

T.C.
HACETTEPE ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

ÇOKLU TEDAVİ ETKİNLİĞİNİN
KARŞILAŞTIRILMASINDA
PROPENSITY SKOR AĞIRLIKLANDIRMASI

Osman DEMİR

Biyoistatistik Programı
DOKTORA TEZİ

ANKARA

2015

T.C.
HACETTEPE ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

ÇOKLU TEDAVİ ETKİNLİĞİNİN
KARŞILAŞTIRILMASINDA
PROPENSITY SKOR AĞIRLIKLANDIRMASI

Osman DEMİR

Biyoistatistik Programı
DOKTORA TEZİ

TEZ DANIŞMANI
Prof. Dr. Osman SARAÇBAŞI

ANKARA
2015

Anabilim Dalı: Biyoistatistik

Program: Biyoistatistik

Tez Başlığı : Çoklu Tedavi Etkinliğinin Karşılaştırılmasında Propensity Skor
Ağırlıklandırması

Öğrenci Adı-Soyadı: Osman DEMİR

Savunma Sınavı Tarihi: 06/10/2015

Bu çalışma jürimiz tarafından doktora tezi olarak kabul edilmiştir.

Jüri Başkanı: Prof. Dr. Reha ALPAR
Hacettepe Üniversitesi

Danışman: Prof. Dr. Osman SARAÇBAŞI
Hacettepe Üniversitesi

Üye: Doç. Dr. Erdem KARABULUT
Hacettepe Üniversitesi

Üye: Doç. Dr. Serdal Kenan KÖSE
Ankara Üniversitesi

Üye: Doç. Dr. Mehtap AKÇİL OK
Başkent Üniversitesi

ONAY:

Bu tez, Hacettepe Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliği'nin ilgili maddeleri uyarınca yukarıdaki jüri üyeleri tarafından uygun görülmüş ve Enstitü Yönetim Kurulu kararıyla kabul edilmiştir.


Prof. Dr. Ersin FADILIOĞLU
Enstitü Müdürü.

TEŞEKKÜR

Çalışmamda benden desteklerini esirgemeyen, kıymetli danışman hocam Sayın Prof. Dr. Osman Saraçbaşı'na teşekkürlerimi sunarım. Doktora programına başladığımdan bu zamana beni teşvik eden ve yardımlarını esirgemeyen Yrd. Doç. Dr. Anıl Dolgun hocama ve ailesine teşekkürlerimi sunarım. Tez izleme komitemde bulunan Doç. Dr. Erdem Karabulut ve Doç. Dr. Serdal Kenan Köse hocalarıma katkılarından dolayı teşekkür ederim. Yine yardımlarını esirgemeyen araştırma görevlileri Gökmen Zararsız ve Dinçer Göksülük'e teşekkür ederim. Üzerimde emeği olan büyüğüm Doç. Dr. İlker Etikan'a ve iki hayat ağacım, eşim Özlem'e, oğlum afacan Batu'ya teşekkür ederim.

ÖZET

Demir, O. Çoklu Tedavi Etkinliğinin Karşılaştırılmasında Propensity Skor Ağırlıklandırması. Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Programı Doktora Tezi, Ankara, 2015. Rasgeleleştirmenin sağlanamadığı yarı deneysel ya da gözlemsel çalışmalarda bireysel farklılıklardan dolayı tedavi kollarına atamada dengesizlik oluşmaktadır. Bu dengesizlik, tedavi etkinliğinin belirlenmesinde yanlılığa neden olmaktadır. Bu çalışma, dengesizliği ortadan kaldırmak için propensity skorlarının kullanılmasını önermektedir. Tedavi kollarındaki bireylerin gruplara atanma olasılıkları olan propensity skorları üzerinden ağırlıklandırma yaparak bu dengesizliğin giderilmesi sağlanmaktadır. İki'den fazla tedavi kolu olduğu durumda propensity skorlarını kestirmek için genelleştirilmiş boosted model ve çok kategorili lojistik regresyon yöntemleri kullanılmaktadır. Bu yöntemlerin denge üzerindeki performansları Monte Carlo benzetimi yardımıyla karşılaştırılmaktadır. Ayrıca yöntemlerde nedensel etkiler kullanılarak tedavilerin yanıt üzerindeki etkisi incelenmektedir. Yöntemlere göre dengenin karşılaştırılması ve tedavilerin yanıt üzerindeki etkisinin incelenmesi için yapılan benzetim çalışmasının ilk kısmındaki denge değerlerine ilişkin bulgulara, genelleştirilmiş boosted model çok kategorili lojistik regresyona göre daha başarılıdır. Yanıt üzerindeki etki için elde edilen bulgulara ise senaryolara göre her iki yöntemde de tedavilerin etkilerinin farklılık gösterdiği görülmektedir. Çalışmanın uygulama kısmında vitamin B12 üzerinde etkisi olduğu düşünülen tedavi değişkeni olarak alınan yaşanan yer değişkeni ve 8 ortak değişkenle oluşturulan ağırlıksız modelden elde edilen yaş, MCV, ve FOLAT değişkenlerine ilişkin önemlilik ağırlıklandırma sonrasında da bozulmamaktadır. Sonuç olarak; tedavi kollarına rasgele atamanın yapılamadığı durumda, bağımlı değişken üzerinde etki araştırıldığında çeşitli senaryolar altında yöntemlerden elde edilen sonuçlar hakkında bilgi verilmektedir.

Anahtar Kelimeler: Genelleştirilmiş boosted model, çok kategorili lojistik regresyon, propensity skoru, propensity skor ağırlıklandırması, ortalama tedavi etkisi (OTE), tedavi edilenler arasında ortalama tedavi etkisi (TOTE)

ABSTRACT

Demir, O. Propensity Score Weighting For The Comparison of Multiple Treatment Effects. Hacettepe University, Institute of Sciences, Ph.D. Thesis, Ankara, 2015. In non-randomized quasi-experimental or observational studies, unbalance may occur in treatment arm assignments due to individual differences. This unbalance leads to bias in determining treatment efficiency. This study suggests using propensity scores to eliminate unbalance. This unbalance's elimination is ensured by weighting on propensity scores which are the assignment probabilities of individuals in treatment arms. When there are more than two treatment arms, generalized boosted model and multinomial logistic regression are used to estimate the propensity scores. Performances of these methods on unbalance are compared via a Monte Carlo simulation. Also, treatment effects on outcome are investigated in these methods using causal effects. In the first part of simulation study, which is conducted for the comparison of unbalance according to methods, generalized boosted model is more successful than multinomial logistic regression. In the second part of the simulation study where the treatment effects on outcome are compared, it is found that the treatment effects of two methods are different from each other. In the application part of the study, significant effect of age, MCV and FOLAT variables on vitamin B12 obtained from unweighted model which is formed with 8 covariates and living place is taken as treatment variable have changed after weighting. As a results; in situation that random assignment can not be done and the treatment effect on the dependent variable is investigated, under various scenarios, information about results from methods are provided.

Key Words: Generalized boosted model, multinomial logistic regression, propensity score, propensity score weighting, average treatment effect (ATE), average treatment effect on the treated (ATT).

İÇİNDEKİLER

	Sayfa
ONAY SAYFASI.....	iii
TEŞEKKÜR.....	iv
ÖZET.....	v
ABSTRACT.....	vi
İÇİNDEKİLER	vii
SİMGELER VE KISALTMALAR.....	ix
ŞEKİLLER.....	xi
TABLolar	xiii
1 GİRİŞ.....	1
2 GENEL BİLGİLER.....	5
2.1 Nedensel Çıkarım.....	5
2.2 Tedavi Etkileri	6
2.2.1 Ortalama Tedavi Etkisi (OTE).....	7
2.2.2 Tedavi Edilenler Arasında Ortalama Tedavi Etkisi	7
3 PROPENSITY SKORU.....	9
3.1 Propensity Skorunda Gösterimler	10
3.2 Propensity Skor Yöntemleri.....	10
3.2.1 Propensity Skoru ile Tabakalama.....	11
3.2.2 Propensity Skoru ile Eşleştirme	11
3.2.3 Propensity Skoru ile Ortak Değişken Düzeltmesi.....	11
3.2.4 Propensity Skoru ile Ağırlıklandırma	12
3.2.5 Varsayımlar	14
3.3 Propensity Skor Kestirim Yöntemleri.....	15
3.3.1 Genelleştirilmiş Boosted Model (GBM)	15
3.3.2 Çok Kategorili Lojistik Regresyon	16
4 GEREÇ VE YÖNTEM.....	18
4.1 R’de Kullanılan Paketler.....	18
4.2 Benzetim Çalışması	18
4.3 Benzetim Algoritmaları.....	18
4.3.1 Ortak Değişken ve Tedavi Değişkeni Türetimi.....	20
4.3.2 Yanıt Değişken Türetimi.....	21

4.3.3	Veri Setlerinin Türetimi	21
4.4	Performans Ölçüleri	21
4.4.1	Standartlaştırılmış yanlılık	21
4.4.2	Kolmogorov Smirnov İstatistiği.....	22
4.5	Gerçek Veri Seti Çalışması.....	23
5	BULGULAR.....	26
5.1	Denge Değerlendirmesi Sonuçları	27
5.1.1	n=500 için Senaryoların Denge Değerlendirmesi	27
5.1.2	n=1000 için Senaryoların Denge Değerlendirmesi	41
5.1.3	n=2000 için Senaryoların Denge Değerlendirmesi	55
5.2	Yanıt Değişken Üzerine Olan Etkilerin Değerlendirilmesi	69
5.2.1	n=500 için Etkilerin Değerlendirilmesi.....	69
5.2.2	n=1000 için Etkilerin Değerlendirilmesi.....	71
5.2.3	n=2000 için Etkilerin Değerlendirilmesi.....	73
5.3	Gerçek Veri Seti Uygulaması	75
5.3.1	Uygulama Verisi için MLR ve GBM ile Denge Değerlendirmesi.....	76
5.3.2	Örtüşme (Overlap) Varsayımının MLR ve GBM için Görsel Değerlendirmesi.....	80
6	TARTIŞMA.....	84
7	SONUÇ VE ÖNERİLER.....	87
	KAYNAKLAR	89
	EKLER	
	EK 1: Etik Kurul İzni	
	EK 2: Veri Türetim Senaryoları	
	EK 3: Veri Türetiminde Kullanılan Katsayılar	

SİMGELER VE KISALTMALAR

ESS_{m_k}	k grubuna ilişkin örnek büyüklüğü
FOLAT	Folik asit
GBM	Genelleştirilmiş Boosted Model
HCT	Hematokrit
HGB	Hemoglobin
KS	Kolmogorov Smirnov istatistiği
Mak.es	En büyük standartlaştırılmış yanlılık (etki büyüklüğü)
Mak.ks	En büyük Kolmogorov Smirnov istatistiği
MCHC	Ortalama Alyuvar Hemoglobin Derişimi
MCV	Ortalama Alyuvar Çapı
Mean.es	Ortalama standartlaştırılmış yanlılık
Mean.ks	Ortalama Kolmogorov Smirnov istatistiği
MLR	Çok kategorili lojistik regresyon
MSB_{cm_k}	c ortak değişkeni için k tedavisini alanlarda mutlak standartlaştırılmış yanlılık
OTE	Ortalama tedavi etkisi
$\hat{\mu}_{m_k}$	k grubuna ilişkin kestirilen ortalama
$p_{m_k}(X_j)$	Ortak değişken vektörlerine göre k grubuna atanma olasılığı
TOTE	Tedavi edilenler arasında ortalama tedavi etkisi
$w_j[m_k]$	k grubuna ilişkin ağırlıklar
$Y[m_k]$	k grubundaki yanıt

$YF_{cm_k}(x)$ c ortak değişkeni için k grubundaki yoğunluk fonksiyonu

ŞEKİLLER

	Sayfa
3.1. Tedavi etkinlikleri kestiriminde akış şeması	17
4.1. Benzetim verisi için değişken yapısı	19
4.2. Uygulama veri seti yapısı	24
5.1. Yöntemlere göre denge değerleri (S1, n=500)	28
5.2. Yöntemlere göre denge değerleri (S2, n=500)	30
5.3. Yöntemlere göre denge değerleri (S3, n=500)	32
5.4. Yöntemlere göre denge değerleri (S4, n=500)	34
5.5. Yöntemlere göre denge değerleri (S5, n=500)	36
5.6. Yöntemlere göre denge değerleri (S6, n=500)	38
5.7. Yöntemlere göre denge değerleri (S7, n=500)	40
5.8. Yöntemlere göre denge değerleri (S1, n=1000)	42
5.9. Yöntemlere göre denge değerleri (S2, n=1000)	44
5.10. Yöntemlere göre denge değerleri (S3, n=1000)	46
5.11. Yöntemlere göre denge değerleri (S4, n=1000)	48
5.12. Yöntemlere göre denge değerleri (S5, n=1000)	50
5.13. Yöntemlere göre denge değerleri (S6, n=1000)	52
5.14. Yöntemlere göre denge değerleri (S7, n=1000)	54
5.15. Yöntemlere göre denge değerleri (S1, n=2000)	56
5.16. Yöntemlere göre denge değerleri (S2, n=2000)	58
5.17. Yöntemlere göre denge değerleri (S3, n=2000)	60
5.18. Yöntemlere göre denge değerleri (S4, n=2000)	62

