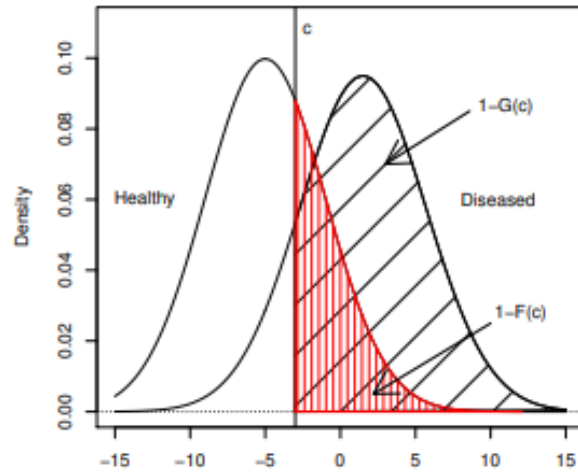
Şekil 23. ROC -AUC yorumlama

Özellikleri ve içerdiği anlam itibarıyla ROC eğrileri ve eğrinin altındaki alanın (AUC) çeşitli modellerle incelenmiştir. Elde edilen veriler, ROC eğrisinin şekline bağlı olarak, tahmin edilen toplumun sınıfsal dağılımını tanımlayan parametreler açısından yorumlanmasına yardımcı olmaktadır. ROC analizi modelleri ve AUC hesaplama teknikleri AUC nin yorumlanması için rehberlik ettiği görülür. Hesaplama kullanılan modeller yorumun niteliğini ve güvenilirliğini de doğrudan etkiler. Bu nedenle, sonuçların genelliği hiçbir şekilde garanti edilmez. Sınırsız tahminler için, asimetrik bir ROC eğrisi, temel dağılımlar için eşit olmayan genişlikler önerir. Daha büyük ortalamaya sahip sınıf 1 olarak etiketlenmişse, tepeye bir içbükeylik $w_0 > w_1$ ve tabana konkavlık, $w_0 < w_1$. Diğer bir yaklaşımla, ampirik bir ROC eğrisindeki herhangi bir asimetriyi açıklamaya çalışırken, temel dağılımların genişliklerinin incelenmesi tavsiye edilir. Eksenlerle üst üste binme miktarı da genişliklerdeki farkın bir ölçüsüdür. Köşegenin bir ROC eğrisi ile kesilmesi miktarın $|\delta c \delta w|$ bazı kritik değerlerden daha küçük olduğunu açıklar. Örneğin, dağılımlar gaussiyse, o zaman bu kritik değer yaklaşık 2'dir. Sınırlı tahminlerde eksenlerde çakışma beklenmez. Ampirik bir ROC grafiğinde önemli bir çakışma olması, temel dağılımların önemli ölçüde farklı olduğunu göstermektedir. Sınırlı ve sınırsız tahminler için AUC, temel dağılımların parametrelerinden elde edilen doğal tahmin kalitesi ölçülerine bağlı olarak doğrusal olmayan bir şekilde artmaktadır. Bazı gerçek veri modellerinde, doğrusal olmama durumundan daha fazlasını gösterir. Doğrusal olmama

durumu, performans arttıkça performansın değerlendirilmesinde AUC'nin etkinliğini azaltacak şekildedir. Bu nedenle, ampirik olarak gözlemlenen bire yakın AUC değerlerinin sık görülmesi, birçok tahminin “makul” kalitede olduğunun bir göstergesidir.

3.4. Parametrik Yaklaşımlar – Binormal

X ve Y rastgele iki bağımsız değişken iken hasta popülasyon için ($D=0$) ve sağlıklı popülasyon için ($D=1$) olarak ele alalım. Genellik kaybı olmadan ve uygun bir kesme noktası c için, test sonucu c 'den büyükse pozitif, aksi takdirde negatif olarak tanımlarız. F ve G 'yi ise sırasıyla X ve Y rastgele değişkenlerinin dağılım fonksiyonları olduğunu kabul edelim. Bu durumda testin duyarlılığı $Se(c) = 1 - G(c)$ ile verilir ve özgüllük $Sp(c) = F(c)$ olarak tanımlanır.



Şekil 24. Binormal dağılım

-15 -10 -5 0 5 10 15 0,00 0,02 0,04 0,06 0,08 0,10 x Yoğunluk c Sağlıklı
Hastalıklı $1 - G(c)$ $1 - F(c)$

ROC Eğrisi Tahmini: ROC eğrisi, $-\infty \leq c \leq \infty$ için $1 - Sp(c)$ 'ye karşı $Se(c)$ 'nin grafiği olarak veya eşit olarak (2.1) $ROC(t) = 1 - GF^{-1}(1 - t)$, $t \in [0, 1]$ üzeri, burada $F^{-1}(1 - t) = \inf x \in R: F(x) \geq 1 - t$. ROC eğrisi, X ve Y değişkenlerinin dönüşümünü artıran herhangi bir monoton altında değişmekte ve değişmemektedir. Literatürde, eğrinin altındaki alan (AUC) veya Youden indeksi ($\max\{Se(c) + Sp(c) - 1\}$) gibi çeşitli ROC eğrisi özet ölçümleri önerilmiştir. Bir testin ayırmacı doğruluğunun özeti olarak kabul edilirler. AUC, (2.2) $AUC = \int_0^1 ROC(u) du$. ROC eğrisini tahmin etmek için farklı yaklaşımlar AUC'nin farklı tahminlerine yol açar. AUC, rastgele seçilmiş bir grup hastalısız ve hastalıklı bireyde, tanısal test değerinin hastalıklı denek için daha yüksek olma olasılığı, yani $AUC = P(YX)$ olasılığı olarak

yorumlanabilir. 1'e yakın AUC değerleri, testin veya markerin yüksek tanısal doğruluğunu gösterir. Ampirik ROC eğrisinin alanı, AUC için tarafsız bir parametrik olmayan tahmin sağlayan Mann-Whitney U istatistiğine eşittir (Faraggi ve Reiser, 2002). Bamber'in (1975) seminal çalışmasından bu yana, birkaç yazar, örneğin aşağıda tarif edilecek çekirdek yöntemini kullanarak yumuşatılmış ROC eğrileri elde etmek için parametrik olmayan yaklaşımı geliştirmeyi önermiştir. ROC eğrisinin parametrik tahmini de aktif bir araştırma alanıdır ve F ve G için çeşitli öneriler düşünülmektedir. En yaygın olarak kullanılan parametrik ROC modeli, bir sonraki bölümde açıklanan binormaldir.

Hastaliksız ve hasta popülasyonlarda F ve G bilindiğinde parametrik yöntemler kullanılır. Binormal model yaygın olarak kabul edilir ve hem hastalıklı hem de hastaliksız test sonuçları normal dağılımları takip ettiğinde uygulanabilir (Faraggi ve Reiser, 2002). Veriler gerçekten binormal ise veya logaritma veya karekök gibi bir Box-Cox dönüşümü veriyi binormal yapıyorsa, ilgili parametreler hastalıklı ve hasta olmayan popülasyonlarda test değerlerinin ortalamaları ve varyansları ile kolayca tahmin edilebilir. X ve Y, ortalama değerleri μ_0 , μ_1 ve varyansları σ_0^2 , σ_1^2 olan bağımsız normal değişkenler olsun. Böylece, ROC eğrisi şu şekilde özetlenebilir:

$$\text{ROC}(t) = \Phi\left(\frac{a + b \Phi^{-1}(t)}{\sigma_1}\right), 0 \leq t \leq 1$$

Burada, Φ standart normal dağılım fonksiyonudur ve a ve b sırasıyla $a = (\mu_1 - \mu_0) / \sigma_1$ ve $b = \sigma_0 / \sigma_1$ ile verilen ayırma ve simetri katsayılarıdır. Bu durumda, AUC, (3.2) $AUC = \Phi\left(\frac{a}{\sqrt{1 + b^2}}\right)$ ile verilen kapalı bir forma sahiptir.

Bir ROC eğrisine uyacak binormal tahminin seçimi genellikle teorik hususlar, matematiksel izlenebilirlik, normal modele aşinalık veya sadece uygunluk ile gerekçelendirilir. Hanley (1988), bu tahminin kullanılması lehine en yaygın argümanları özetleyen bir tablo sunar ancak bazı yazarlar, binormal tahminin sağlam olduğunu da iddia ediyor. Tahminde sağlam kelimesinin birçok farklı anlamı olabilir. Burada sağlam istatistikler anlamında kullanılır, yani normal olmayan bir dağılımdan gelen belirli bir miktarda gözlemin varlığında binormal tahmin edicinin güvenilir sonuçlar vereceği anlamına gelir. Son zamanlarda, sağlık bilimlerinde kullanılan parametrik veya yarı parametrik modellerin yanlış tanımlanmasının etkisi önem kazanmaktadır, çünkü uygulayıcılar teorik modellerin gerçekliğin yalnızca yaklaşık değerleri olduğunu ve model farklılıkları altında güvenilir sonuçlar veren istatistiksel prosedürlerin gerçek problemleri çözmek için gerekli olduğunu farkındadır. ROC eğrisinin binormal tahmini

durumunda, Swets (1986) gibi yazarlar, "Deneysel psikoloji ve çeşitli pratik alanlardan alınan deneysel ROC'ların (...) binormal grafiğe iyi oturduğunu ..." öne sürer. Bu ifade, "... binormal tabanlı uyumların kesinlikle tüm pratik amaçlar için yeterince iyi olduğunu" iddia eden Hanley (1988) tarafından güçlendirilmiştir. Hajian-Tilaki vd. (1997), "Sonuçlar, LARROC programında uygulandığı gibi ikili normal model kullanılıyorsa, AUC'nin binormaliteden sapmalara karşı sağlam olduğunu" öne sürmektedir. Swets (1986) ve Hanley (1988) 'de sağlamlık, binormal tahminin teorik ROC eğrisi veya gözlemlenen derecelendirme yöntemi ile karşılaştırıldığında "doğru görünen bir ROC eğrisine uyma yeteneği" olarak anlatılır. Üstelik binormal tahminin, verilerin normallik varsayımını karşılamadığı durumlarda geçerli çıkarımlar üretme yeteneği tartışılmıştır.

3.5. Non-Parametrik ROC Eğrisi

3.5.1. Ampirik tahmin yöntemleri

Parametrik olmayan tahmin yöntemlerinin en pratik olanı "Ampirik" tahmin yöntemidir. ROC eğrisinin ampirik tahmini şu şekilde verilir:

$$ROC(t) = 1 - Ge(Fe^{-1}(1 - t))$$

burada Fe^{-1} ve Ge sırasıyla ampirik kuantil fonksiyonu ve hastalıklı/sağlıklı popülasyonla ilişkili ampirik dağılım fonksiyonunu göstermektedir. Kabaca, t için verilen herhangi bir değer için örnek noktaların yüzde oranı t değerinden küçük yada t değerine eşitliği ampirik dağılım fonksiyonu olarak tanımlanabilir.

Ampirik ROC eğrisi, ampirik dağılım fonksiyonunun birçok özelliğini korur ve teorik eğriye yakın bir eğri olarak karşımıza çıkar (Hsieh and Turnbull, 1996). Bununla birlikte yorumlayıcının bazı dezavantajları da vardır. Özellikle küçük örneklem gruplarında büyük değişkenlikler dağılımın yorumlanmasını zorlaştırır. Bu durum endüstrinin diğer alanlarında mesela makine eğitimlerinde, veri madenciliğinde veya finans alanında önemli bir problem olmayabilir. Ancak klinik çalışmalar çalışmanın içeriği itibarıyla zaman zaman büyük örneklem grubuna imkan tanımayabilir. Bu durumda yetersiz yada güvenilir yorumlama yapma imkanı zorlaşır. Tüm bunlarla birlikte, tahmini ROC eğrisi devamlılık arzetsmeyebilir ve bu durum yorumun daha karmaşık olmasına neden olabilir. (Jokiel-Rokita and Pulit, 2013). Daha net ROC eğri beklentileri için "kernel smooting" yöntemleriyle daha esnek bir yorum ortamı oluşturulabilir (Lloyd,

1998; Lloyd and Yong, 1999). Ya da ampirik dağılım fonksiyonunun daha yumuşak türleri tercih edilebilir(Jokiel-Rokita and Pulit, 2013).

3.5.2. Çekirdek yoğunluk tahmin yöntemi (Kernel estimator)

Lloyd (1998) tarafından geliştirilen bu yöntem esas olarak ROC eğrisinin tahmininde ampirik tahmin yöntemlerinin esneklik eksikliğinin giderilmesi amacıyla kullanılır. Çekirdek yoğunluk tahmin yöntemi basit, çok yönlü, oldukça iyi teorik ve pratik özelliklere sahiptir (Silverman, 1986; Tenreiro, 2010).

Çekirdek tabanlı ROC eğri tahmininin çoğunlukla en hassas yönü, optimum bant genişliği seçimidir (Zhou ve Harezlak, 2002; Hall ve Hyndmann, 2003; Zhou ve diğerleri, 2011; Jokiel-Rokita ve Pulit, 2013). Bu, K seçimi ile birleştğinde tahminin özelliklerini belirler. Zou vd. (1997), f ve g 'yi tahmin etmek için asimptotik olarak optimal bant genişliklerini kullanmıştır. Lloyd (1998), ROC eğrisi doğrudan bu kümülatif dağılım fonksiyonlarına bağlı olduğundan, F ve G 'yi tahmin etmek için asimptotik olarak optimal olan bant genişlikleri seçerek önceki öneriyi geliştirmiştir. Lloyd ve Yong (1999), çekirdek yoğunluğu tahmin edicilerinin deneysel olanların üstesinden nasıl geldiğini gösterir. Qiu ve Le (2001), G 'ye bir çekirdek dağılım fonksiyonu tahminine ve $F - 1$ 'e yerel bir düzleştirme kuantil fonksiyonu tahminine dayanan bir ROC eğri tahmini önermiştir. Peng ve Zhou (2004), X ve Y desteğinin sınırına yakın daha iyi performansa sahip olan, optimal asimptotik bir şekilde tahmin edilen tek bir bant genişliğini içeren başka bir çekirdek tahminini tanıtmıştır. Kol'a'cek ve Karunamuni (2009), sınır etkilerini ortadan kaldıran ROC eğrisi için ilgili çekirdek tabanlı bir tahminci önermiştir. Bu yaklaşımların aksine, Jokiel-Rokita ve Pulit (2013), bir simülasyon çalışmasına göre küçük örneklem büyüklükleri için deneysel ve çekirdek tahmin ediciden daha iyi performans gösteren ampirik ROC eğrisinin düzleştirilmiş bir versiyonuna dayanan oldukça tutarlı bir tahminci önermiştir. AUC'yi tahmin etmek için çekirdek tabanlı tahmin ediciler de kullanılabilir. Birkaç parametrik ve parametrik olmayan yöntemi karşılaştıran bir simülasyon çalışması için bkz. Fabsic (2012).

3.6. Bayes Yaklaşımları

Bayes yöntemleri, maksimum olabilirlik yöntemlerine alternatif olarak ROC eğrisi tahmininde kullanılır. Bayes yaklaşımları, tahmin sürecine önceki bilgilerin girilmesini sağlayarak elde edilen çıkarımların belirsizliğini azaltır. Bu husus, özellikle sağlıklı - hastalıklı grupları ayırmak için ve hastalığa ait bir altın standart testinin olmadığı durumlarda değer kazanır. Zorlu süreçlerle zaman kaybetmemek, ayrıca bir teknik altyapı oluşturmamak ve maliyet kazanımlarıyla ön plana çıkar. Bu tür çalışmalarda bireylerin gerçek durumu bilinmemeden, önceki dağılımlarda kullanılan veriler yardımıyla daha doğru sonuca ulaşmaya çalışılır. Bayes modelleri, ROC eğrisi ve AUC gibi diğer özet ölçümler için güvenilirlik aralıkları elde edilmesini sağlar. Bayes çalışmalarını avantajlı hale getiren en önemli özellik parametrelerin arka dağılımlarından elde edilen bilgilerle güvenilirlik alanlarının belirlenmesidir. Güvenirlik alanları, Bayes yaklaşımlarıyla olasılıksal yorumlamayı da ekler. Böylece alanın güvenilirliği artmış kabul edilir. Buna ilave olarak gelecekteki belirsiz grupların sağlık durumlarına ilişkin öngörü olasılıkları da belirlenebilir. Ayrıca Bayes perspektifi ile hiyerarşik yapıların modele dahil edilmesi, karmaşık tasarımların modellenmesinde büyük kolaylıklar sağlar (Ishwaran and Gatsonis, 2000; O'Malley and Zou, 2006; Johnson and Johnson, 2006). Tanı testlerinin hastalıklı ve sağlıklı popülasyonları ayırt etme yeteneğinin çeşitli faktörlerden etkilenebildiği bilinen bir olgudur (Pepe, 2003). Bununla birlikte ortak değişkenlerinin etkisinin değerlendirilmesi, farklı popülasyonlara karşı test yeterliliği hakkında daha doğru ve faydalı bilgiler elde edilmesine yardımcı olabilir (de Carvalho et al., 2013). Aynı zamanda stabil olmayan değişkenlerin tahminlerde ihmal edilmesi test performansı hakkında önyargılı çıkarımlara neden olabileceğine de dikkat çekmek gerekir. Eş değişkenli etkilerin ROC eğrileri üzerindeki değişimleri birkaç çalışmada gösterilmiştir (Peng ve Hall, 1996; Branscum et al., 2008; Carvalho et al, 2013). Klasik olarak, Bayes metoduyla yapılan ROC eğrisi tahminleri özellikle parametrik verilerle gerçekleşmiş olsa da son zamanlarda yarı-parametrik ve parametrik olmayan olgularda da aynı yaklaşımlarla çalışmalar yapılmaya devam edilmektedir.