5.19.	Yöntemlere göre denge değerleri (S5, n=2000)	64
5.20.	Yöntemlere göre denge değerleri (S6, n=2000)	66
5.21.	Yöntemlere göre denge değerleri (S7, n=2000)	68
5.22.	n=500 için tedavi etkileri	70
5.23.	n=1000 için tedavi etkileri	72
5.24.	n=2000 için tedavi etkileri	74
5.25.	1'e karşı 2. tedavi grubunun OTE değerlendirmesi	77
5.26.	1'e karşı 3. tedavi grubunun OTE değerlendirmesi	77
5.27.	1'e karşı 2. tedavi grubunun TOTE değerlendirmesi	78
5.28.	1'e karşı 3. tedavi grubunun TOTE değerlendirmesi	79
5.29.	Tedavi gruplarına göre GBM ve MLR ile kestirilen propensity skorlarına ilişkin kutu çizgi grafiği	80
5.30.	Ağırlıklandırma sonrası sabite ilişkin kestirimler (Tokat merkez referans)	82
5.31.	Ağırlıklandırma öncesi ve sonrası etkiler	83

TABLOLAR

	Sayfa
4.1. Bireylere ilişkin deęişkenlerin tanımlayıcı istatistikleri	25
4.2. Vitamin B12 üzerinde etkisi olduęu düşünölen deęişkenlere ilişkin doğrusal regresyon modeli sonuçları	25
5.1. MLR ve GBM'ye ilişkin özet denge deęerleri (S1, n=500)	27
5.2. MLR ve GBM'ye ilişkin özet denge deęerleri (S2, n=500)	29
5.3. MLR ve GBM'ye ilişkin özet denge deęerleri (S3, n=500)	31
5.4. MLR ve GBM'ye ilişkin özet denge deęerleri (S4, n=500)	33
5.5. MLR ve GBM'ye ilişkin özet denge deęerleri (S5, n=500)	35
5.6. MLR ve GBM'ye ilişkin özet denge deęerleri (S6, n=500)	37
5.7. MLR ve GBM'ye ilişkin özet denge deęerleri (S7, n=500)	39
5.8. MLR ve GBM'ye ilişkin özet denge deęerleri (S1, n=1000)	41
5.9. MLR ve GBM'ye ilişkin özet denge deęerleri (S2, n=1000)	43
5.10. MLR ve GBM'ye ilişkin özet denge deęerleri (S3, n=1000)	45
5.11. MLR ve GBM'ye ilişkin özet denge deęerleri (S4, n=1000)	47
5.12. MLR ve GBM'ye ilişkin özet denge deęerleri (S5, n=1000)	49
5.13. MLR ve GBM'ye ilişkin özet denge deęerleri (S6, n=1000)	51
5.14. MLR ve GBM'ye ilişkin özet denge deęerleri (S7, n=1000)	53
5.15. MLR ve GBM'ye ilişkin özet denge deęerleri (S1, n=2000)	55
5.16. MLR ve GBM'ye ilişkin özet denge deęerleri (S2, n=2000)	57
5.17. MLR ve GBM'ye ilişkin özet denge deęerleri (S3, n=2000)	59
5.18. MLR ve GBM'ye ilişkin özet denge deęerleri (S4, n=2000)	61

5.19.	MLR ve GBM'ye ilişkin özet denge değerleri (S5, n=2000)	63
5.20.	MLR ve GBM'ye ilişkin özet denge değerleri (S6, n=2000)	65
5.21.	MLR ve GBM'ye ilişkin özet denge değerleri (S7, n=2000)	67
5.22.	Senaryolara ilişkin tedavi etkileri (n=500)	69
5.23.	Senaryolara ilişkin tedavi etkileri (n=1000)	71
5.24.	Senaryolara ilişkin tedavi etkileri (n=2000)	73
5.25.	MLR ve GBM'ye ilişkin özet denge değerleri (n=5448)	76
5.26.	Ağırlıklandırmadan önce ve sonra tüm değişkenlerin yanıt üzerinde tedavinin etkisi	81

1 GİRİŞ

Bilimsel çalışmanın yapıldığı her disiplinde, toplanan verilerin uygun herhangi bir istatistiksel yöntemle değerlendirilebiliyor olması, sonuçların uygulanabilirliğine ve dolayısıyla bilgi edinmeye katkı sağlamaktadır. Bilimsel sonuçların geçerliğinin olması için ise verinin toplanma şekli büyük önem arz etmektedir. Dolayısıyla çalışma düzeninin ve yürütülme biçiminin sonuçların geçerliği üzerinde etkisi bulunmaktadır.

Tedavinin, yanıt değişken üzerindeki etkisi belirlenirken, tedavi kollarına atanan bireylerin özelliklerinden kaynaklı bir yanlılık ortaya çıkabilmektedir. Bu yanlılığın en aza indirilmesi için kullanılabilecek çalışma düzeni olan deneysel çalışmalarda amaç gruptaki bireylerin benzer özelliklerde olmasını sağlamaktır. Dolayısıyla yanıt değişken üzerindeki tedavi etkinliği yansız bir şekilde değerlendirilebilmektedir (1). Gruplara rasgele atamanın sağlanabildiği ve birbirine benzer grupların oluşturulmaya çalışılarak tedavilerin etkilerinin değerlendirildiği çalışmalar rasgele kontrollü çalışmalar olarak adlandırılır. Burada bireylerin gruplara atamasında araştırmacı ya da hasta faktörünü en aza indirmek gerekmektedir. Özellik olarak birbirine benzer gruplar oluşturulduğundan müdahalenin ya da tedavinin etkinliğinin yansız olarak değerlendirilmesi mümkün olabilmektedir. Bu bağlamda, rasgele kontrollü çalışmalar bilimsel olarak sonuçlarına güvenilirliğin en fazla olduğu ve klinik denemeler için altın standardın olduğu denemelerdir (2,3).

Gözlemsel çalışmalarda tedavi kollarına atanmış olan bireylerin özellikleri dikkate alınmadığından gruplarda bir dengesizlik söz konusu olmaktadır. Dolayısıyla gruplar arasındaki fark, yanlış bir gruba veya tedaviye atfedilebilir. Bu yanlılık beraberinde sistematik hataya neden olacaktır. Dolayısıyla araştırmacının oluşturacağı gruplardaki bireylere ilişkin özelliklerin benzer olması sonuçlarda yanlılığı en aza indirecektir.

Gözlemsel çalışmada ise araştırmaya katılan bireylere herhangi bir müdahale söz konusu değildir. Tanımlayıcı, geriye dönük, kesitsel ve ileriye dönük olmak üzere alt başlıklar halinde incelenebilecek olan gözlemsel çalışmaların hepsinde de olayın gelişimine dışardan bir girişimde bulunulmaz (4). Gözlemsel çalışmalarda

olaylar kontrol altına alınamadığından ve aynı koşulları sağlamak mümkün olmadığından nedensellik tam olarak açıklanamaz. Örneğin hastalık ve etken arasındaki ilişkinin kontrol altına alınamayan farklı bir faktörden mi kaynaklanıp kaynaklanmadığı bilgisine ulaşılamaz. Deneysel çalışmalarda ise bunu belirlemek daha kolaydır (3).

Tıp alanında çoğu araştırma gözlemsel olarak yürütülmektedir. Bununla birlikte maliyet ve etik nedenlerden dolayı deneysel çalışmanın yapılamadığı durumlarla da karşılaşmaktadır. Dolayısıyla rasgeleleştirme yapılmadan tasarlanan müdahaleli bir çalışma olan yarı deneysel ya da müdahalenin yapılmadığı gözlemsel çalışmalardan yararlanılmaktadır. Yarı deneysel ve gözlemsel çalışmaların kullanımında tedavi kollarına atamadaki dengesizlik, yanlılık ortaya çıkmaktadır. Bunun yanında verinin değerlendirilme sürecinde kullanılacak olan istatistiksel yöntemlere ilişkin varsayımlar ve kısıtlılıklar araştırmacıları, bu varsayımlar ve kısıtlılıkları hafifletecek ya da ortadan kaldıracak alternatif yöntemlere itmektedir. Ayrıca günümüz teknolojisindeki özellikle donanımlı bilgisayarların ve yazılımların bu alternatif yöntemlerin elde edilmesinde büyük öneme sahip olduğu açıktır.

Yarı deneysel ya da gözlemsel çalışmalardaki gibi rasgeleleştirmenin sağlanamadığı durumlarda tedavi kollarına atamadaki dengesizliği gidermek ve dolayısıyla tedavilerin yanıt üzerindeki etkisini incelemek için alternatif bir yöntem olan propensity skor teknikleri kullanılmaktadır. Bu tekniklerin genel amacı; tedaviye atanmadaki dengesizlikleri kontrol altına alarak sistematik hatayı azaltacak şekilde propensity skorunun kullanımını sağlamaktadır. Bu özelliği ile gözlemsel çalışma, rasgeleleştirmenin sağlandığı deneysel çalışmaya benzerlik gösterecektir (5). Propensity skoru, tedavi kollarında benzer olacak şekilde eşleştirme, tabakalama, ortak değişken düzeltilmesi ve ağırlıklandırma için kullanılmaktadır (6,7).

Gürtekin (2011), “*Propensity Skor Yönteminin Meme Kanserinde Yaş ile Dokudaki ER, PR, CEB-B2 Değişiminin Analizinde Kullanılması*” adlı doktora çalışmasında sistematik hatayı azaltmak için lojistik regresyon analizi sonrası elde ettiği propensity skorunu kullanarak tabakalama yöntemini uygulamaktadır. Propensity skorları kullanılarak elde edilen yeni örneklemlemlerle analizler tekrarlanarak yaş ile pozitif ve negatif şeklinde ikili olarak kodlanan östrojen (ER),

progesteron (PR) reseptörleri ve büyüme geni (Cerb-B2) arasında istatistiksel bir anlamlılığın olmadığı bulunmaktadır (8) .

Kandemir (2008), “Gözleme dayalı çalışmalarda propensity skor ve sosyal bilimlerde bir uygulama” adlı doktora çalışmasında ISO 9001:200 KYS belgesi olan ve olmayan şeklinde ikili olarak kodlanan bağımlı değişkenin, lojistik regresyon analizinden elde edilen propensity skorları yardımıyla eşleştirme yöntemi sonrasında KOBİ çalışanları arasında ilgili alt boyutlar açısından anlamlı farklılıklar olmadığını göstermişlerdir (9).

Altunok (2006), “Gözleme dayalı çalışmalarda propensity skor ve tıp bilimlerinde bir uygulama” adlı çalışmasında delirium tanısı alan ve almayan hastalar için lojistik regresyon analiziyle propenstiy skor tabakalamayı kullanarak risk faktörlerinin yeni oluşan örnekleme analizini gerçekleştirmektedir. Uygulanan teknik sonrasında risk faktörlerinin anlamlılıklarında değişiklik gözlenmiştir (10).

McCaffrey vd’nin 2013 yılında gerçek veri seti ile yaptıkları çalışmada, çoklu tedavi için genelleştirilmiş boosted ve çok kategorili lojistik regresyon kullanarak propensity skor ağırlıklandırma ile tedavi etkinlikleri değerlendirilmiştir (7).

Lee vd’nin 2009 çalışmasında yaptıkları benzetim çalışmasında, tedavi değişkenin iki kategorili olduğu durumda lojistik regresyon ve karar ağaçları yöntemleri karşılaştırılmaktadır (6).

Alan yazında yapılan çalışmalarda çok kategorili tedavi değişkeni için ağırlıklandırma yöntemini kullanarak Genelleştirilmiş boosted model (GBM) ve çok kategorili lojistik regresyonun (MLR) denge üzerindeki performanslarını ve tedavi etkinliklerini değerlendiren bir çalışma bulunmamaktadır.

Bu tez çalışmasının giriş bölümünde çalışmanın amacı, kapsamı, hipotezleri, önceki çalışmalar ile deneysel ve gözlemsel çalışmalar hakkında özet bilgiler verilmektedir. İkinci bölümde nedensel çıkarım ve tedavi etkileri üzerine genel bilgileri içermektedir. Üçüncü bölümde propensity skorunun ne olduğu ve propensity skor metotlarından bahsedilmektedir. Ayrıca yöntemlerin uygulamasında kullanılan gerçek veri tanıtılmıştır. Dördüncü bölümde R yazılımında kullanılan paketlerden ve benzetim çalışmasının yapısından denge değerlerini karşılaştırmak için kullanılan

metriklerden ve gerçek veri setine ilişkin çalışmadan bahsedilmektedir. Beşinci bölümde senaryolara göre GBM ve MLR'den elde edilen denge değerlerindeki değişim ve yanıt üzerindeki etkiler incelenmektedir. Son olarak gerçek bir veri seti üzerinde propensity skor ağırlıklandırma ile GBM ve MLR'den elde edilen sonuçlar değerlendirilmektedir.

Bu tez çalışmasında bireylerin tedavi kollarına atanmasındaki dengesizliği gidermek için propensity skor kullanımı önerilmektedir. Bireylerin tedavi gruplarına atanma olasılıkları olan propensity skorlarıyla ağırlıklandırarak oluşturulan dengelenmiş yapay gruplar üzerinden tedavi etkinliğini belirlemek bu tez çalışmasının amacını oluşturmaktadır. İki'den fazla tedavi kolu olduğu durumda çok kategorili lojistik regresyon ve genelleştirilmiş boosted model, propensity skorlarını elde etmek için kullanılacaktır. GBM ve MLR'den elde edilen propensity skorlarıyla ağırlıklandırma sonrası yöntemlerin denge performansları karşılaştırılacaktır. Buna göre GBM'nin genel olarak denge performansları açısından MLR'ye göre daha başarılı sonuçlar göstereceği ve farklı senaryolar altında yöntemlerin tedavi etkinliğini farklı belirleyeceği hipotez edilmektedir. Bu hipotezleri test edebilmek için farklı veri senaryoları altında veri türetilecek, GBM ve MLR ile propensity skorları elde edilerek ağırlıklandırma yapılacak, ağırlıklandırma sonrası denge değerleri ve tedavi etkinlikleri karşılaştırılacaktır.

2 GENEL BİLGİLER

2.1 Nedensel Çıkarım

İki ya da daha fazla değişkenin (olayın) birbirini etkilemesi sonucu değişkenler (olaylar) arası ilişkiden ve dolayısıyla etkileyen ve etkilenen değişkenler olarak adlandırılan kavramlar ortaya çıkmaktadır. Bu kavramlar deneysel bilim çerçevesinde olayların ya da bireylerin gözlenebilir özelliklerinin tanımlanmasını, olaylar arasındaki ilişkilerin tanımlanmasını ve daha önemlisi olaylar arasındaki nedensel ilişkilerin tanımlanmasını gerektirmektedir (4,11). Değişkenler arasındaki gözlenebilir bağlantı olarak tanımlanabilen ilişki için göz önünde bulundurulması gereken nokta, bu ilişkiyle birlikte nedenselliğin var olmasıdır. Bir değişkenin başka bir değişken ya da değişkenler tarafından etkileniyor olması, değişkenler arasındaki nedenselliğin bir göstergesi olmayabilir. Dolayısıyla mutlak bir ilişkiden çok, bu ilişkinin ilgili alan içerisinde mantıksal yoruma sahip bir neden-sonuç ilişkisi içermesi gereklidir (12). Örneğin planlanan bir çalışma çerçevesinde cebinde çakmak bulunan kişilerin akciğer kanseri olması ile sonuçlanan bir etkiye sahip olduğu gibi ilişki bulunabilir. Burada düşünülmesi gereken bilgi, çakmanın varlığı ile akciğer kanseri olması arasında nedensel olmayan ilişkiden çok çakmak taşıyan kişilerin bir sigara içicisi olduğu ve bu durumun akciğer kanseri riskini artırdığı neden-sonuç ilişkisidir. Bu örnekten de görüldüğü gibi sonuca etki eden bir neden olması ve nedenin sonuçtan önce olması gerekmektedir (13).

Nedenselliğin olduğu yerde, sonuç üzerindeki nedensel etkilerin kestirilmesinde karşıolgusallık (counterfactual) kavramı ortaya çıkmaktadır. Karşıolgusallık çalışma gruplarındaki bireylerin verilen her bir tedavi için potansiyel bir yanıtı sahip olacağı ideal durumunu içermektedir. Örneğin tedavi almayan (örneğin sigara içen) bir birey için gözlenebilir yanıt değişkeni (bebeğin doğum ağırlığı) Y_0 olduğu düşünölsün. Şayet aynı birey tedaviyi almış (sigara içmeyen) olsaydı yanıt değişkeninin Y_1 olduğu düşünölsün, Y_1 yanıt değişkeni bu birey için karşıolgusaldır denir. Buradaki problem bireyin asıl tedaviyi aldığı durumda gözlenebilir bir potansiyel yanıtı sahip olması ve aksi durumda kayıp veri içermesidir. Buna; Rubin'in nedensel çıkarımının temel sorunu (Rubin's fundamental problem of causal inference) denmektedir (14,15). Bu problemin çözümü için

potansiyel yanıtlar üzerinden kestirimde bulunacak olan tedavi etkileri kullanılmaktadır (16).

2.2 Tedavi Etkileri

Tedavi etkisi, ilgilenilen bağımlı değişken ya da yanıt değişken (örneğin, sistolik kan basıncı) üzerindeki tedavi kolunun (ilaç alma) etkisi olarak tanımlanabilir. Alan yazında tedavi değişkeni genel olarak ikili olarak çalışılmaktadır (17). Bu tez çalışmasında ise tedavi değişkeni üç kategorili olarak ele alınmaktadır. Gruplardaki her bir birey için bu tedavileri alıp almama durumuna göre potansiyel sonuçlar ortaya çıkacaktır. Gruplardaki her bir birey için potansiyel yanıtlar $Y(m_k)$, $k = 1,2,3$ olarak düşünüldüğünde bir birey için aldığı asıl tedavi $Y(m_1)$ iken aynı bireye ilişkin karşıolgusallar $Y(m_2)$ ve $Y(m_3)$ olacaktır. Dolayısıyla $Y(m_1), Y(m_2), Y(m_3)$ yanıtları aynı birey için potansiyel yanıtlar olacaktır. Birey için nedensel etkiyi ortaya koymak için tedavi kollarına karşılık gelen potansiyel yanıtlar arasındaki farkın bulunması gereklidir. Burada bir birey karşılaştırılacak olan tüm tedavi kollarında gözlemlenemediğinden dolayı bireylere ilişkin nedensel etki kestirilemeyecektir. Bu etkileri kestirmek için ortalama tedavi etkisi (OTE) ve tedavi edilenler arasında ortalama tedavi etkisi (TOTE) adı verilen tedavi etkisi kestiricileri kullanılmaktadır (16,18). Bu iki tedavi etkisi gruplar ne kadar birbirine benzer ise o kadar yakın sonuçlar vermektedir. Dolayısıyla rasgele kontrollü bir çalışmada bu iki değer aynı çıkmaktadır (5).

İlgili disiplinde hangi tedavi etkisinin kullanılacağı, o disiplinde sorulan soruya ve seçilen popülasyona göre değişmektedir. Uygulanacak tedaviler popülasyondaki her bir birey için uygun görülebiliyorsa burada OTE kestirimi daha uygundur. Örneğin madde kullanımının bırakılmasına yönelik uygulanacak bir yoğun tedavi programındaki etkinlik kestirildiğinde, çalışmadaki bireylerin programa katılımları ve tamamlamalarında potansiyel olarak büyük engel yaşamaları bekleneceğinden, programın mevcut tüm madde bağımlılarına uygulandığındaki etkinliğini kestirmeye çalışmak gerçekçi olmayabilir. Bu durumda TOTE kestiriminin uygulanması daha uygundur (5,7). Sonuç olarak çalışma grubundaki tüm bireylerin faydalanacağı bir durum varsa OTE, bireylerin kendilerinin atanabileceği bir durum söz konusu ise TOTE kestirimini kullanmak daha faydalı olacaktır.

2.2.1 Ortalama Tedavi Etkisi (OTE)

İkili durumda tedavi ve kontrol grubu için gruplar arasındaki yanıt üzerinden elde edilen ortalama farktır. Yani popülasyondan rasgele seçilen bir birey için beklenen kazanç da denilebilmektedir (7).

$$E(F[1,0]) = E(Y[1] - Y[0]) \quad (2.1)$$

K tedavi için düşünüldüğünde $K(K - 1)/2$ tane OTE bulunacaktır. Çalışma üç tedavi grubu için düşünüldüğünde $\frac{3(3-1)}{2} = 3$ tane OTE vardır. Bunlar;

$$E(F[m_1, m_2]) = E(Y[m_1] - Y[m_2]) = E(Y[m_1]) - E(Y[m_2]) = \mu_{m_1} - \mu_{m_2} \quad (2.2)$$

$$E(F[m_1, m_3]) = E(Y[m_1] - Y[m_3]) = E(Y[m_1]) - E(Y[m_3]) = \mu_{m_1} - \mu_{m_3} \quad (2.3)$$

$$E(F[m_2, m_3]) = E(Y[m_2] - Y[m_3]) = E(Y[m_2]) - E(Y[m_3]) = \mu_{m_2} - \mu_{m_3} \quad (2.4)$$

Eşitliklerde;

m_k , $k = 1,2,3$: Tedavi gruplarını

F : Tedaviler arası fark

$Y[m_k]$: İlgili tedavi grubundaki yanıtı ve

μ_{m_k} : İlgili tedavi grubundaki ilgili değişken için ortalamayı temsil etmektedir.

2.2.2 Tedavi Edilenler Arasında Ortalama Tedavi Etkisi

Tedavi edilenler arasında ortalama tedavi etkisi ise gerçekten tedavi edilenler için tedavideki ortalama kazanç olarak tanımlanmaktadır. Populasyonda asıl olarak tedavi alan bireyler arasından rasgele seçilen bir birey için beklenen tedavi etkisidir (17).

K tedavi için düşünüldüğünde $K(K - 1)$ tane TOTE bulunacaktır. Çalışma üç tedavi grubu için düşünüldüğünde $3 \cdot (3 - 1) = 6$ tane TOTE vardır: Bunlar;

1. tedavi ile tedavi edilenler arasında 2. ve 1. tedavi arasındaki ortalama fark:

$$\mu_{m_1, m_2} - \mu_{m_1, m_1}, \quad (2.5)$$

1. tedavi ile tedavi edilenler arasında 3. ve 1. tedavi arasındaki ortalama fark :

$$\mu_{m_1,m_3} - \mu_{m_1,m_1}, \quad (2.6)$$

2. tedavi ile tedavi edilenler arasında 2. ve 1. tedavi arasındaki ortalama fark :

$$\mu_{m_2,m_2} - \mu_{m_2,m_1}, \quad (2.7)$$

2.tedavi ile tedavi edilenler arasında 3. ve 2. tedavi arasındaki ortalama fark :

$$\mu_{m_2,m_3} - \mu_{m_2,m_2}, \quad (2.8)$$

3. tedavi ile tedavi edilenler arasında 3.ve 1. tedavi arasındaki ortalama fark :

$$\mu_{m_3,m_3} - \mu_{m_3,m_1}, \quad (2.9)$$

3. tedavi ile tedavi edilenler arasında 3. ve 2. tedavi arasındaki ortalama fark :

$$\mu_{m_3,m_3} - \mu_{m_3,m_2}, \quad (2.10)$$

şeklindedir.

3 PROPENSITY SKORU

Yarı deneysel yada gözlemsel çalışmalarda tedavi kolları ve yanıt değişkeni ile ilişkili olan ortak değişken etki karışımına neden olmaktadır. Verilen tedavinin yanıt değişkeni üzerindeki etkisi ortak değişkenin düzeylerine göre değişebilmektedir yani yanıt üzerinde tedavinin etkisi karışmaktadır (19). Etki karıştırıcının olduğu çalışmalarda ortak değişkenden kaynaklanan dengesizliği gidermek için iki yaklaşım kullanılmaktadır. Bunlardan ilki genelleştirilmiş doğrusal modellerin kullanıldığı ortak değişken düzeltmeli klasik yöntemlerdir. Burada tedavi, yanıt ve ortak değişkenler arasındaki ilişki matematiksel olarak gösterilmektedir. Bu ilişki kullanılarak yanıt değişken üzerinde tedavi etkinliği, bu etkiye etki ettiği düşünülen ortak değişkenlerin düzeltilmesi ile kestirilmeye çalışılmaktadır. İkincisi ise rasgeleleştirmenin mümkün olduğu ya da uygun deneysel bir çalışma tasarımının kullanıldığı çalışmalarda böyle bir etkileşim yoktur ve dolayısıyla yanıt üzerinde tedavinin etkinliği yansız bir şekilde bulunabilmektedir.

Tedavi etkinliğini incelemek için yapılan gözlemsel bir çalışmada ortak değişkenlerden kaynaklanan dengesizliği kontrol altına alan bir diğer yöntem ailesi ilk olarak Rosenbaum ve Rubin tarafından gösterilen propensity skoru kullanımına dayanmaktadır (14). Propensity skoruna dayanan bu yöntemler son yıllarda yaygın bir şekilde kullanılmaktadır. Tedavi gruplarında benzer propensity skor değerlerine sahip bireyler alınarak yine propensity skorları üzerinden eşleştirme, tabakalama, regresyon düzeltmesi ve ağırlıklandırma yaparak tedavi etkinlikleri bulunmaktadır (7,20).