3.6.1. Parametrik Yaklaşım

Bayes yöntemi ile ROC eğrisi kestirimi ilk olarak regresyon modellemesi esasına dayanmaktadı. (Peng and Hall, 1996; Hellmich et al., 1998). ROC eğrileri ve AUROC hesaplamalarının bayes modellemeleri için sıralı test sonuçlarına Probit bağlantılı genellenmiş lineer regresyon formülleri uygulanır. Hasta ve sağlam popülasyon dağılımında Peng ve Hall (1996) tarafından tanımlanmış olan bi-normal dağılım anlayışına rağmen diğer parametrik dağılımlar da yaklaşımlar arasında yer almaktadır. Çalışmacılar, gizli dağılımlarda gözlemlenemeyen sürekli verilere imkan sağlayan veri yerleştirme teknikleri kullanarak gözlemlerden elde edilen ve doğası gereği sıralı olan bu verilerin işlenmesindeki zorluklardan kısmen kurtulmuş olurlar. Bu serilerde zaman zaman bilgi değeri olmayan öncelikler kullanılabilir.

Bu ordinal regresyon modeli, tek bir ortak değişkenin değerinin değiştirilmesinden kaynaklanan ROC eğrilerinde gözlemlenen değişiklikleri açıklayabilir. ROC eğrisi analizine bazı regresyon yaklaşımları hiyerarşik yapıları göz önünde bulundurur (O'Malley and Zou, 2006; Johnson and Johnson, 2006). Bayes'in çok değişkenli hiyerarşik dönüşüm modeli, O'Malley ve Zou (2006) tarafından, değişkenlerle kümelenmiş sürekli tanı testi verilerine dayanılarak geliştirilmiştir. Bu yaklaşım, örneğin farklı hastanelerde (kümelenme) buna odaklanmış hastalardan (bireysel seviye) toplanan radyolojik veriler gibi kümelenmiş yanıtları olan çok düzeyli veriler bağlamında yararlıdır. Bu yaklaşım araştırmacılara tanı testlerinin doğruluğunu ve iç içe geçmiş tanı testlerinin tanımlamalarına imkan tanır.

Araştırmacılar, kümeler arasındaki heterojenliği elde etmek için sonuçların bir kümeye özgü dönüşümünün uygulanabileceğini ve çoklu korelasyon sonuçların kullanılabileceğini belirtmektedir. Bu metodoloji, çok merkezli bir klinik çalışmadan toplanan prostat kanseri biyopsi verilerine uygulandı. Johnson ve Johnson (2006) radyolojide sıkça gözlemlenen ve birkaç radyologun sıralı bir ölçekle aynı kişiden alınan birden fazla veriyi derecelendirdiği bir durumu ele aldı. Çok oranlı korelasyonlu sıralı verilerin analizi için bir Bayes hiyerarşik gizli değişken modeli kullanmayı öngördüler. Üç farklı varyasyon kaynağı belirlendi: Hastaların karakteristik farkları, tanı için kullanılan muayene farklılıkları ve değerlendiricilerin farklılıkları olarak modellediler. Bu simülasyon çalışmasında, bu yöntemin, en çok bilinen ve yaygın olarak kullanılan çok

katmanlı korelasyonlu veri analizi modellerinden çok daha verimli olduğunu ortaya koydular (Dorfman et al., 1992).

3.6.2. Semiparametrik ; Nonparametrik Yaklaşımlar

Son yıllarda ROC eğrisi hesaplaması için Bayes'in yarı-parametrik ve parametrik olmayan modelleri kullanılmaktadır (Erkanli et al., 2006; Wang et al., 2007; Gu et al., 2008; Branscum et al., 2013). Bu metodolojiler hala geliştirilmekte olup üzerinde önemli çalışmalar yapılmaktadır.

Parametrik olmayan Bayes yöntemleri ile, sabit parametrik model ve bunun sonucu olarak multimodal olma ve çarpıklık gibi standart olmayan veri özelliklerinin yakalanmasındaki zorluklar göz önüne alınarak getirilen kısıtlamaların engel oluşturmalarını aşarak çözümleme imkanı oluşturmaları amaçlanmaktadır.

Geleneksel parametrik yapının aksine, parametrik olmayan değerlendirmede, belirli bir parametrik dağılım kümesinin dikkate alınmaması ile verilerin daha esnek bir modellenmesini sağlar. Parametrik olmayan yaklaşımda, tüm olasılık ölçüm alanı üzerine önceden bir dağılım belirtilmesini gerektiren bir modelleme yapısı ortaya koyar.

In'acio (2012) tarafından işaret edildiği gibi, bu modelde parametrelerin olmaması anlamına gelmez, aksine (muhtemelen) sonsuz sayıda parametre içerir. Bu çerçevede, Dirichlet prosesleri, Dirichlet proseslerinin karışımları, Polya ağaçları ve Polya ağaçlarının karışımları sıklıkla kullanılmaktadır. Dirichlet processinin karışımlarını temel alan Bayes yarı-parametrik ROC eğrisi tahmin yöntemi, Erkanlı ve ark. tarafından geliştirilmiştir (2006). Gibbs örnekleme iskeleti, Dirichlet prosesi modelinin karışımlarının arka plan dağılımını elde etmek için kullanılır, böylece hassasiyet, özgüllük, ROC eğrileri ve EAA'nın da arka plan tahminleri elde edilebilir. Araştırmacılar, altın standart bir tanı testi mevcut olmadığında bile sonuçların hala geçerli olduğunu göstermiştir. Ayrıca, daha önce bahsedilmiş olan çekirdek yoğunluğu tahmin yaklaşımı ile yakıdan ilişkilidir. Hanson ve arkadaşları (2008) sürekli serolojik verilerin analizinde Dirichlet process ve Polya dağılımından yararlanmışlardır. Bu yöntemde tanımlanmış olan yenilik, enfekte ve enfekte olmamış popülasyonların serolojik değer dağılımları için stokastik bir sıralama kısıtlaması içermesidir. Esas olarak enfekte olan popülasyon verilerinin olmayanlara göre yüksek olma eğiliminden dolayı oldukça makul bir varsayımdır Araştırmacılara göre bu yaklaşımın sağladığı avantaj EAA 'nın her zaman 0,5'ten büyük olmasını garanti etme imkanı sağlar ve ROC eğrisinin asla ana diagonalin

altına düşmediğini gösterir. İki modelde süt ineklerinde görülen Johne hastalığı verilerinde kullanılmıştır. Nitel olarak benzer çıkarımlar elde edilir ve serolojik testlerin doğruluğuna ilişkin aynı sonuçlar her iki uygulamadan da çıkarılabilir. Non- parametrik Bayes yöntemi kullanılarak yapılan ROC eğrisi kestiriminde, ortak değişkenlerin etkisini inceleyen az sayıda çalışma vardır.. Bu konu ilk olarak Carvalho ve arkadaşları tarafından açıklanmıştır(2013). Model, bağımlı Dirichlet işlemlerine dayanmaktadır ve her gruptaki tüm dağılımın ortak değişkenlerin bir fonksiyonu olarak sorunsuz bir şekilde değişmesine izin verir. Bu yaklaşımda birden fazla ve sürekli kategorik kestirim imkanı sunar. Bu çalışmaların devamı niteliğinde olan bazı karşılaştırma çalışmaları yapılmıştır. Pepe (1998) ve Rodr'íguez-Alvarez' ve arkadaşlarının yaptıkları çalışmalar karşılaştırılmıştır. Bu çalışmalar temelinde B-Spline eğrileri , yarı-parametrik yaklaşım ile non-parametrik çekirdek yoğunluk tahmin yöntemleri ile yapılmış olan çalışmalardır. Sonuç olarak önerilen modelin doğrusal olmayan senaryolar ve küçük örnek boyutları için kullanılan yöntemlerin performans olarak daha iyi olduğunu göstermiştir. Bu modelin diyabet tanısında kullanılması önerilmiştir.