Rasgeleleştirmenin mümkün olmadığı durumda kullanılabilecek olan model-tabanlı bir regresyon çözümlemesine göre propensity skor yöntemlerinin avantajları vardır. Tek bir skor yardımıyla tüm değişkenler özetlenebilmektedir. Klasik bir regresyon tabanlı analizde modele ilave edilecek değişken sayısında kısıtlılık vardır. Ayrıca regresyon analizinden beklenen en az değişkenle tedavi etkinliğinin belirlenmesi gibi bir fayda, ortak değişken sayısının kısıtlılığına neden olmaktadır (19,21). Ayrıca standart bir doğrusal regresyon analizinin uygulanamadığı durumlarda, doğrusal olmayan model belirlemede de zorluk ortaya çıkarmaktadır. Regresyon çözümlemesinde yanıt değişkene ilişkin ortalama modellenirken

propensity skor metotlarında böyle bir durum yoktur. Dolayısıyla model belirlemede yanlışlık yaşanmamaktadır. Son olarak regresyon tabanlı bir modelin ileriye dönük kestirim yapması mümkün iken propensity skor yöntemleri gözlenen veri üzerinden bunu yapar yani ileriye dönük bir kestirimi yoktur (7,22).

Bir dengeleme skoru olan propensity skoru; gözlenen ortak değişken üzerinden bireyin tedavi grubuna atanma olasılığıdır. m tedavi grupları olmak üzere propensity skoru aşağıdaki eşitlikle hesaplanır:

$$\pi_j = e(x_j) = \Pr(T_j = m | X_j) \quad (3.1)$$

Eşitlik (3.1)'de, X_j , C tane gözlenen ortak değişkene ilişkin vektörü ifade etmektedir. Propensity skoru gözlenen X_j ortak değişkenlerinin tedavi kollarında benzer dağılım göstermesine yardımcı olmaktadır (5).

3.1 Propensity Skorunda Gösterimler

OTE ve TOTE kestirimlerinin elde edilebilmesi için gözlenen veri için ayrı bir gösterim kullanılacaktır. $j = 1, 2, \dots, n$ olmak üzere j . birey için gözlenen tedavi durumu T_j ile gösterilsin. $m_k \in \{1, \dots, K\}$ tedavi gruplarını göstermek üzere j bireyi gerçekte m tedavisini alıyorsa $T_j = m_k$ şeklinde gösterilecektir. Y_j aynı şekilde asıl tedaviyi alan j . bireyi göstermektedir. Burada dikkat edilmesi gereken her bir bireyin K potansiyel sonucu vardır. Herbir birey için K tedaviden sadece biri gözlenebilir yanıt sahiptir (Y_j) ve diğer tedaviler altında gözlenmiş olsalardı ortaya çıkacak yanıt değişkenler karşılığusallar adını alacaktır. Daha genel olarak j . birey için $T_j = m_k$ ise $Y_j = Y_j[m_k]$ yanıtını göstermektedir. Popülasyondaki ortak değişkenlerin gösterimi için ise X_j kullanılacaktır (7).

3.2 Propensity Skor Yöntemleri

Propensity skorlarını bulmak için genellikle lojistik regresyon analizinden yararlanılmaktadır. Tedavi değişkenin bağımlı değişken ve ortak değişkenlerin bağımsız değişken olarak modele konması ile elde edilir. Kurulan modelden elde edilen tedavi gruplarına atanma olasılıkları kestirilen propensity skorlarıdır. Yanıt değişken üzerindeki tedavinin etkinliğini etkileyecek ortak değişkenlerin etkisini

kaldırmak için propensity skorunu kullanan dört farklı propensity skor yöntemi vardır. Bu yöntemlerin hepsi tedavi gruplarındaki propensity skorlarına ilişkin dağılımların benzer olmasını amaçlar (23).

3.2.1 Propensity Skoru ile Tabakalama

Tüm örneklem kullanılarak herhangi bir yöntemle kestirilen propensity skor değerleri yani tedavi grubuna atanma olasılıklarına göre bireyler sıralanır. Daha sonra bu olasılıklar daha önce belirlenen bir eşik değerine göre tabakalara ayrılır. Bu eşik değer genellikle propensity skorlarının beşte birlikleri kullanılarak beş eşit alt grup oluşturacak şekilde tabakalandırılır. Aynı tabakada olan tedavi gruplarındaki bireyler direkt olarak karşılaştırılır. Bu yöntemde etki karıştıran değişkenden kaynaklanan yanlılığın yaklaşık olarak %90 ortadan kaldırılacağı gösterilmektedir (24,25).

3.2.2 Propensity Skoru ile Eşleştirme

Kestirilen propensity skorunun kullanıldığı bir diğer yöntem olan eşleştirmede, tedavi grubundaki bir bireye ait propensity skor değerine en yakın diğer gruptan bireyin eşleştirmesi ile yapılır. Eşleştirme dışında kalan bireyler ihmal edilir. İhmal edilme genellikle kontrol grubundan olmaktadır. Dolayısıyla kontrol grubunun fazla olduğu çalışmalar için tercih edilecek bir yöntemdir (25,26).

3.2.3 Propensity Skoru ile Ortak Değişken Düzeltmesi

Propensity skoru ile tabakalama, eşleştirme ve ağırlıklandırma yönteminden farklı olarak, propensity skoru ile ortak değişken düzeltmesi yöntemi yanıt değişkenin modellenemediği bir ilişki ortaya koymaktadır. Kestirilen propensity skorunun regresyon modelinde kullanılmasından dolayı da standart regresyon düzeltmesinden farklıdır (22). X_j ortak değişkenler ve $\beta_0, \beta_j, \beta_T$ regresyon katsayıları T tedavi değişkeni ve ε hata terimi olmak üzere

$$Y = \beta_0 + \beta_j X_j + \beta_T T + \varepsilon \quad (3.2)$$

Eşitlik 3.2'deki standart bir regresyon modeline, kestirilen propensity skorlarına ilişkin değişken, bir terim olarak ilave edilerek yanıt değişken üzerinde tedavi etkinliği belirlenir. β_{PS} regresyon katsayısı ve $p(T)$ propensity skor kestirimleri olmak üzere

$$Y = \beta_0 + \beta_j X_j + \beta_T T + \beta_{PS} p(X) + \varepsilon \quad \text{dur.} \quad (3.3)$$

Propensity skorunun modelde olmasıyla β_T regresyon katsayısının etkisi düzelecektir (22).

3.2.4 Propensity Skoru ile Ağırlıklandırma

Alanyazında propensity skor ağırlıklandırma, ters olasılıklı ağırlıklandırma ya da ters olasılıklı tedavi ağırlıklandırma olarak da adlandırılan bu yöntemin tabakalama ya da eşleştirmede olduğu gibi benzer propensity skorları üzerinden herhangi bir alt gruptaki bireyleri karşılaştırmaya çalışmaz. Kestirilen propensity skorlarını bireyleri ağırlıklandırmak için kullanmaktadır. Bunu yaparken ise tedavi kollarındaki ortak değişkenlerin dengelendiği yapay ya da yalancı bir popülasyon oluşturmaktadır. Alan çalışması analizlerinde olduğu gibi ağırlıklar kullanarak, kimi gözlemlere daha az, kimisine daha fazla ağırlık vererek tedavi etkisini ölçer. Propensity skor değerleri özellikle çok küçük olduğunda, bu değerlerin çarpmaya göre tersi kullanılacağından büyük ağırlıkların bulunmasına neden olmaktadır. Bunun üstesinden gelmek için ise ağırlıklar için aşırı değerlerden arındırılmış bir veri seti ile çalışılması önerilse de, bu durum tedavi etkilerinin kestiriminde yanlılığa neden olabilmektedir (5,23,27).

OTE ve TOTE kestirimlerinde her bir ortak değişken için propensity skorları kullanılarak elde edilecek ağırlıklar yardımıyla yeni ortalamalar kestirilmektedir.

K tedavi kolunun olduğu bir çalışma düzeninde OTE için kestirim yapıldığında, burada $K = 3$ alındığında Eşitlik 2.2, 2.3 ve 2.4'te verildiği gibi $(\mu_{m_1} - \mu_{m_2}, \mu_{m_1} - \mu_{m_3}, \mu_{m_2} - \mu_{m_3})$ $3(3-1)/2=3$ kestirim yapılabilir. j . bireyin X_j ortak değişkenleri için m tedavisine atanma olasılığını $p_{m_k}(X_j)$ propensity skorunu gösterebilir.

$$p_{m_k}(X_j) = \Pr(T_j = m_k | X_j) \quad (3.4)$$

ve tedavi grubundaki bireylerin ağırlıkları,

$$w_j[m_k] = \frac{1}{p_{m_k}(X_j)} \quad (3.5)$$

olmak üzere ağırlıklandırılmış ortalama,

$$\hat{\mu}_{m_k} = \frac{\sum_{j=1}^n T_j[m_k] Y_j w_j[m_k]}{\sum_{j=1}^n T_j[m_k] w_j[m_k]} \quad (3.6)$$

olarak verilmektedir. Burada $T_j[m_k]$: j . bireyin aldığı tedavi grubu, Y_j : j . bireye ilişkin yanıt değişken değeridir. Kontrol grubu olarak alınacak tedavi grubu dışındaki grup için kullanılacak ağırlık

$$w_j[m_k] = \frac{1}{1 - p_{m_k}(X_j)} \quad \text{'dir.} \quad (3.7)$$

Kontrol grubu için kestirilecek ağırlıklandırılmış ortalama için de Eşitlik 3.6 kullanılmaktadır.

TOTE kestirimi için ikili etkiler Eşitlik 2.5'ten 2.10'a kadar olmak üzere $3(3-1)=6$ tane kestirim yapmak mümkündür. Kestirim yapmak için Eşitlik 2.5 ve 2.6 ($\mu_{m_1, m_2} - \mu_{m_1, m_1}$, $\mu_{m_1, m_3} - \mu_{m_1, m_1}$) düşünüldüğünde yani 1 tedavisini alanlarda 1 ve 2 tedavisi arasındaki farklılık karşılaştırılmak istendiğinde nedensel etki

$$\hat{\mu}_{m_1, m_2} = \frac{\sum_{j=1}^n T_j[m_1] Y_j}{\sum_{j=1}^n T_j[m_1]} \quad (3.8)$$

eşitliği kullanılarak kestirilmektedir. Bu da tedaviyi alan bireylerden elde edilecek olan ağırlıksız ortalamaları ifade etmektedir. 2. tedaviyi alan bireyleri ağırlıklandırmak için

$$\frac{p_{m_2}(X_j)}{1 - p_{m_2}(X_j)} \quad (3.9)$$

oranı kullanılmaktadır (7).

Ağırlıklandırma sonrasında yapay olarak oluşturulan örneklem gruplarının büyüklüklerini, ağırlıkları kullanarak hesaplamak mümkündür.

$$ESS_{m_k} = \frac{(\sum_{j=1}^N T_j[m_k]w_j)^2}{\sum_{j=1}^N T_j[m_k]w_j^2} \quad (3.10)$$

k gruba ilişkin örnek büyüklüğü Eşitlik 3.10'daki gibi hesaplanabilir.

3.2.5 Varsayımlar

Propensity skor yöntemlerinde öncelikle tedavi kollarının dolaylı olarak diğer bir tedavi altındaki gözlemleri etkilememesi gerekmektedir. Bir gözleme ilişkin potansiyel yanıtların diğer gözlemlerin aldığı tedaviden etkilenmemesi anlamına da gelmektedir (28,29). Bunun dışında yöntemlerin aşağıda belirtildiği gibi iki önemli varsayımı bulunmaktadır.

Koşullu bağımsızlık varsayımı

Bilinmeyen ya da ölçülemeyen etki karıştırıcının olmadığı anlamını taşıyan bu varsayımda, ortak değişken vektörü için potansiyel tedavi yanıtlarının, atama yönteminden bağımsız olması anlamına gelmektedir. Koşullu bağımsızlığa ilişkin bağıntı Eşitlik 3.11'de verilmektedir. K tedavi kolu için düşünüldüğünde;

$$Y_{m_1}, Y_{m_2}, \dots, Y_{m_k} \perp T[m_k] \mid X \quad (3.11)$$

Rasgele atamanın olduğu bir çalışma düzeninde tam bir bağımsızlık söz konusudur (Eşitlik 3.12). Yani tedaviye katılımın yanıtlara bağlı olmaması anlamına gelmektedir (28).

$$Y_{m_1}, Y_{m_2}, \dots, Y_{m_k} \perp T[m_k] \quad (3.12)$$

Özet olarak; etki karıştırıcı olarak düşünülen değişkenlerin propensity skor modelinde kullanılması gerektiğini ifade eden bir varsayım olarak da düşünülebilir (7). Bir ortak değişkenin etki karıştırıcı olup olmadığına karar vermek için ilgili disiplin içerisindeki araştırmacının da bilgisine başvurmak daha doğru olacaktır. Bu

anlamda modele alınacak olan bu tür ortak değişkenler direkt olarak çalışmada veri üzerinde ölçülememektedir.

Pozitiflik varsayımı

Eşleştirme ya da örtüşme olarak da bilinen bu varsayım da her bir tedaviyi alan bireylerin olasılık değerlerinin pozitif olması gerektiğini ifade eder (Eşitlik 3.13).

$$0 < pr(T_j = m_k | X) < 1 \quad (3.13)$$

Tedavi gruplarından elde edilen propensity skorlarına ilişkin dağılımın tedavi gruplarına yakın olması ya da örtüşmesi pozitiflik varsayımının sağlandığını göstermektedir.

3.3 Propensity Skor Kestirim Yöntemleri

Propensity skorlarını kestirmek için geleneksel istatistiksel yöntemler olan lojistik regresyon, diskriminant analizi ve kümeleme analizi gibi yöntemlerden yararlanıldığı gibi, yapay zeka algoritmalarından olan CART, budanmış CART, bagged CART, random forest ve GBM gibi metotlardan da yararlanılmaktadır. Bu çalışma kapsamında MLR ve GBM yöntemleri kullanılmaktadır.

3.3.1 Genelleştirilmiş Boosted Model (GBM)

GBM ya da genelleştirilmiş boosted regresyon ağaçları da denen bu model, kestirim yada sınıflama yapmak için yinelemeli modeller kurmaktadır. Başlangıçta bir regresyon ağacı ile başlanarak her bir tekrarda bu ağaca yeni bir regresyon ağacı eklenir. Kurulan her bir modelin veriye uyumuna bakılarak log olabilirlikte en büyük artışı sağlayan model kullanılır (7,30). Doğrusal olmayan ilişkileri yakalayabilecek şekilde çoklu regresyon modelleriyle kestirim yapmaktadır. Burada en uygun tekrar (ağaç) sayısını bulmak için kullanılan durma kuralları vardır. Bunlar standartlaştırılmış yanlılık ve Kolmogorov Smirnov istatistiğinden elde edilen değerlerdir. Bu değerler dengeyi ortaya koyan performans ölçüleri ya da metriklerdir. Her bir değişken için hesaplanan maksimum standartlaştırılmış yanlılık (Mak.es), ortalama standartlaştırılmış yanlılık (mean.es), maksimum Kolmogorov Smirnov

istatistiği (Mak.ks) yada ortalama Kolmogorov Smirnov istatistiği (Mean.ks) değerleri kullanılarak en uygun ağaç sayısı bulunur. Gereğinden fazla ağaç sayısı model karmaşık hale geldikçe değişken değerlerinde bozulmalara neden olabilmektedir (31).

3.3.2 Çok Kategorili Lojistik Regresyon

Çok kategorili lojistik regresyon yanıt (bağımlı) değişkenin ikiden fazla kategoriye sahip olduğu durumda ortak (bağımsız) değişkenler ile yanıt değişken arasındaki ilişkinin belirlendiği bir yöntemdir. Doğrusal regresyondan farklı olarak, bağımlı değişkenin kategorilerinden birine atanma olasılığı aşağıdaki Eşitlik ile kestirilmektedir (19). Bu olasılık Eşitlik 3.14'te verilmektedir.

$$\pi_j(X_j) = P(T_j = m_k | X_j) = \frac{e^{\beta_{m_k} X_j}}{1 + \sum_{m_k=1}^{K-1} e^{\beta_{m_k} X_j}} \quad (3.14)$$

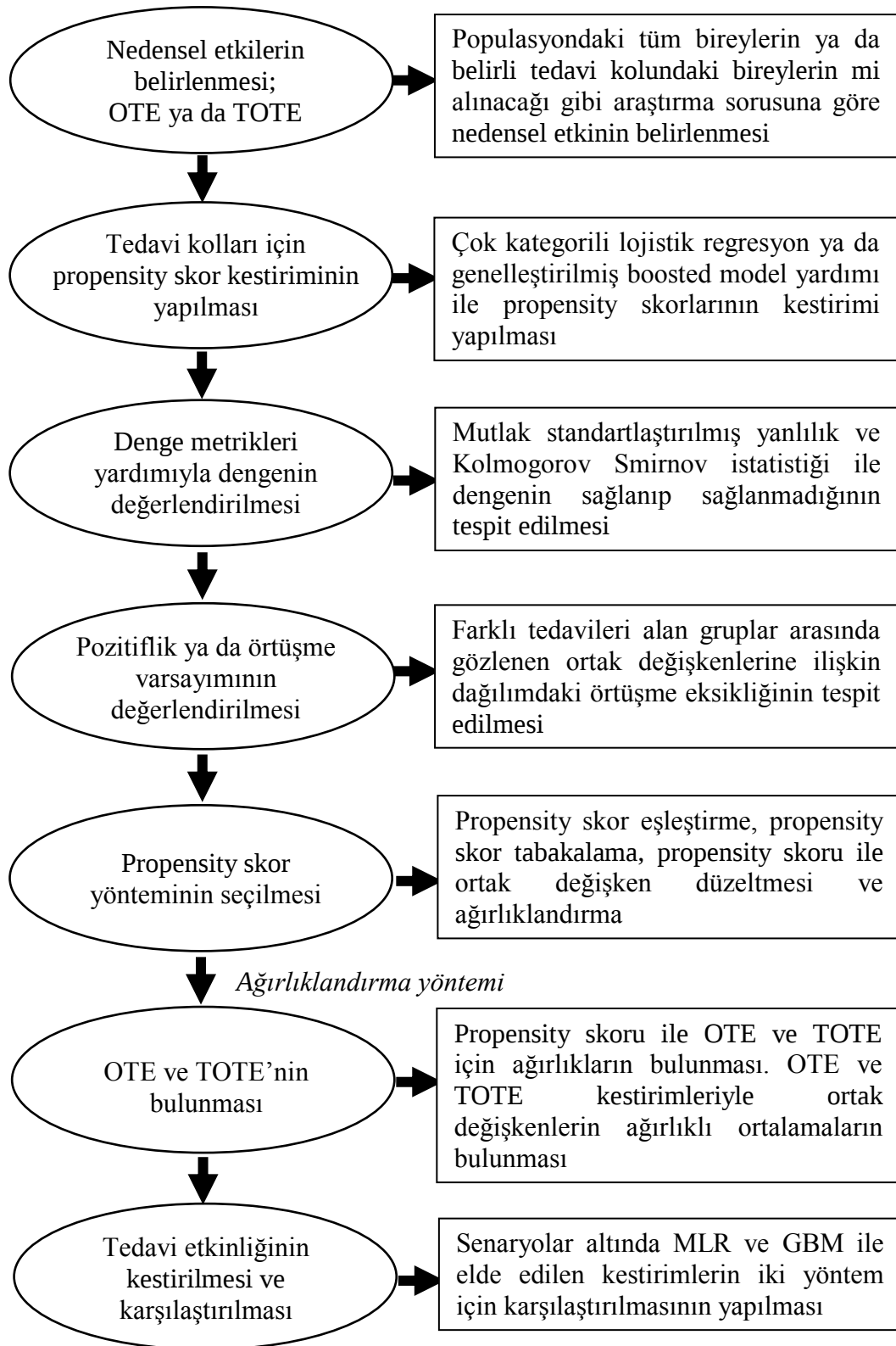
Eşitlikte;

$k = 1, 2, \dots, K - 1$ olmak üzere m_k : tedavi kollarını ifade eder.

Ayrıca tedavi kollarına atanma olasılıklar toplamı 1' e eşittir. Eşitlik 3.15'de verilmiştir.

$$\sum_{m_k=1}^K P(T_j = m_k | X_j) = 1 \quad \text{'dir.} \quad (3.15)$$

Bu bölümde verilen bilgilerin tez çalışması içerisinde nasıl ve hangi sıra ile kullanıldığının anlaşılması açısından, çalışmada nedensel etkilerin belirlenmesinde kullanılan akış şeması Şekil 3.1'de verilmiştir.



Şekil 3.1. Tedavi etkinlikleri kestiriminde akış şeması

4 GEREÇ VE YÖNTEM

4.1 R’de Kullanılan Paketler

Analizler, R 3.2.0 programı ile yapıldı. GBM için “Twang”, Çok kategorili Lojistik regresyon için “nnet” paketleri kullanıldı (32,33). Veri setlerini benzetim yardımıyla türetilebilmek için R programlama dili kullanıldı. Benzetim çalışmasında “car”, “doSNOW”, “foreach”, “digest”, “survey”, “foreign”, “ggplot2”, “reshape2” R paketlerinden yararlanıldı (34-40).

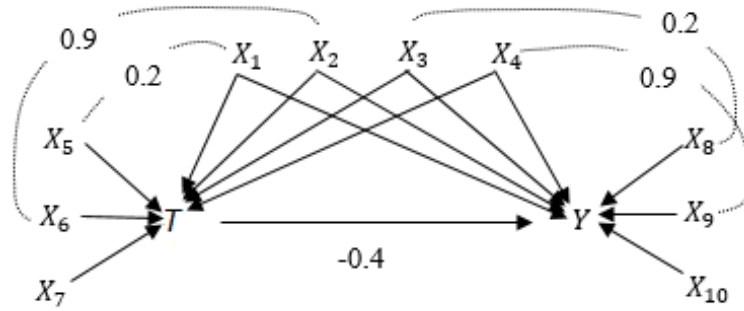
4.2 Benzetim Çalışması

Yedi senaryo altında türetilen veri setlerinde GBM ve çok kategorili lojistik regresyon yöntemlerinin ters olasılıklı ağırlıklandırma yardımıyla karşılaştırmasında 1000 tekrarlı Monte Carlo simülasyon yöntemi kullanıldı. Herbir senaryo için $n=500$, 1000, 2000 örneklem büyüklüğüne sahip veri setleri türetildi. Karşılaştırmalarda parametre kestirimleri, mutlak standartlaştırılmış yanlışlık ve Kolmogorov Smirnov testi ölçütleri kullanıldı. Senaryolar Ek 2’de verilmektedir.

4.3 Benzetim Algoritmaları

Propensity skor modellerini oluştururken, karesel ve etkileşim terimlerinin dikkate alınma durumuna göre, doğrusallığın ve toplamsallığın farklılaştığı yedi senaryoda genelleştirilmiş boosted ve çok kategorili lojistik regresyon modellerinin performansları değerlendirilmiştir. Bu modeller ile kestirim yapılırken ters ağırlıklandırma yöntemi kullanılmaktadır.

Benzetim çalışması için veri yapısı, Setoguchi (2008) vd’nin gerçek bir veri setinden elde ettiği yapı kullanılmaktadır. Bu çalışmada farklılık olarak tedavi değişkeni üç kategori olarak alınmaktadır. Veri yapısı Şekil 4.1’deki gibi olacaktır.



Şekil 4.1. Benzetim verisi için değişken yapısı

Şekil 4.1’de tedavi ve yanıt değişken üzerinde etkisi önemli olan değişkenler ok ile gösterilmektedir. Şekil 4.1’de,

T: üç kategorili tedavi (etken, asıl bağımsız değişken),

Y: yanıt değişken,

$X_1 - X_4$: hem etkeni hem de sonucu etkileyen ortak değişkenler,

$X_5 - X_7$: sadece etkenle ilişkili kestiriciler,

$X_8 - X_{10}$: sadece yanıt değişkenle ilişkili kestiriciler,

$X_1, X_3, X_5, X_6, X_8, X_9$: iki kategorili değişkenler,

Y, X_2, X_4, X_7, X_{10} : sürekli değişkenlerdir.

İlk olarak standart normal dağılım kullanılarak $X_1, X_3, X_4, X_7, X_{10}$ değişkenleri türetilacaktır. X_1 değişkeni kullanılarak arasındaki korelasyon 0,2 olacak şekilde X_5 değişkeni elde edilecektir. Korelasyon 0,2 olacak şekilde X_3 yardımıyla X_8 değişkeni elde edilecektir. Korelasyon 0,9 olacak şekilde X_2 yardımıyla X_6 değişkeni elde edilecektir. Son olarak korelasyon 0,9 olacak şekilde X_4 yardımıyla X_9 değişkeni elde edilecektir. $X_1, X_3, X_5, X_6, X_8, X_9$ değişkenlerini dikotom hale dönüştürmek için herbir değişkenin gözlemleri ilgili değişkenin ortalaması ile karşılaştırılacak ve ilgili değişkenin gözlemi değişken ortalamasından büyük ise o gözleme 1 değilse 0 atanarak değişken ikili yapıya dönüştürülecektir.

Çoklu tedavi T değişkeni X_i nin bir fonksiyonu olarak çok kategorili lojistik model kullanılarak modellenenecektir. Bu fonksiyonun eşitliği (gerçek propensity

skorları) aşağıda verilen yedi senaryo altında elde edilecektir. Senaryolara ilişkin kullanılan katsayılar ve Eşitlikler Ek 2 ve Ek 3’de verilmektedir.

Senaryolar

1. Sadece ana etkilerin olduğu toplamsallık ve doğrusallığın olduğu model
2. Bir karesel terimin olduğu model (hafif doğrusal olmayan model)
3. Üç karesel terimin olduğu model (ılımlı doğrusal olmayan model)
4. Üç tane iki yönlü etkileşim terimli hafif toplamsal olmayan model
5. Dört tane iki yönlü etkileşim terimli ve bir karesel terimin olduğu hafif toplamsal ve doğrusal olmayan model
6. 10 tane iki yönlü etkileşim terimli ılımlı toplamsal olmayan model
7. 10 tane iki yönlü etkileşim terimli ve üç tane iki yönlü etkileşim terimli ılımlı toplamsal ve doğrusal olmayan modeldir.