3.6.3. Altın Kural Eksiği

Bilindiği gibi tıpta kullanılan testler kusursuz değildir. Bayes metodolojisi özellikle bu testler için kısmen uygundur (Krzanowski and Hand, 2009). Kusursuz olmayan bu testlerin referans sonuçlarının nonparametrik uzantı modellemeleri, gerçek fakat bilinmeyen hastalık durumunu ifade etmek için yeterli veriyi sunabilir. Altın standardın bulunmadığı durumlarda Maksimum olabilirlik veya Bayes kestirim yöntemlerini kullanarak, ikili tanı testlerinin performansının değerlendirmesi hakkında çok sayıda literatür bulunmaktadır. Aynı şekilde Altın standart veri analizinin olmadığı durumlarda da benzer şekilde ikili tanı testlerinde bayes metodolojisinin geçerli olduğu ortaya konmuştur (Choi ve ark. (2006)). Veriler bi-normal varsayıma göre sınıflandırılabilir Ancak bu varsayım altın standart bir testin yokluğunda uygun bir transformasyon yapılabiliyorsa geçerli olabilir. Simülasyon verileri yada daha önce yapılmış olan altın standart olarak kabul edilebilecek veriler bu dönüşüm için yeterli düzeyde referans olabilir. Esas olarak bu yöntem başlangıçta altın standart bir durum için formüle edilmiş olmasına rağmen altın standart yokluğunda uygulanmak üzere hafifçe değiştirilmiştir. Bu metod, hastalıklı ve sağlıklı popülasyon gruplarının örtüşmesi beklenenin üzerinde büyük olduğunda, gerçek hastaların durumuna karar vermekte zorluk

çıkarabilir. Wang ve ark. (2007) , iki test arasında koşullu bağımsızlık varsayarak, kusurlu bir ikili referans testiyle karşılaştırılarak yeni bir sıralı veya sürekli ölçekli tanı testinin ROC eğrisini tahmin etme problemini araştırmıştır. Tanımlanabilirlik sorunları, farklı prevalanslara sahip en az iki popülasyondan veri gerektirir. Bu yöntem, multinomial bir model olup, test değerlerine karşılık gelen dağılımların şekli ile ilgili herhangi bir varsayıma ihtiyaç duymaz. Bu durumda ROC eğrisinin tekdüzeliğinin korunması veya sağlanması avantajları elde edilmiş olur. Both Choi ve ark. (2006) ile Wang ve ark. (2007) sığırlarda görülen Johne hastalığı ile ilgili olarak farklı veri kümeleri kullanmıştır. Tıpta ise bu modellerden, akciğer lifi sayısı verilerinden asbest maruz kalma olasılığını belirlemek için yararlanılmıştır. Model, bireylerde aynı testin tekrarlanan ölçümleri arasındaki korelasyonları kabul eder. Branscum ve ark. (2008), bir altın standardın yokluğunda ROC analizi ve hastalık teşhisine Bayes parametrik olmayan ve yarı parametrik yaklaşımlar önermektedir. Hastalık riski ve ROC eğrisinin olasılıklarını tahmin etmek için Polya dağılımlarının karışımlarını kullanan parametrik olmayan bir model önerilmektedir. Bu modelin semiparametric uzantılarında önerilen çözümler arasındadır. Bu semiparametrik modeller hastalık durumu ile ilgili ek bilgiler içermektedir. Bu modellerde iki tür veri kullanılır. Bunlar: standart değişken veriler ve ek ikili tanı testlerinden elde edilen verilerdir. Bu tür ek bilgiler, denekleri sağlıklı veya hastalıklı olarak doğru bir şekilde sınıflandırma özelliğini artırır. Bu da altın standart olgular ile altın standart bulunmadığında non-parametrik modelleme arasında bir modelleme sürecine yol açar. Bu, diyagnostik tahmin doğruluğunu arttırmak için risk faktörleri ve ek tanı testlerinin sonuçları hakkında mevcut tek bir çerçevede birleştirmeyi sağlayan çok esnek bir yaklaşımdır.

4. ROC EĞRİSİ ALTI ALAN HESAPLAMA YÖNTEMLERİNİN KARŞILAŞTIRILMASI:

Bu çalışmada ROC eğrisi altı alan hesaplaması için, Erişkinlerde Bipolar bozuklukta dürtüsellik “Barratt Dürtüsellik Ölçeği_(BIS; Barratt Impulsiveness Scale)”, “Dikkat Eksikliği Hiperaktivite Bozukluğu Kendi Bildirim Ölçeği_(ASRS; Adult ADHD Self-Report Scale)” ve Kişilerin kendi tutumları, görüş ve ilgilerini tanımlarken kuallandığı “Mizaç ve Karakter Envanteri_(TCI; Temperament and Character Inventory:)” formlarının toplam skorları örnek veri setimizi oluşturmaktadır. Çözümlemede parametrik “Binormal” hesaplama yöntemi ile bulunann AUC değeri ile “Binomial exact_De Long” yaklaşımı ve nonparametric (Mann-Whitney) çözümleme yöntemlerinin sonucu olan AUC değerlerinin birbiriyle karşılaştırılması amaçlanmıştır. Bu hesaplama yöntemlerinin veri setine hangisinin uygun olabileceğinin belirlenmesi için normallik denetlemesi ve diğer denetleme yaklaşımları uygulanmıştır.

4.1. Normal Dağılıma Uygunluk Denetlemesi

4.1.1. Merkezi ve dağılım ölçütlerine göre karar verilmesi.

- Ortalama, median ve mod birbirine yakın olmalıdır.
- Standart Sapma Ortalamanın yaklaşık $\frac{1}{4}$ ü olmalıdır.

4.1.2. Grafikler; histogram, normal olasılık grafiği, Skewness ve kurtosis ile Normale uygunluk değerlendirmesi.

- ‘Histogramı’ veya "normal olasılık grafiği" (normal probability plot) ile,
- Çarpıklık ve Basıklık değerleri kullanılabilir. Skewness ve Kurtosis sıfıra yakın olmalıdır.

4.1.3. Hipotez Testleri ile Denetleme

Kolmogorov Smirnov, Shapiro-Wilks’, vb. test kullanılabilir. $p < 0.05$ ise, dağılım normal değildir. Denek sayısı yüksek ise normal dağılım testleri " $p < 0.05$ " verebilir, bu durumlarda grafik değerlendirmeler dikkate alınmalıdır.

5. GEREÇ VE YÖNTEM

5.1. Klinik Veriler:

Erişkinlerde Bipolar bozuklukta “Barratt Dürtüsellik Ölçeği (BIS; Barratt Impulsiveness Scale)”, “Dikkat Eksikliği Hiperaktivite Bozukluğu (DEHB) Kendi Bildirim Ölçeği (ASRS; Adult ADHD Self-Report Scale)” ve Kişilerin kendi tutumları, görüş ve ilgilerini tanımlarken kuullandığı “Mizaç ve Karakter Envanteri (TCI; Temperament and Character Inventory)” formları ile toplanan, hasta (n=116) ve kontrol (n=69) olmak üzere toplam 185 bireyin, gruplar arası karşılaştırma ve sosyo-demografik özellikler ile ölçek sonuçları arasındaki ilişkilerinin araştırılması çalışmasının verileridir.

	KONTROL GRUBU			HASTA GRUBU		
	N	Mean	SD	N	Mean	SD
ASRS_Toplam	78	12,19	7,33	121	22,76	7,83
Barrat_Toplam	66	47,59	8,44	108	83,85	20,72
MKE_Toplam	36	108,42	6,37	87	118,67	14,21

Tablo 11. Sağlam ve Hasta gruplarının tanımlayıcı değerleri.

SAĞLAM	Tanımlayıcı Değerler			Most Extreme Differences			Kolmogorov-Smirnov	
	N	Mean	SD	Absolute	Positive	Negative	z	p
ASRS_Toplam	69	11,87	7,08	0,12	0,12	-0,06	1	0,271
Barrat_Toplam	66	47,56	8,48	0,12	0,12	-0,07	0,94	0,34
MKE_Toplam	35	108,29	6,41	0,09	0,09	-0,07	0,55	0,918

Tablo 12. Sağlam Grup için “One-Sample Kolmogorov-Smirnov Testi” kullanılarak yapılan normallik denetleme sonuçları.