Y yanıt değişkeni ise T ve X_i nin bir fonksiyonu olarak modellenenecektir. Yanıt değişkene ilişkin kullanılan modeller ve katsayılar Ek 2 ve Ek 3’de verilmektedir.

4.3.1 Ortak Değişken ve Tedavi Değişkeni Türetimi

İlk olarak standart normal dağılım kullanılarak $X_1, X_3, X_4, X_7, X_{10}$ değişkenleri türetilenecektir. X_1 değişkeni kullanılarak arasındaki korelasyon 0.2 olacak şekilde X_5 değişkeni elde edilecektir. Korelasyon 0.2 olacak şekilde X_3 yardımıyla X_8 değişkeni elde edilecektir. Korelasyon 0.9 olacak şekilde X_2 yardımıyla X_6 değişkeni elde edilecektir. Son olarak korelasyon 0.9 olacak şekilde X_4 yardımıyla X_9 değişkeni elde edilecektir. $X_1, X_3, X_5, X_6, X_8, X_9$ değişkenlerini dikotom hale dönüştürmek için herbir değişkenin gözlemleri ilgili değişkenin ortalaması ile karşılaştırılacak ve ilgili değişkenin gözlemi değişken ortalamasından büyük ise o gözleme 1 değilse 0 atanarak değişken ikili yapıya dönüştürülecektir.

Çoklu tedavi T değişkeni X_i nin bir fonksiyonu olarak multinomial lojistik model kullanılarak modellenenecektir. Bu fonksiyonun eşitliği (gerçek propensity skorları) Ek 3’de verilen yedi senaryo altında elde edilecektir.

4.3.2 Yanıt Değişken Türetimi

Y yanıt değişkeni ise T ve X_i nin bir fonksiyonu olarak modelleneyecektir. Eşitlik 4.1'deki kullanılan katsayılar Ek 3'de verilmektedir.

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \alpha_5 X_8 + \alpha_6 X_9 + \alpha_7 X_{10} + \gamma_1 T \quad (4.1)$$

4.3.3 Veri Setlerinin Türetimi

1000 tekrarlı 7 farklı senaryo oluşturularak $n=500, 1000, 2000$ örneklem büyüklüğüne sahip veriler türetilenektir. Senaryolar Ek 2'de verilmektedir.

4.4 Performans Ölçüleri

GBM ve MLR'den elde edilen propensity skorları kullanılarak bireyler ağırlıklandırıldıktan sonra her bir ortak değişken için ağırlıklandırmadan önceki ve sonraki değerleri üzerinden hesaplanacak iki performans ölçüsü ya da metrik vardır. Bunlar standartlaştırılmış yanlılık ve Kolmogorov Smirnov istatistiğidir.

4.4.1 Standartlaştırılmış yanlılık

Bireylerin ağırlıklandırılmasından önceki tedavi ve kontrol olarak seçilen gruplar arasındaki mutlak fark ile ağırlıklandırdıktan sonraki tedavi ve kontrol olarak seçilen gruplar arasındaki mutlak farkın değişimi incelenir. Mutlak farkın azalması grupların benzer olduğunu ve dengenin sağlandığını göstermektedir. OTE için tedavi ve kontrol grubu için ağırlıklandırılmış ortalamalar arasındaki farkın ortak standart sapmaya bölünmesi ile elde edilir. TOTE için ise tedavi ve kontrol için ağırlıklandırılmış ortalamalar arasındaki farkın tedavi grubu standart sapmasına bölünmesi ile elde edilir. OTE ve TOTE için standartlaştırılmış yanlılıklara ilişkin çoklu tedavi için genelleştirme Eşitlik 4.2 ve 4.4'deki gibidir.

OTE için mutlak standartlaştırılmış yanlılık (MSB), $c = 1, 2, \dots, C$ ortak değişken sayısını ifade etmek üzere

$$MSB_{cm_k} = \left| \bar{X}_{cm_k} - \bar{X}_{cm_p} \right| / \hat{\sigma}_{cm_p} \quad \text{'dir.} \quad (4.2)$$

Eşitlikte;

$\bar{X}_{cm_k}$: c. ortak değişken için k tedavisini alanlar için ortalama,

$\bar{X}_{cm_p}$: c. ortak değişken için popülasyona ilişkin ortalama,

$\hat{\sigma}_{cm_p}$: c. ortak değişken için popülasyona ilişkin standart sapmadır.

$\bar{X}_{cm_k}$ tedavi grubu olarak alındığında Eşitlik 3.5'in gösterim olarak farklı bir yazımı ile ilgili ortak değişken için ağırlıklı ortalama

$$\bar{X}_{cm_k} = \frac{\sum_{j=1}^n T_j[m_k] \bar{X}_{cj} w_j[m_k]}{\sum_{j=1}^n T_j[m_k] w_j[m_k]} \quad \text{'dir.} \quad (4.3)$$

TOTE için mutlak standartlaştırılmış yanlılık,

$$MSB_{cm_k} = |\bar{X}_{cm_k} - \bar{X}_{cm_t}| / \hat{\sigma}_{cm_k} \quad \text{'dir.} \quad (4.4)$$

Eşitlikte;

$\bar{X}_{cm_k}$: c. ortak değişken için k tedavisini alanlar için ortalama

$\bar{X}_{cm_t}$: c. ortak değişken için t tedavi grubuna ilişkin ortalama

$\hat{\sigma}_{cm_k}$: c. ortak değişken için k tedavi grubuna ilişkin standart sapma

Standartlaştırılmış yanlılık değerleri dengeyi değerlendirmek için kullanılmaktadır. Genel olarak, 0,20'den daha düşük standartlaştırılmış ortalama farklar küçük olarak düşünülür, 0,40 orta seviye, 0,60 büyük olarak düşünülmektedir (41).

Benzetim çalışmasında her bir yöntemde, 1000 farklı veri setinden elde edilen standartlaştırılmış yanlılık değerlerinin ortalamaları alınmaktadır.

4.4.2 Kolmogorov Smirnov İstatistiği

Kolmogorov Smirnov (KS) istatistiği tedavi kollarındaki örneklemelerin ağırlıklandırılmış deneysel dağılım fonksiyonlarına dayanmaktadır. Bu test tedavi kolları arasında ortak değişkenlere ilişkin dağılımın örtüşüp örtüşmediğini belirlemeye yardımcı olmaktadır (42). Ayrıca tüm dağılımı karşılaştırmak mümkündür.

c ortak değişken ve m_1 tedavi ve m_2 kontrol olarak düşünüldüğünde

$$I(X_{jc} \leq x) = \begin{cases} 1, & X_{jc} \leq x \\ 0, & X_{jc} > x \end{cases} \quad (4.5)$$

için olasılık yoğunluk fonksiyonu

$$YF_{cm_k}(x) = \frac{\sum_{j=1}^n w_j [m_k] T_j [m_k] I(X_{jc} \leq x)}{\sum_{j=1}^n w_j [m_k] T_j [m_k]} \quad (4.6)$$

olarak tanımlanmaktadır. Aynı ortak değişken için Kolmogorov Smirnov istatistiği

$$KS_c = \sup_x |YF_{cm_1}(x) - YF_{cm_2}(x)| \quad \text{'dir.} \quad (4.7)$$

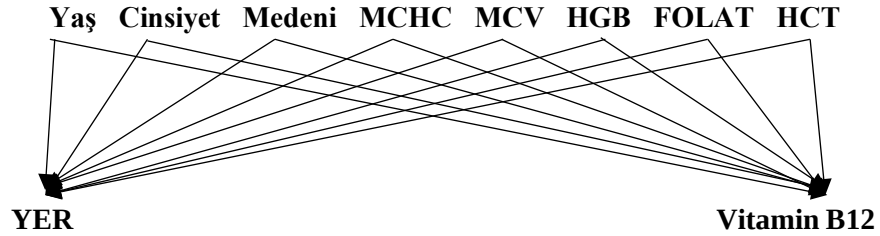
Eşitlik 4.7'de gösterildiği gibi tedavi kolları arasındaki en büyük fark değeri Kolmogorov Smirnov istatistiğini vermektedir. KS istatistiği ile denge değerlendirmesinde büyük örneklerde 0,10'dan büyük KS istatistiğinin dengenin bozulduğunu göstermektedir. Her bir yöntemde, 1000 farklı veri setinden elde edilen KS değerlerinin ortalamaları alınmaktadır.

4.5 Gerçek Veri Seti Çalışması

Gaziosmanpaşa Üniversitesi Sağlık Araştırma ve Uygulama merkezine 2010-2015 yılları arasında başvuran ve aynı anda Vitamin B12, MCHC (Mean Corpuscular Hemoglobin Concentration-Ortalama Alyuvar Hemoglobin Derişimi), MCV (Mean Corpuscular Volume-Ortalama Alyuvar Çapı), HGB (Hemoglobin), FOLAT (Folik Asit), HCT (Hematokrit) laboratuvar tetkiklerinin çalışıldığı 5549 birey örneklem olarak alınmaktadır. Çalışmanın etik kurul izni Gaziosmanpaşa Üniversitesi Klinik Araştırmalar Etik Kurulundan alınmıştır (Proje no: 15-KAEK-160).

Çalışmada, bireylerin yaşadıkları yere göre beslenme alışkanlarının farklılaşabileceği göz önüne alınarak, Tokat Merkez, Tokat İlçeleri ve Tokat Dışı olarak oluşturulan gruplar tedavi kolları olarak düşünülmektedir. Böylece bireylerin yaşadıkları yerin üç gruba (Tokat Merkez, Tokat İlçeleri ve Tokat Dışı) ayrıldığı ve tedavi değişkeni olarak düşünülen yer değişkeninin, yanıt değişkeni (vitamin B12)

üzerinde etkisi araştırılmaktadır. Çalışmada vitamin B12 yanıt değişkeni üzerinde etkili olduğu düşünülen diğer laboratuvar tetkikleri ile beraber yaş, cinsiyet ve medeni durum değişkenleri de ortak değişken olarak alınmaktadır. Kısaca düşünülen model Şekil 4.2’de verilmektedir.



Şekil 4.2. Uygulama veri seti yapısı

Analizde kullanılan değişkenlerin yaşanan yer grubuna göre dağılımları Tablo 4.1’de verilmektedir. Buna göre sürekli veriler için MCHC, MCV, HGB, HCT ve vitamin B12 değişkenlerinin ortalamaları, yaşanan yere göre farklılık göstermemektedir ($p>0,05$). Cinsiyet ve medeni gruplara göre dağılımında farklılık bulunmaktadır ($p<0,05$). Değişkenleri kontrol altına alarak vitamin B12 yanıt değişkeni üzerindeki etkisini görebilmek için yapılacak çoklu doğrusal regresyon analizi sonuçları da Tablo 4.2’de verilmektedir. Kurulan modelin anlamlı olduğu görülen ($F=19,612$; $p<0.001$) Tablo 4.2’ye göre yaş, MCV ve FOLAT değişkenlerinin vitamin B12’nin kestiriminde anlamlı olduğu görülmektedir ($p<0,05$). Burada doğrusal regresyon analizinin çok değişkenli normal dağılım sağlanması, varyans kovaryans matrisinin homojenliği gibi varsayımlarından ve modele ilave edilecek değişken sayısının sınırlı olabilmesi gibi kısıtlılıklardan dolayı alternatif bir analiz uygulanabilir. Bu durumda propensity skor analizinden yararlanılarak, tedavi değişkeni olarak düşünülen, yaşanan yer değişkeninin vitamin B12 üzerindeki gerçek etkisinin ortaya konulması sağlanacaktır.

Tablo 4.1 Uygulama verisinde kullanılan değişkenlerin yaşadığı yere göre tanımlayıcı istatistikleri ve hipotez testleri (n=5448)

Değişkenler	Yaşadığı Yer			F	p
	Tokat Merkez (n=1354)	Tokat İlçe (n=1806)	Tokat Dışı (n=2288)		
YAŞ	31,99±13,38 ^a	38,39±17,73 ^b	34,83±16,64 ^c	61,344	<0,001
MCHC	32,82±1,65	32,79±1,75	32,85±1,67	0,570	0,566
MCV	83,66±7,42	84,05±7,64	84,04±7,52	1,322	0,267
HGB	13,35±2,14	13,43±2,16	13,47±2,09	1,443	0,236
FOLAT	8,18±3,2 ^a	8,77±3,35 ^b	8,48±3,23 ^c	12,629	<0,001
HCT	40,76±5,5	40,84±5,53	41,03±5,51	1,170	0,310
Vitamin B12	313,5±139,58	317,7±150,35	318,16±145,03	0,483	0,617
Cinsiyet	n(%)	n(%)	n(%)	χ^2	p
	Erkek	Kadın			
Medeni Durum	377(27,8)	623(34,5)	748(32,7)	17,691	<0,001
Hal	977(72,2)	1183(65,5)	1540(67,3)		
Bekar	678(50,1)	757(41,9)	1194(52,2)	43,308	<0,001
Evli	676(49,9)	1049(58,1)	1094(47,8)		

a,b,c: Farklı üst indisler istatistiksel olarak farklı grupları göstermektedir (p<0,05).