HASTA	Tanımlayıcı Değerler			Most Extreme Differences			Kolmogorov-Smirnov	
	N	Mean	SD	Absolute	Positive	Negative	z	p
ASRS_Toplam	116	23,01	7,72	0,1	0,1	-0,08	1,11	0,174
Barrat_Toplam	107	84,22	20,45	0,09	0,06	-0,09	0,95	0,329
MKE_Toplam	87	118,67	14,21	0,08	0,05	-0,08	0,72	0,681

Tablo 13. Hasta Grup için "One sample Kolgomorov-Smirnov Testi" kullanılarak yapılan normallik denetleme sonuçları.

Sağlam ve hasta olgularda ayrı ayrı yapılan denetleme sonucunda tüm değişken ve altgruplarda dağılımın normal ($p>0.05$) olduğu gözlenmiştir.

5.2. Yöntem:

ASRS, BARRAT VE MKE ölçeklerinin tanı performanslarının belirlenmesi amacıyla yapılan ROC Curve analizi ve eğri altı alan (AUC) hesaplama yöntemlerinin bulguları farklı yazılımlarla (SPSS, Stata, Medcalc) karşılaştırılmıştır.

Çalışmanın biyoistatistiksel çözümlemesinde, ele alınan ölçütler Ortalama, Standart Sapma, frekans ve yüzde değerleri ile tanımlanmıştır.

Gruplar arası frekans ve yüzdelerin kıyaslanmasında **Ki-kare ve Fisher kesin olasılık testi** kullanılmıştır.

Normal dağılıma sahip değişken ortalamalarının karşılaştırılması için,

-iki grup (hasta-kontrol) kıyaslamasında **t, z testi**, ve bazı denetlemeler sonucunda gerek duyulan (denek sayısı ve homojenlik sorunları nedeniyle) deney kurgularında nonparametrik “**Mann-Whitney U**” kullanılmıştır.

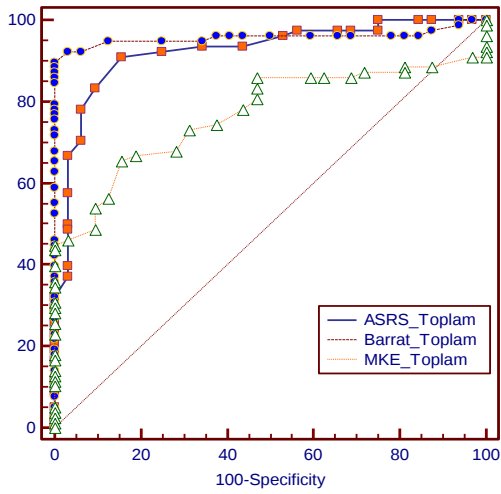
Yorumlamalarda anlamlılık sınırı **p=0.05** alınmıştır.

Biyoistatistiksel analizlerde **SPSS** (Sürüm:21.0) paket programı kullanılmıştır

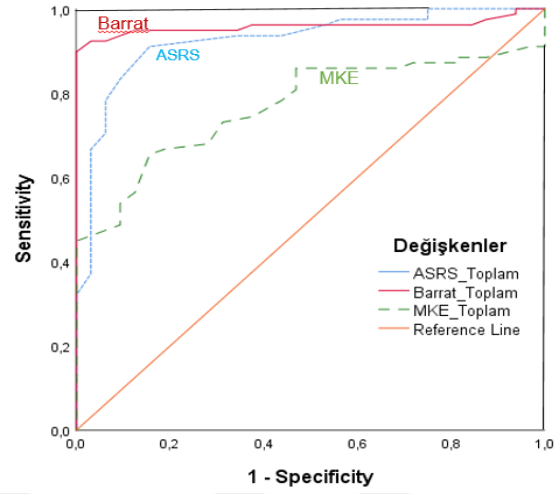
6. BULGULAR

ROC eğrisi analizi yalnızca duyarlılık ve özgüllük değerlerini kullanarak tanı belirlemenin sakıncalarından kurtulmak için geliştirilmiş bir değerlendirme yöntemidir (Metz, 2006 ve 1978).

Binormal ROC eğrisi grafiği



Nonparametrik ROC eğrisi grafiği



Şekil 25. Binormal ROC ve Nonparametrik ROC eğrisi Grafikleri

Değişkenler	AUC	SE _{AUC} ^a	p ^b	95% Güven Sınırları	
				Alt sınır	Üst sınır
ASRS_Toplam	,925	,028	,000	,870	,980
Barrat_Toplam	,957	,021	,000	,915	,998
MKE_Toplam	,766	,045	,000	,679	,853

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

Tablo 14. Araştırma değişkenlerinin ROC analizi bulguları (Nonparametrik) .

ASRS, Barrat ve MKE değişkenlerinin tümünün tanı performansı nonparametric çözümlemeye göre referans alandan anlamlı olarak farklı bulunmuştur (p=0.000). Bu sonuç üç ölçütünde başarılı olduğunun kanıtıdır.

6.1. ROC eğrilerini Karşılaştırılması

Çalışma verileri ile yapılan ROC analizi sonucunda ASRS, BARRAT ve MKE ölçeklerinin toplam skorlarının tanı ayırımında oldukça başarılı (AUC değerleri sırasıyla 0.925, 0.957 ve 0.766) sonuçlar verdikleri gözlenmiştir (tüm değişkenler için $p=0.000$).

ASRS_Toplam ~ Barrat_Toplam	
Difference between areas	0,0333
Standard Error ^c	0,0298
95% Confidence Interval	-0,0251 to 0,0916
z statistic	1,116
Significance level	P = 0,2643

Tablo 15. ASRS-Toplam ve Barrat-Toplam ölçeklerinin değerlendirme tablosu

ASRS ve BARRAT toplam skorlarının AUC değerleri sırasıyla 0.925, 0.957 bulunmuş olup ikiside yüksek doğrulukla tanı koyabilmektedirler ve aralarındaki fark anlamlı değildir ($p=0.2643$). Ancak aşağıdaki iki tablonun sonucuna göre ASRS ve BARRAT ölçeklerinin tanı belirleme performansının MKE skor toplamalarında anlamlı olarak daha yüksek oldukları sonucu elde edilmiştir (sırasıyla $p=0.0055$ ve $p=0.0001$).

ASRS_Toplam ~ MKE_Toplam	
Difference between areas	0,156
Standard Error ^c	0,0561
95% Confidence Interval	0,0457 to 0,266
z statistic	2,775
Significance level	P = 0,0055

Tablo 16. ASRS-Toplam ve MKE-Toplam ölçeklerinin değerlendirme tablosu

Barrat_Toplam ~ MKE_Toplam	
Difference between areas	0,189
Standard Error ^c	0,0490
95% Confidence Interval	0,0929 to 0,285
z statistic	3,858
Significance level	P = 0,0001

Tablo 17. BARRAT-Toplam ve MKE-Toplam ölçeklerinin değerlendirme tablosu

MKE toplam skor değişkeninin gerçekte kullanılabilir düzeyde (AUC=0.766) bir ölçüt olmakla birlikte ASRS ve BARRAT ölçeklerinden anlamlı olarak daha az başarılı olduğu sonucuna varılmıştır.

6.2. ROC EĞRISI ALTI ALAN HESAPLAMA YÖNTEMLERİNİN KARŞILAŞTIRILMASI

Çalışma verileri kullanılarak parametrik “Binormal” hesaplama yöntemi ile bulunann AUC değeri ile “Binomial exact_De Long” yaklaşımı ve nonparametric (Mann-Whitney) çözümleme yöntemi ile hesaplanan AUC değerlerinin birbiriyle karşılaştırılmış, hesaplamalar adım adım olmak üzere sonuç ve bulgular aşağıdaki tablolarda sunulmuştur.

ASRS (n=185)

	AUC	SE _{AUC}
1-Binormal	0.9074	0,0296
2-Binomial exact_De Long	0.8910	0,0259
3-Nonparametric	0.9249	0,028

Tablo 18. ASRS, AUC değerleri karşılaştırma tablosu

ASRS, Fark ve Karşılaştırma Sonuçları

	Fark	SD	Z	P
1-2	0,0164	0,0393	0,417	0,677
1-3	0,0175	0,0407	0,429	0,668
2-3	0,0339	0,0381	0,889	0,374

Tablo 19. ASRS, AUC değerleri arasındaki farklar tablosu

Yukarıdaki iki tablonun ilkinde, ASRS değişkeninin hasta ve sağlam ayırımındaki başarı düzeyinin araştırıldığı ROC eğrisi çözümlemesinin sonuçlarına göre parametrik ve nonparametrik temelli 3 farklı yöntemin AUC ve SE değerleri, ikinci tabloda ise AUC değerleri arasındaki fark, fark dizilerinin standart sapması ve AUC’lerin birbiriyle karşılaştırma Sonuçlarına ait z ve p değerleri bulunmaktadır. P değerleri “Binormal” “Binomial exact_De Long” ve nonparametric (Mann-Whitney) yöntemlerinin verdiği

sonuçlar (AUC) arasında 1-2, 1-3 ve 2-3 kıyaslamalarında herhangi anlamlı bir farka rastlanmadığını (üç kıyaslamada da $p>0.05$) göstermektedir.

BARRAT (n=173)

	AUC	SE _{AUC}
1-Binormal	0.9007	0.0282
2-Binomial exact_De Long	0.9410	0.0181
3-Nonparametric	0.9566	0,0210

Tablo 20. BARRAT, AUC değerleri karşılaştırma tablosu

BARRAT, Fark ve Karşılaştırma Sonuçları

	Fark	SD	Z	P
1-2	0,0403	0,0335	1,203	0,229
1-3	0,0559	0,0352	1,590	0,112
2-3	0,0156	0,0277	0,563	0,574

Tablo 21. BARRAT, AUC değerleri arasındaki farklar tablosu

Yukarıdaki iki tabloda da olanında, BARRAT değişkeninin hasta ve sağlam tanı performansı için ROC eğrisi analizi bulgularının birbirleri ile karşılaştırma sonuçlarına ait değerlendirmede “Binormal” “Binomial exact_De Long” ve nonparametric (Mann-Whitney) yöntemlerinin AUC’leri arasında 1-2, 1-3 ve 2-3 kıyaslamalarında anlamlı fark olmadığı (üç kıyaslamada da $p>0.05$) saptanmıştır.

MKE (n=122)

	AUC	SE _{AUC}
1-Binormal	0.8263	0,1905
2-Binomial exact_De Long	0.7670	0,045
3-Nonparametric	0.7670	0,045

Tablo 22. MKE, AUC değerleri karşılaştırma tablosu

MKE, Fark ve Karşılaştırma Sonuçları

	Fark	SD	Z	P
1-2	0,0593	0,1960	0,303	0,762
1-3	0,0593	0,1960	0,303	0,762
2-3	0,0009	0,6360	0,014	0,989

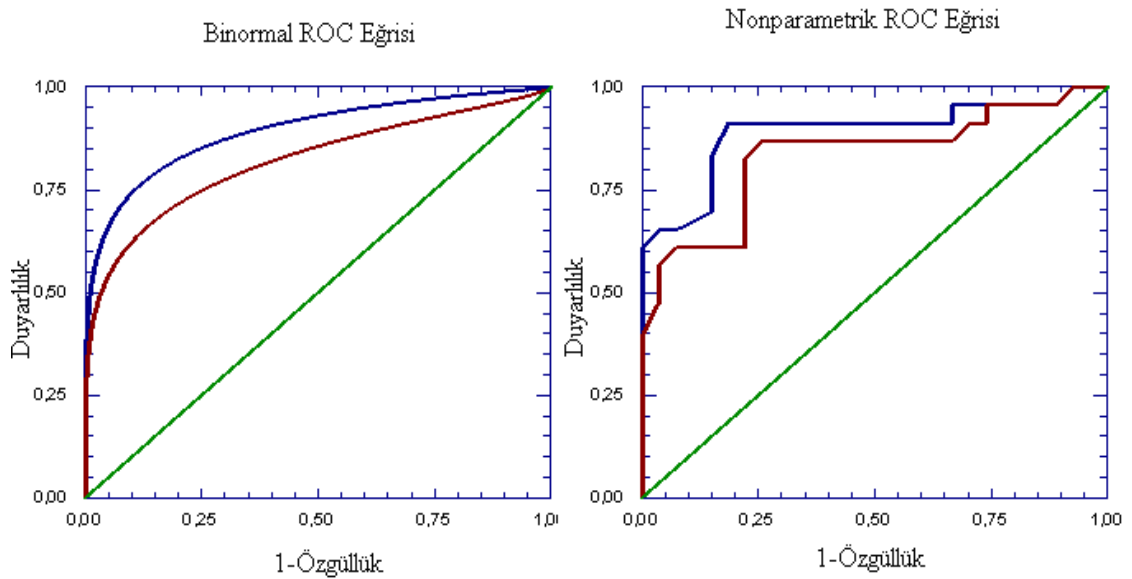
Tablo 23. MKE, AUC değerleri arasındaki farklar tablosu

Yukarıdaki MKE değişkenine ait iki tabloda hasta ve sağlam tanı ayırımı için ROC eğrisi çözümlemesi ile “Binormal” “Binomial exact_De Long” ve nonparametric (Mann-Whitney) yöntemlerinin birbirleri arasında (1-2, 1-3 ve 2-3 kıyaslamalarında) gözlenen farkların biyoistatistiksel olarak anlamlı olmadığı (üç kıyaslamada da $p>0.05$) sonucuna varılmıştır.

7. TARTIŞMA

ROC eğrisi hesaplama prosedürü ve sunduğu görselin izlenmesi açık ve kolaydır. Mümkün olan tüm sınır değerlere (cut-off point) ilişkin testin tanısal yeterlilik düzeyini ve doğruluğunu göstermektedir. Tek bir tanısal karar eşiğinin seçilmesini gerektirmez ve prevalanstan bağımsızdır. Testler arasında direkt görsel kıyaslamaya olanak verebilecek özelliktedir. Histogram grafiklerinde olduğu gibi, sonuçların gruplandırılarak değerlendirilmesini gerektirmez. Duyarlılık ve özgüllük değerleri grafikten doğrudan elde edilebilir (Taga ve ark., 2000).

Biyoistatistik hesaplamalarda parametrik ve parametrik olmayan yöntemleri içeren çok sayıda istatistik paket program kullanılmaktadır. Bazı paket programlarda iki yöntemden yalnızca birisi kullanıma sunulmaktadır. Örneğin SPSS programında ROC eğrisi altındaki alan hesaplaması için yer alan yöntem parametrik olmayan Mann Whitney yaklaşımıdır. Parametrik (Binormal) yöntem ile hesaplama seçeneği bulunmamaktadır.



Şekil 26. Binormal ROC ve Non parametrik ROC eğrisi grafikleri

Normal dağılımlı parametrik verilere daha uygun olduğu kabul edilen ve yukarıdaki örnekte görüldüğü gibi hesaplamada kullanılan binormal yöntem ile elde edilen grafikte gözlemlendiği gibi eğriler düzleştiği için ROC eğrisi altındaki alan anlamlı olmasa da belli bir düzeyde farklılaşmaktadır. Nonparametrik yöntemde ise eğri düzleştirilmeden gerçek alan hesaplandığı için ve koşul olacak varsayımlar da gerektirmediği için genelde nonparametrik yöntem, binormal yöntemle tercih edilmektedir.

Çalışmamızdaki üç yöntemden elde edilen AUC düzeyleri karşılaştırıldığında anlamlı bir farklılığın olmadığı sonucuna varılmıştır. Dolayısıyla eldeki veri özelliklerine göre hangi yöntemin uygulanması kullanım kolaylığı sağlıyorsa o yöntem tercih edilebilecektir.

Hasta ve sağlıklılara ilişkin tanı testi sonuçlarının her ikisi de normal dağılımlıyken büyük örneklem genişliğine sahip olanlarda parametrik yöntem daha yansız olma eğilimindedir, ancak parametrik olmayan yöntemin de çok rahatlıkla kullanılabileceği görülmektedir.

Sonuç olarak; Çalışmamızdaki üç yöntemden elde edilen ROC çözümlemesi sonucu AUC düzeyleri karşılaştırıldığında anlamlı bir farklılık olmadığı sonucuna varılmıştır. Dolayısıyla eldeki veri özelliklerine göre hangi yöntemin uygun olduğuna karar verilirse veya özgün uygulama kolaylığı sağladığına inanılan çözümleme yöntemin tercih edilebileceği kanısına varılmıştır.