Tablo 4.2 Vitamin B12 üzerinde etkisi olduğu düşünülen değişkenlere ilişkin doğrusal regresyon modeli sonuçları

Değişken	b _i	S(b _i)	Beta	VIF	t	p
Sabit	296,866	45,029			6,593	<0,001
Tokat İlçe	-5,754	5,215	-0,019	1,604	-1,103	0,270
Tokat Dışı	0,460	4,938	0,002	1,581	0,093	0,926
Cinsiyet	4,884	5,015	0,016	1,459	0,974	0,330
Medeni Durum	2,716	5,213	0,009	1,807	0,521	0,602
YAŞ	1,350	0,163	0,153	1,914	8,283	<0,001
MCHC	0,441	1,460	0,005	1,619	0,302	0,763
MCV	-1,305	0,316	-0,068	1,507	-4,131	<0,001
HGB	0,263	1,729	0,004	3,602	0,152	0,879
FOLAT	3,205	0,620	0,072	1,095	5,167	<0,001
HCT	0,636	0,647	0,024	3,383	0,983	0,326
n=5448	s=143,061	R=0,187	R ² =0,035	(F=19,612; p<0,001)		

Yer tedavi (Tokat Merkez(referans)-Tokat ilçe-Tokat dışı) değişkeni göstermelik değişken olarak kodlanmaktadır. Cinsiyet için erkek grubu referans, medeni durum için bekar grubu referans seçilmiştir.

5 BULGULAR

Çalışmada her bir senaryo altında $n=500$, 1000 ve 2000 örneklem büyüklüğüne sahip her biri için 1000 farklı veri seti türetilmektedir. Senaryolardan elde edilen gerçek propensity skor değerleri tedavi değişkenin türetimi için kullanılmaktadır. Özet denge değerleri olarak hazırlanan tablolarda tedavi kollarındaki örneklem büyüklükleri, 1000 farklı veri setinden elde edilen herbir tedavi grubundaki örneklemelerin kendi gruplarındaki örneklem büyüklüklerinin ortalamaları alınarak bulunmaktadır. Aynı şekilde diğer özet istatistikler için de 1000 farklı veri setinden elde edilen değerlerin ortalamaları kullanılmaktadır.

Tezin benzetim çalışması kısmı iki ana aşamadan oluşmaktadır. İlk aşamada, belirlenen senaryolar altında oluşturulan modellerden propensity skor kestirim yöntemlerinden GBM ve MLR ile propensity skorlarını kestirip, propensity skor ağırlıklandırma yöntemi yardımıyla ağırlıkların elde edilmesidir. Elde edilen ağırlıklar yardımıyla ortak değişkenlere ilişkin denge metrikleri değerlendirilmektedir.

İkinci aşamada ise, elde edilen bu ağırlıklar yardımıyla, ortalama tedavi etkisi kestiricisi ve tedavi edilenler arasında ortalama tedavi etkisi kestiricisi (OTE, TOTE) yani her iki kestiriciye göre tedavinin yanıt üzerindeki etkisi değerlendirilmektedir.

5.1 Denge Değerlendirmesi Sonuçları

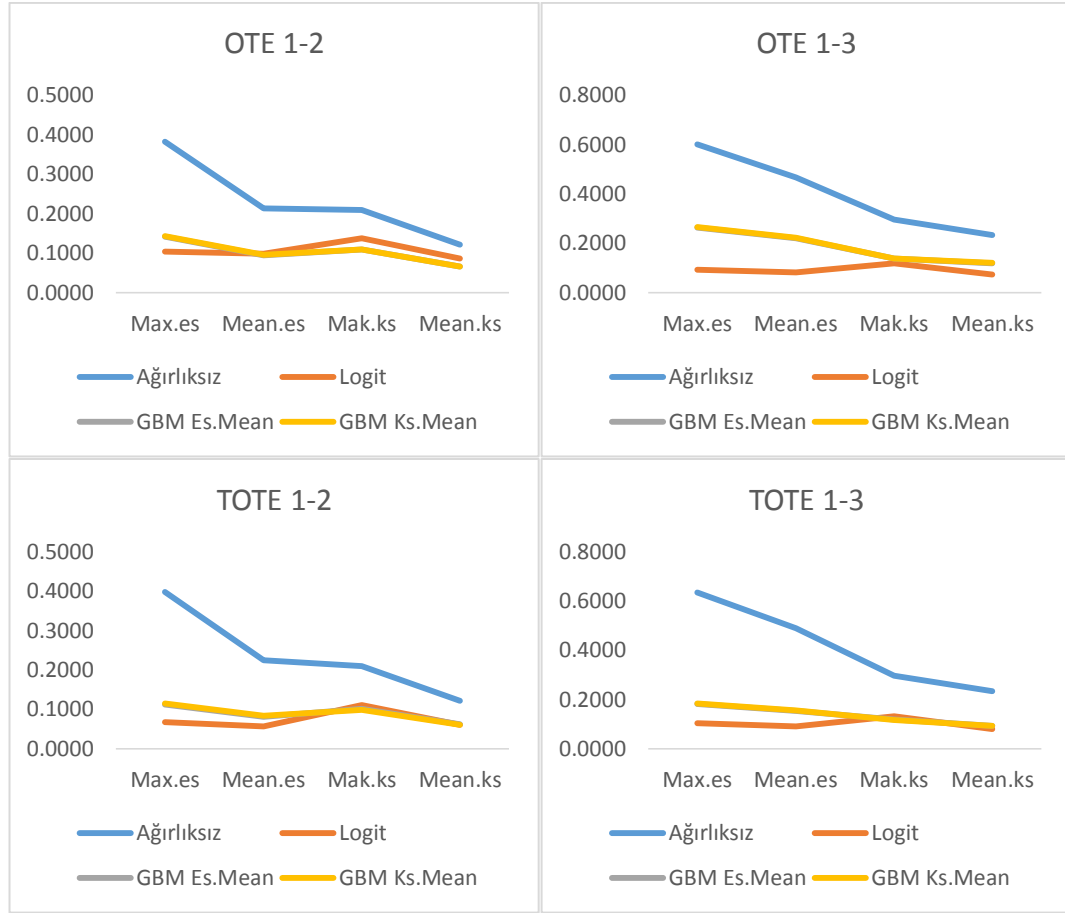
5.1.1 n=500 için Senaryoların Denge Değerlendirmesi

Senaryo 1, n=500

Ana etkilerin olduğu model

Tablo 5.1. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 1; n=500)

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	139,005	101,669	139,005	101,669	0,3820	0,2132	0,2084	0,1206
	1-3	139,005	259,326	139,005	259,326	0,6001	0,4665	0,2963	0,2328
Ağırlıksız TOTE	1-2	139,005	101,669	139,005	101,669	0,3968	0,2232	0,2084	0,1206
	1-3	139,005	259,326	139,005	259,326	0,6333	0,4874	0,2963	0,2328
MLR OTE	1-2	139,005	101,669	62,060	72,385	0,1034	0,0978	0,1369	0,0854
	1-3	139,005	259,326	62,060	159,162	0,0930	0,0821	0,1185	0,0725
MLR TOTE	1-2	139,005	101,669	139,005	64,088	0,0663	0,0566	0,1096	0,0607
	1-3	139,005	259,326	139,005	53,115	0,1028	0,0902	0,1310	0,0787
GBM Mean.es OTE	1-2	139,005	101,669	100,529	82,716	0,1416	0,0943	0,1094	0,0662
	1-3	139,005	259,326	100,529	217,432	0,2634	0,2204	0,1373	0,1189
GBM Mean.ks OTE	1-2	139,005	101,669	100,868	82,348	0,1426	0,0951	0,1086	0,0661
	1-3	139,005	259,326	100,868	218,122	0,2661	0,2216	0,1382	0,1192
GBM Mean.es TOTE	1-2	139,005	101,669	139,005	60,638	0,1111	0,0805	0,1008	0,0611
	1-3	139,005	259,326	139,005	80,624	0,1810	0,1536	0,1175	0,0921
GBM Mean.ks TOTE	1-2	139,005	101,669	139,005	62,086	0,1142	0,0826	0,0980	0,0601
	1-3	139,005	259,326	139,005	82,366	0,1823	0,1550	0,1152	0,0916

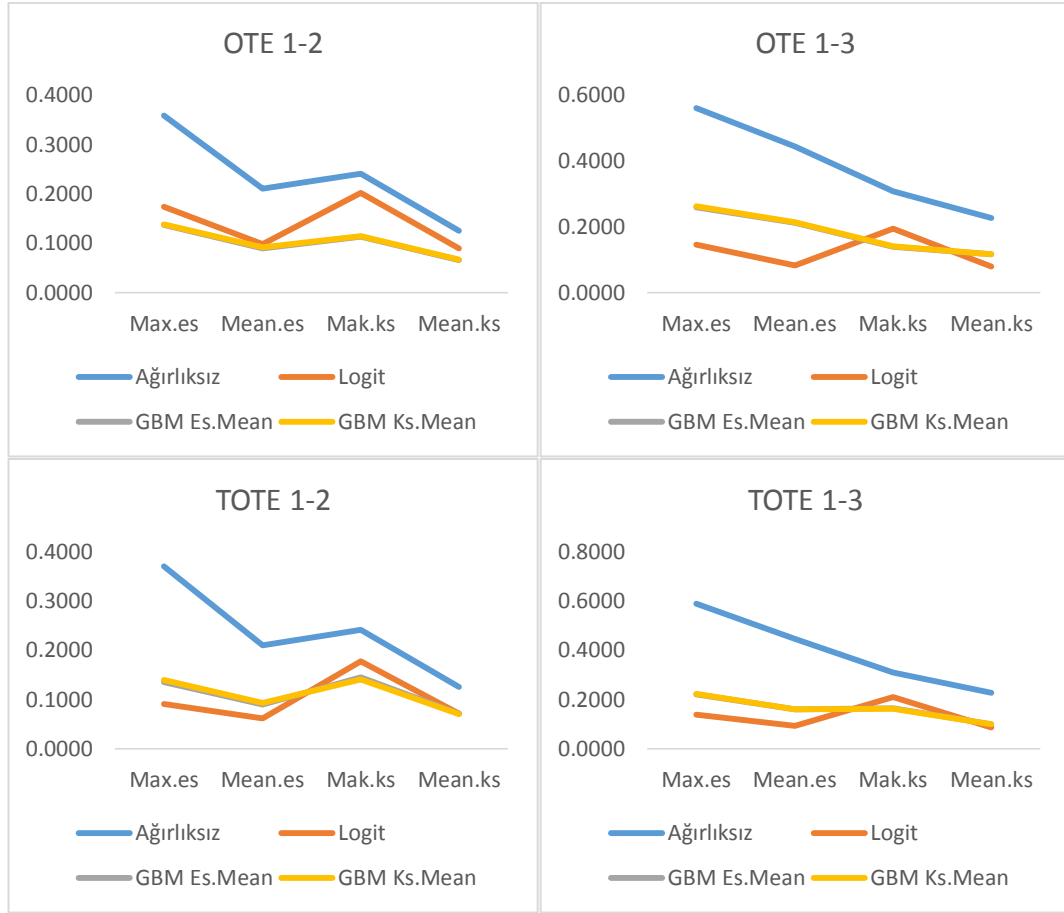


Şekil 5.1. Yöntemlere göre denge değerleri (S1, n=500)

Mak.es, means.es, mak.ks ve mean.ks değerlerinin ağırlıklandırma sonrasında yaklaşık olarak 0.20'nin altına düştüğü görülmektedir. Bu düşüş MLR'de mak.es ve mean.es için genel olarak daha çoktur. Dağılımları karşılaştırmaya yönelik KS istatistiği değerleri kullanıldığından GBM'den elde edilen denge skorları MLR'ye göre daha yakın çıkmıştır. Yani ortak değişkenlerden elde edilen standartlaştırılmış yanlılık ve Kolmogorov Smirnov istatistiği denge skorlarına ilişkin bu istatistiklerin en büyük değerler ve ortalama değerleri GBM için daha benzerdir. GBM için elde edilen grafikteki düzgünlük bunu ifade etmektedir (Tablo 5.1 ve Şekil 5.1).