KAYNAKÇA

1. Wilks, D. S. (1995) *Statistical Methods in the Atmospheric Sciences*. Academic Press, NY. 467 pp.
2. Wilks, D. S. (2001) A skill score based on economic value for probability forecasts. *Meteorol. Appl.*, 8, 209-219.
3. Zhou, X-H., D. K. McClish, and N. A. Obuchowski (2002) *Statistical Methods in Diagnostic Medicine*. John Wiley & Sons, pp. 464.
4. Zou, K. H. (2001) <http://splweb.bwh.harvard.edu:8000/pages/pp1/zou/roc.htm>
5. Bamber, P. (1975). The area above the ordinal dominance graph and the area below the receiver operating characteristic graph, *Journal of Mathematical Psychology*, 12, 387-415.
6. Branscum, A. J.; Johnson, W. O. and Baron, A.T. (2013). Robust medical test evaluation using flexible bayesian semiparametric regression models, *Epidemiology Research International*, ID 131232.
7. Branscum, A. J.; Johnson, W. O.; Hanson, T. E. and Gardner, I. A. (2008). Bayesian semiparametric ROC curve estimation and disease diagnosis, *Statistics in Medicine*, 27, 2474–2496.
8. Choi, Y.-K.; Johnson, W. O.; Collins, M. T. and Gardner, I. (2006). Bayesian inferences for receiver operating characteristic curves in the absence of a gold standard, *Journal of Agricultural, Biological, and Environmental Statistics*, 11, 210–229.
9. Collison, P. (1998). Of bombers, radiologists, and cardiologists: time to ROC, *Heart*, 80, 215–217.
10. ROC Curve Estimation: An Overview
11. de Carvalho, V. I.; Jara, A.; Hanson, T. E. and de Carvalho, M. (2013). Bayesian nonparametric ROC regression modeling, *Bayesian Analysis*, 8, 623–646.
12. Devlin, S. A.; Thomas, E. G. and Emerson, S. S. (2013). Robustness of approaches to ROC curve modeling under misspecification of the underlying probability model, *Communications in Statistics—Theory and Methods*, 42, 3655–3664.
13. Dorfman, D. D.; Berbaum, K. S. and Metz, C. E. (1992). Receive operating characteristic rating analysis: generalization to the population of readers and patients with the jackknife method, *Investigative Radiology*, 27, 723–731.
14. Dorfman, D. D.; Berbaum, K. S.; Metz, C. E.; Lenth, R. V.; Hanley, J. A. and Dagga, H. A. (1996). Proper receiver operating characteristic analysis: the bigamma model, *Academic Radiology*, 4, 138–149.
15. Erkanli, A.; Sung, M.; Costello, J. E. and Angold, A. (2006). Bayesian semi-parametric ROC analysis, *Statistics in Medicine*, 25, 3905–3928.
16. Fabsic, P. (2012). Comparing the Accuracy of ROC Curve Estimation Methods, Master Thesis, Swiss Federal Institute of Technology, Zurich.
17. Farcomeni, A. and Ventura, L. (2010). An overview of robust methods in medical research, *Statistical Methods in Medical Research*, 21, 111–133.
18. Faraggi, D. and Reiser, B. (2002). Estimation of the area under the ROC curve, *Statistics in Medicine*, 21, 3093–3096.
19. Gonçalves, L.; Subtil, A.; Oliveira, M. R.; Rosario, V.; Lee, P. ´ and Shaio, M.-F. (2012). Bayesian latent class models in malaria diagnosis, *PLoS ONE*, 7, e40630.
20. Greco, L. and Ventura, L. (2011). Robust inference for the stress-strength reliability, *Statistical Papers*, 52, 773–788.

21. Gu, J.; Ghosal, S. and Roy, A. (2008). Bayesian bootstrap estimation of ROC curve, *Statistics in Medicine*, 27, 5407–5420.
22. Hajian-Tilaki, K. O.; Hanley, J. A.; Joseph, L. and Collet, J. P. (1997). A comparison of parametric and nonparametric approaches to ROC analysis of quantitative diagnostic tests, *Medical Decision Making*, 17, 94–102.
23. Hall, P. G. and Hyndmann, R. J. (2003). Improved methods for bandwidth selection when estimating ROC curves, *Statistics and Probability Letters*, 64, 181–189.
24. Hand, D. J. (2009). Measuring classifier performance: a coherent alternative to the area under the ROC curve, *Machine Learning*, 77, 103–123.
25. Hanley, J. A. (1988). The robustness of the “binormal” assumptions used in fitting ROC curves, *Medical Decision Making*, 8, 197–203.
26. Hanson, T. E.; Kottas, A. and Branscum, A. J. (2008). Modelling stochastic order in the analysis of Receiver Operating Characteristic data: Bayesian non-parametric approaches, *Journal of the Royal Statistical Society. Ser. C*, 57, 207–225.
27. L. Gonçalves, A. Subtil, M. Rosário Oliveira and P. de Zea Bermudez Hellmich, M.; Abrams, K. R., Jones, D. R. and Lambert, P. C. (1998). A Bayesian approach to a general regression model for ROC curves, *Medical Decision Making*, 18, 436–443.
28. Heritier, S.; Cantoni, E.; Copt, S. and Victoria-Feser, M.-P. (2009). *Robust Methods in Biostatistics*, John Wiley & Sons, Chichester.
29. Hong, X.; Chen, S. and Harris, C. J. (2007). A kernel-based two-class classifier for imbalanced data sets, *IEEE Transactions on Neural Networks*, 18, 28–41.
30. Hsieh, F. and Turnbull, B.W. (1996). Nonparametric and semiparametric estimation of the receiver operating characteristic curve, *The Annals of Statistics*, 24, 25–40.
31. Huang, Y. and Pepe, M. S. (2009). A parametric ROC model based approach for evaluating the predictiveness of continuous markers in case-control studies, *Biometrics*, 65, 1133–1144.
32. Hughes, G. and Bhattacharya, B. (2013). Symmetry properties of binormal and bi-gamma receiver operating characteristic curves are described by Kullback–Leibler divergences, *Entropy*, 15, 1342–1356.
33. Hussain, E. (2012). The bi-gamma ROC curve in a straightforward manner, *Journal of Basic & Applied Sciences*, 8, 309–314.
34. Inacio, V. (2012). *Semiparametric and Nonparametric Modeling of Diagnostic Data*, PhD Thesis, University of Lisbon, Portugal.
35. Johnson, T. D. and Johnson, V. E. (2006). A Bayesian hierarchical approach to multirater correlated ROC analysis, *Statistics in Medicine*, 25, 1858–1871.
36. Jokiel-Rokita, A. and Pulit, M. (2013). Nonparametric estimation of the ROC curve based on smoothed empirical distribution functions, *Statistics and Computing*, 23, 703–712.
37. Ishwaran, H. and Gatsonis, A. C. (2000). A general class of hierarchical ordinal regression models with applications to correlated ROC analysis, *The Canadian Journal of Statistics*, 28, 731–750.
38. Kolaček, J. and Karunamuni, R. J. (2009). On boundary correction in kernel estimation of ROC curves, *Austrian Journal of Statistics*, 38, 17–32.
39. Krzanowski, W. J. and Hand, D. J. (2009). *ROC Curves for Continuous Data*, Chapman & Hall/CRC, Boca Raton.
40. Lasko, T. A.; Bhagwat, J. G.; Zou, K. H. and Ohno-Machado, L. (2005). The use of receiver operating characteristic curves in biomedical informatics, *Journal of Biomedical Informatics*, 38, 404–415.

41. Lloyd, C. J. (1998). Using smoothed receiver operating characteristic curves to summarize and compare diagnostic systems, *Journal of the American Statistical Association*, 93, 1356–1364.
42. Lloyd, C. J. and Yong, Z. (1999). Kernel estimators of the ROC curve are better than empirical, *Statistics and Probability Letters*, 44, 221–228.
43. Lusted, L. B. (1960). Logical analysis in roentgen diagnosis, *Radiology*, 74, 178–193.
44. ROC Curve Estimation: An Overview 19
45. Martinez, E. Z.; Louzada-Neto, F. and Pereira, B. B. (2003). A curva ROC para testes diagnósticos (Document in Portuguese), *Cadernos Saúde Coletiva*, 11, 7–31.
46. Marzaban, C. (2004). The ROC curve and the area under it as performance measures, *Weather and Forecasting*, 19, 1106–1114.
47. Morrison, A. M.; Coughlin, K.; Shine, J. P.; Coull, B. A. and Rex A. C. (2003). Receiver operating characteristic curve analysis of beach water quality indicator variables, *Applied and Environmental Microbiology*, 69, 6405–6411.
48. O'Malley, A. J. and Zou, K. H. (2006). Bayesian multivariate hierarchical transformation models for ROC analysis, *Statistics in Medicine*, 25, 459–479.
49. Parodi, S.; Pistoia, V. and Muselli, M. (2008). Not proper ROC curves as new tool for the analysis of differentially expressed genes in microarray experiments, *BMC Bioinformatics*, 8, 410.
50. Peng, F. and Hall, W. J. (1996). Bayesian analysis of ROC curves using Markov-chain Monte Carlo methods, *Medical Decision Making*, 16, 404–411.
51. Peng, L. and Zhou, X.-H. (2004). Local linear smoothing of receiver operating characteristic (ROC) curves, *Journal of Statistical Planning and Inference*, 118, 129–143.
52. Pepe, M. S. (1998). Three approaches to regression analysis of receiver operating characteristic curves for continuous test results, *Biometrics*, 54, 124–135.
53. Pepe, M. S. (2003). *The Statistical Evaluation of Medical Tests for Classification and Prediction*, Oxford University Press, New York.
54. Pundir, S. and Amala, R. (2012). A study on the bi-Rayleigh ROC curve model, *Bonfring International Journal of Data Mining*, 2, 42–47.
55. Qiu, P. and Le, C. (2001). ROC curve estimation based on local smoothing, *Journal of Statistical Computation and Simulation*, 70, 55–69.
56. Rodríguez-Alvarez, M. X.; Roca-Pardinas, J. and Cadarso-Suarez, C. (2011). ROC curve and covariates: extending induced methodology to the non-parametric framework, *Statistics and Computing*, 21, 483–499.
57. Sheather, S. J. (2004). Density estimation, *Statistical Science*, 19, 588–597.
58. Silva-Fortes, C.; Amaral Turkman, M. A. and Sousa L. (2012). Arrow plot: a new graphical tool for selecting up and down regulated genes and genes differentially expressed on sample subgroups, *BMC Bioinformatics*, 13, 147.
59. Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman & Hall, London.
60. Swets, J. A. (1986). Form of empirical ROCs in discrimination and diagnostic tasks: implications for theory and measurement of performance, *Psychological Bulletin*, 99, 181–198.
61. Swets, J. A. (1996). *Signal Detection Theory and ROC Analysis in Psychology and Diagnostics: Collected Papers*, Lawrence Erlbaum Associates, New Jersey.
62. Tang, L.; Du, P. and Wu, C. (2010). Compare diagnostic tests using transformation-invariant smoothed ROC curves, *Journal of Statistical Planning and Inference*, 140, 3540–3551.

63. L. Gonçalves, A. Subtil, M. Rosário Oliveira and P. de Zea Bermudez, Tang, L. L. and Balakrishnan, N. (2011). A random-sum Wilcoxon statistic and its application to analysis of ROC and LROC data, *Journal of Statistical Planning and Inference*, 141, 335–344.
64. Tenreiro, C. (2010). Uma introdução à estimação não-paramétrica da densidade (Document in Portuguese), Edições SPE, Lisboa.
65. Walsh, S. J. (1997). Limitations to the robustness of binormal ROC curves: effects of model misspecification and location of decision thresholds on bias, precision, size and power, *Statistics in Medicine*, 16, 669–679.
66. Wang, C.; Turnbull, B.W.; Grohn, Y. T. and Nielsen, S. S. (2007). Nonparametric estimation of ROC curves based on Bayesian models when the true disease state is unknown, *Journal of Agricultural, Biological, and Environmental Statistics*, 12, 128–146.
67. Weichenthal, S.; Joseph, L.; Bédise, P. and Dufresne, A. (2010). Bayesian estimation of the probability of asbestos exposure from lung fiber counts, *Biometrics*, 66, 603–612.
68. Zhou, X.-H. and Harezlak, J. (2002). Comparison of bandwidth selection methods for kernel smoothing of ROC curves, *Statistics in Medicine*, 21, 2045–2055.
69. Zhou, X-H.; Obuchowski, N. A. and McClish, D. K. (2011). *Statistical Methods in Diagnostic Medicine*, Second Edition, John Wiley & Sons, New Jersey.
70. Zou, K. H.; Hall, W. J. and Shapiro, D. E. (1997). Smooth non-parametric receiver operating characteristic (ROC) curves for continuous diagnostic tests, *Statistics in Medicine*, 16, 2143–2156.
71. Drton, M., C. Marzban, P. Guttorp, and J. T. Schaefer (2003) A Markov Chain Model of Tornado Activity. *Mon. Wea. Rev.*, 131, 2941-2953.
72. Hanley J. A., and B. J. McNeil (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 143, 2936.
73. Hanley J. A. and B. J. McNeil (1983) A method of comparing the areas under Receiver Operating Characteristic curves derived from the same cases. *Radiology*, 148, 839-43.
74. Harvey, L. O. Jr., K. R. Hammond, C. M. Lusk, and E. F. Mross (1992). Application of signal detection theory to weather forecasting behavior. *Mon. Wea. Rev.*, 120, 863-883.
75. Marzban, C., S. Leyton, B. Colman (2003) Nonlinear Post-processing of Model Output: Ceiling and Visibility. NWS/COMET report.
<http://www.nhn.ou.edu/~marzban/cometreport1.pdf>
<http://www.nhn.ou.edu/~marzban/comettalk.pdf>
76. Marzban, C., and A. Witt (2001) A Bayesian Neural Network for Hail Size Prediction. *Wea. Forecasting*, 16, 600-610.
77. Mason, I. (1982) A model for assessment of weather forecasts. *Aust. Meteor. Mag.*, 30, 291-303.
78. Mason, S. J., and N. E. Graham (1999) Conditional Probabilities, Relative Operating Characteristics, and Relative Operating Levels. *Wea. Forecasting*, 14, 713-725.
79. Metz, C. E., B. A. Herman, J. H., Shen (1998) Maximum likelihood estimation of receiver operating characteristic (ROC) curves from continuously-distributed data. *Statistics in Medicine*, 17, 1033-1053.
80. Miettinen, K. (1999) *Nonlinear Multiobjective Optimization*, Kluwer Academic, Boston, 300 pp.
81. Murphy, A. H., and R. L. Winkler (1992) Diagnostic verification of probability forecasts. *Int. J. Forecasting*, 7, 435-455.

82. Murphy, A. H., and R. L. Winkler (1987) A general framework for forecast verification. *Mon. Wea. Rev.*, 115, 1330-1338.
83. Mylne, K. R. (1999) The use of forecast value calculations for optimal decision making using probability forecasts. Preprints, 17th Conf. on Weather Analysis and Forecasting, Denver, CO, Amer. Meteor. Soc., 235239.
84. Richardson (2000) Applications of costloss models. Proc. Seventh ECMWF Workshop on Meteorological Operational Systems, Reading, United Kingdom, ECMWF, 209213.
85. Richardson (2001) Measures of skill and value of ensemble prediction systems, their interrelationship and the effect of ensemble size. *Quart. J. Roy. Meteor. Soc.*, 127, 24732489.
86. Shapiro, D. E. (1999) The interpretation of diagnostic tests. *Statistical Methods in Medical Research*, 8, 113-134.
87. Stephenson, D. B. (2000) Use of the "Odds Ratio" for Diagnosing Forecast Skill. *Wea. Forecasting*, 15, 221-232.



ÖZGEÇMİŞ

Kişisel Bilgiler

Adı	Metin	Soyadı	Vural
Doğ.Yeri	Ankara	Doğ.Tar.	24.09.1966
Uyruğu	TC	TC Kim No	16411498688
Email	drvuralmetin@gmail.com	Tel	05322420779

Eğitim Düzeyi

Mez. Yılı

Yük.Lisans	Boğaziçi Üniversitesi Biyomedikal Mühendisliği	2005
Doktora	İ.Ü. Cerrahpaşa Tıp Fakültesi	1989
Lise	Antalya Lisesi	1983

İş Deneyimi (Sondan geçmişe doğru sıralayın)

Görevi	Kurum	Süre (Yıl - Yıl)
1. 112 Acil Sağlık Hizmetleri	Sağlık Bakanlığı	2020-
2. Centrium Hastanesi	Özel	2018-2019
3. Maya Estetik	Özel	2016-2018
4. Aile Hekimliği	Sağlık Bakanlığı	2010-2016
5. Spectrum Teknoloji	Özel	2008-2010
6. Anatolia Göz Hastanesi	Özel	2006-2008
7. Medikal Otomasyon Sist.	Özel	1998-2007
8. Sağlık Eğitim Fakültesi	Marmara Üni.-Öğretim Görevlisi	1998-2007
9. İstanbul Sağlık Müdürlüğü	Sağlık Bakanlığı	1991-1997
10. Haseki Hastanesi	Sağlık Bakanlığı	1990-1991

Yabancı Dilleri	Okuduğunu Anlama*	Konuşma*	Yazma*	KPDS/ÜDS/ Yökdil Puanı	(Diğer) Puanı
İngilizce	İyi	Orta	İyi		
Almanca	Orta	Orta	Orta		

*Çok iyi, iyi, orta, zayıf olarak değerlendirin

Bilgisayar Bilgisi

Program	Kullanma becerisi
MS Office, işletim sistemi ve donanım	Çok İyi
SPSS ve istatistik paket programları	Orta, Zayıf