Senaryo 2, n=500**Ana etkiler ve bir karesel terimin olduğu model****Tablo 5.2. MLR ve GBM’ye ilişkin özet denge değerleri (Senaryo 2; n=500)**

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	176,985	92,565	176,985	92,565	0,3585	0,2108	0,2409	0,1254
	1-3	176,985	230,450	176,985	230,450	0,5612	0,4438	0,3080	0,2258
Ağırlıksız TOTE	1-2	176,985	92,565	176,985	92,565	0,3691	0,2094	0,2409	0,1254
	1-3	176,985	230,450	176,985	230,450	0,5868	0,4445	0,3080	0,2258
MLR OTE	1-2	176,985	92,565	81,536	69,142	0,1736	0,0984	0,2023	0,0891
	1-3	176,985	230,450	81,536	138,771	0,1454	0,0824	0,1946	0,0784
MLR TOTE	1-2	176,985	92,565	176,985	62,490	0,0908	0,0609	0,1767	0,0709
	1-3	176,985	230,450	176,985	58,407	0,1364	0,0920	0,2084	0,0863
GBM Mean.es OTE	1-2	176,985	92,565	130,921	75,622	0,1374	0,0894	0,1132	0,0659
	1-3	176,985	230,450	130,921	188,566	0,2595	0,2125	0,1386	0,1161
GBM Mean.ks OTE	1-2	176,985	92,565	130,941	75,155	0,1378	0,0910	0,1136	0,0660
	1-3	176,985	230,450	130,941	189,309	0,2624	0,2139	0,1396	0,1164
GBM Mean.es TOTE	1-2	176,985	92,565	176,985	53,883	0,1350	0,0894	0,1446	0,0714
	1-3	176,985	230,450	176,985	76,606	0,2187	0,1582	0,1644	0,0988
GBM Mean.ks TOTE	1-2	176,985	92,565	176,985	55,736	0,1393	0,0922	0,1396	0,0700
	1-3	176,985	230,450	176,985	78,503	0,2208	0,1597	0,1616	0,0982

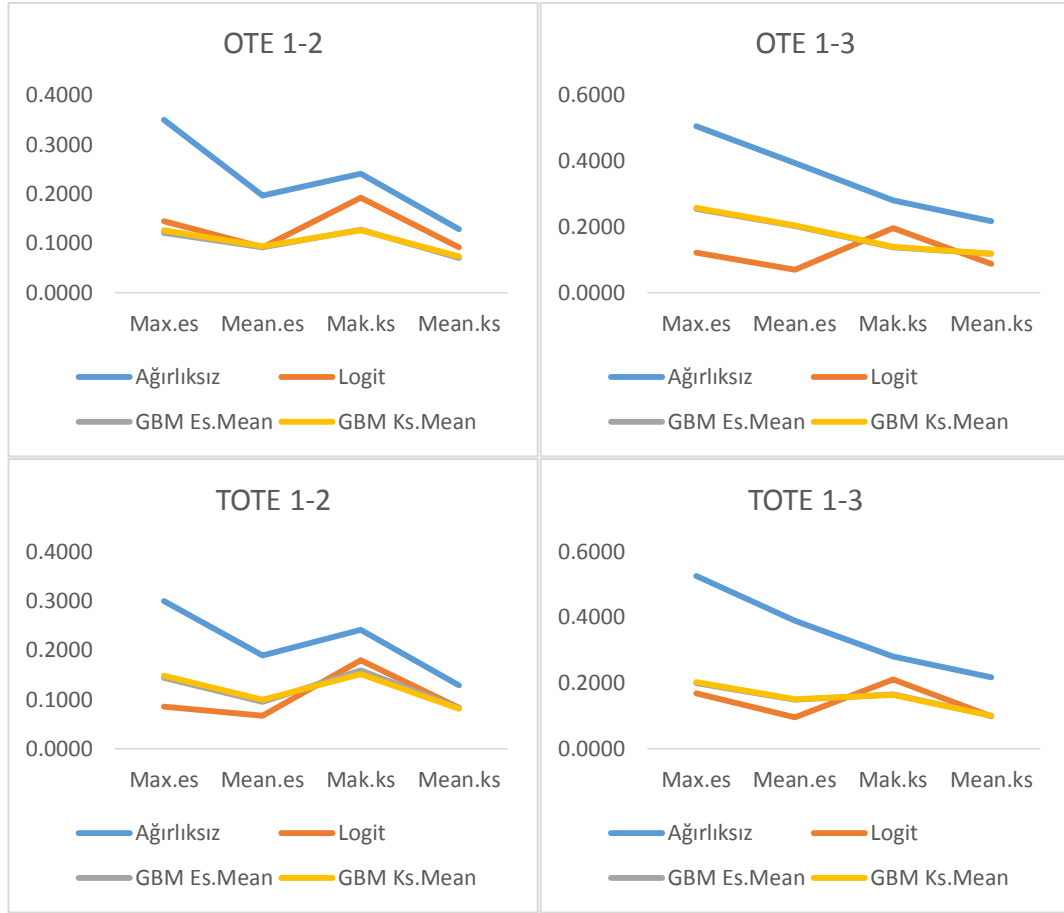


Şekil 5.2. Yöntemlere göre denge değerleri (S2, n=500)

Mak.es, means.es, mak.ks ve mean.ks değerlerinde ağırlıklandırma sonrasında düşüş olmaktadır. Bu düşüş OTE kestirimi için her bir yöntemde 0.20'nin altında iken ve 1-2 tedavi grupları için GBM'deki performansın daha iyi olduğu gözükmemektedir. Diğer ikili karşılaştırmalarda MLR'de, max.es ve mean.es için daha fazla düşüş gözlenirken, mak.ks ve mean.ks için GBM daha düşük denge sonuçları üretmektedir. Genel olarak GBM'deki denge değerleri MLR'ye göre daha benzerdir (Tablo 5.2 ve Şekil 5.2).

Senaryo 3, n=500**Ana etkiler ve üç karesel terimin olduğu model****Tablo 5.3. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 3; n=500)**

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	202,059	69,558	202,059	69,558	0,3504	0,1967	0,2402	0,1282
	1-3	202,059	228,383	202,059	228,383	0,5058	0,3938	0,2805	0,2160
Ağırlıksız TOTE	1-2	202,059	69,558	202,059	69,558	0,2994	0,1888	0,2402	0,1282
	1-3	202,059	228,383	202,059	228,383	0,5250	0,3886	0,2805	0,2160
MLR OTE	1-2	202,059	69,558	106,802	55,006	0,1441	0,0913	0,1918	0,0917
	1-3	202,059	228,383	106,802	152,168	0,1202	0,0696	0,1963	0,0863
MLR TOTE	1-2	202,059	69,558	202,059	50,897	0,0848	0,0665	0,1793	0,0827
	1-3	202,059	228,383	202,059	77,413	0,1686	0,0942	0,2109	0,0977
GBM Mean.es OTE	1-2	202,059	69,558	151,093	55,235	0,1205	0,0914	0,1270	0,0699
	1-3	202,059	228,383	151,093	186,287	0,2538	0,2024	0,1367	0,1179
GBM Mean.ks OTE	1-2	202,059	69,558	150,413	55,112	0,1263	0,0940	0,1256	0,0730
	1-3	202,059	228,383	150,413	186,537	0,2574	0,2040	0,1382	0,1181
GBM Mean.es TOTE	1-2	202,059	69,558	202,059	41,647	0,1429	0,0951	0,1588	0,0823
	1-3	202,059	228,383	202,059	74,778	0,1992	0,1484	0,1653	0,1005
GBM Mean.ks TOTE	1-2	202,059	69,558	202,059	43,018	0,1482	0,0991	0,1512	0,0802
	1-3	202,059	228,383	202,059	76,638	0,2025	0,1501	0,1624	0,0997



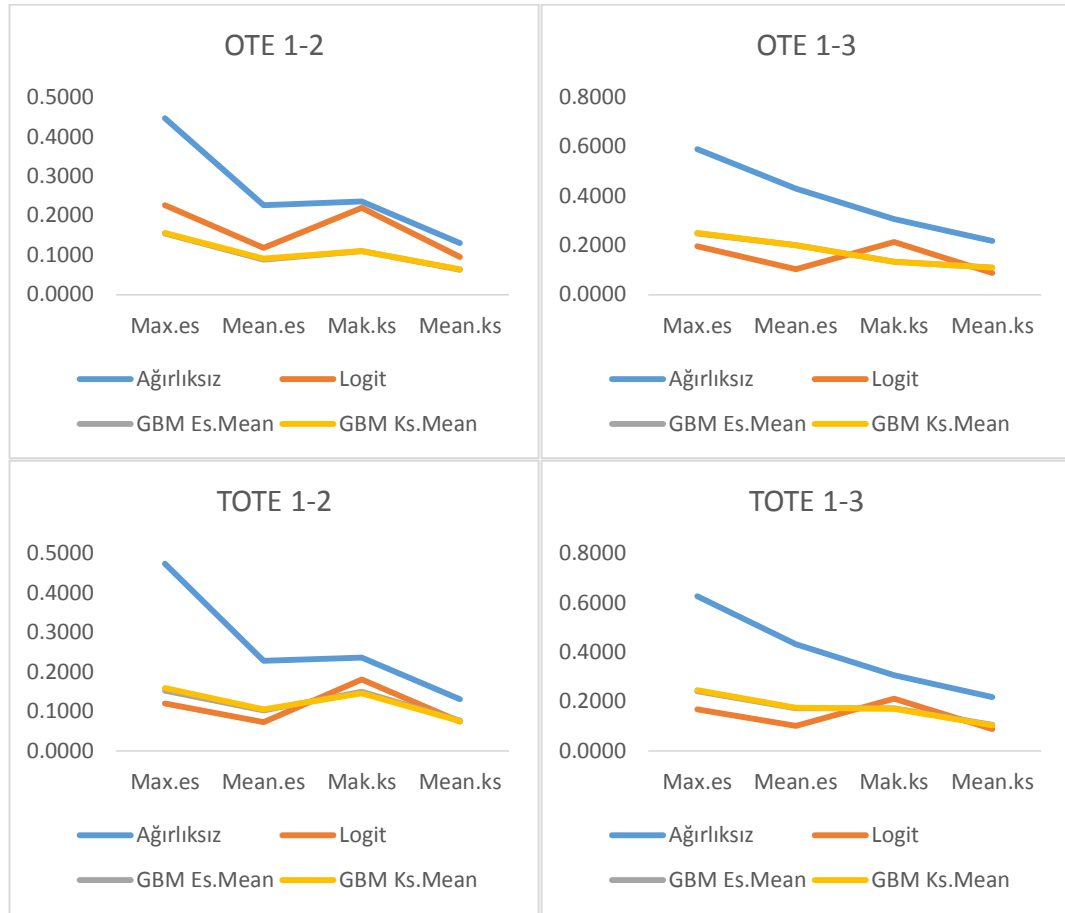
Şekil 5.3. Yöntemlere göre denge değerleri (S3, n=500)

Mak.es, means.es, mak.ks ve mean.ks değerleri ağırlıklandırma sonrasında düşüş olmaktadır. Bu düşüş OTE kestirimi için her bir yöntem için 0.20'nin altında iken ve GBM'deki performansın daha iyi olduğu gözükmemektedir. Diğer ikili karşılaştırmalarda MLR'de, max.es ve mean.es için daha fazla düşüş gözlenirken, mak.ks ve mean.ks için GBM daha iyi sonuç üretmektedir. Genel olarak GBM'deki denge değerleri MLR'ye göre daha benzerdir (Tablo 5.3 ve Şekil 5.3).

Senaryo 4, n=500**Ana etkiler ve üç tane iki yönlü etkileşim terimin olduğu model**

Tablo 5.4. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 4; n=500)

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	185,139	96,959	185,139	96,959	0,4473	0,2262	0,2356	0,1306
	1-3	185,139	217,902	185,139	217,902	0,5905	0,4290	0,3070	0,2177
Ağırlıksız TOTE	1-2	185,139	96,959	185,139	96,959	0,4744	0,2275	0,2356	0,1306
	1-3	185,139	217,902	185,139	217,902	0,6267	0,4322	0,3070	0,2177
MLR OTE	1-2	185,139	96,959	76,335	77,141	0,2270	0,1177	0,2204	0,0958
	1-3	185,139	217,902	76,335	141,444	0,1955	0,1025	0,2125	0,0869
MLR TOTE	1-2	185,139	96,959	185,139	62,380	0,1198	0,0725	0,1800	0,0742
	1-3	185,139	217,902	185,139	63,921	0,1682	0,1004	0,2116	0,0875
GBM Mean.es OTE	1-2	185,139	96,959	136,973	80,203	0,1548	0,0892	0,1104	0,0622
	1-3	185,139	217,902	136,973	179,164	0,2475	0,1998	0,1323	0,1089
GBM Mean.ks OTE	1-2	185,139	96,959	136,658	79,765	0,1561	0,0913	0,1096	0,0648
	1-3	185,139	217,902	136,658	179,812	0,2498	0,2011	0,1336	0,1092
GBM Mean.es TOTE	1-2	185,139	96,959	185,139	53,659	0,1518	0,1014	0,1489	0,0762
	1-3	185,139	217,902	185,139	75,293	0,2422	0,1721	0,1725	0,1041
GBM Mean.ks TOTE	1-2	185,139	96,959	185,139	56,032	0,1587	0,1044	0,1450	0,0749
	1-3	185,139	217,902	185,139	77,711	0,2455	0,1738	0,1693	0,1034

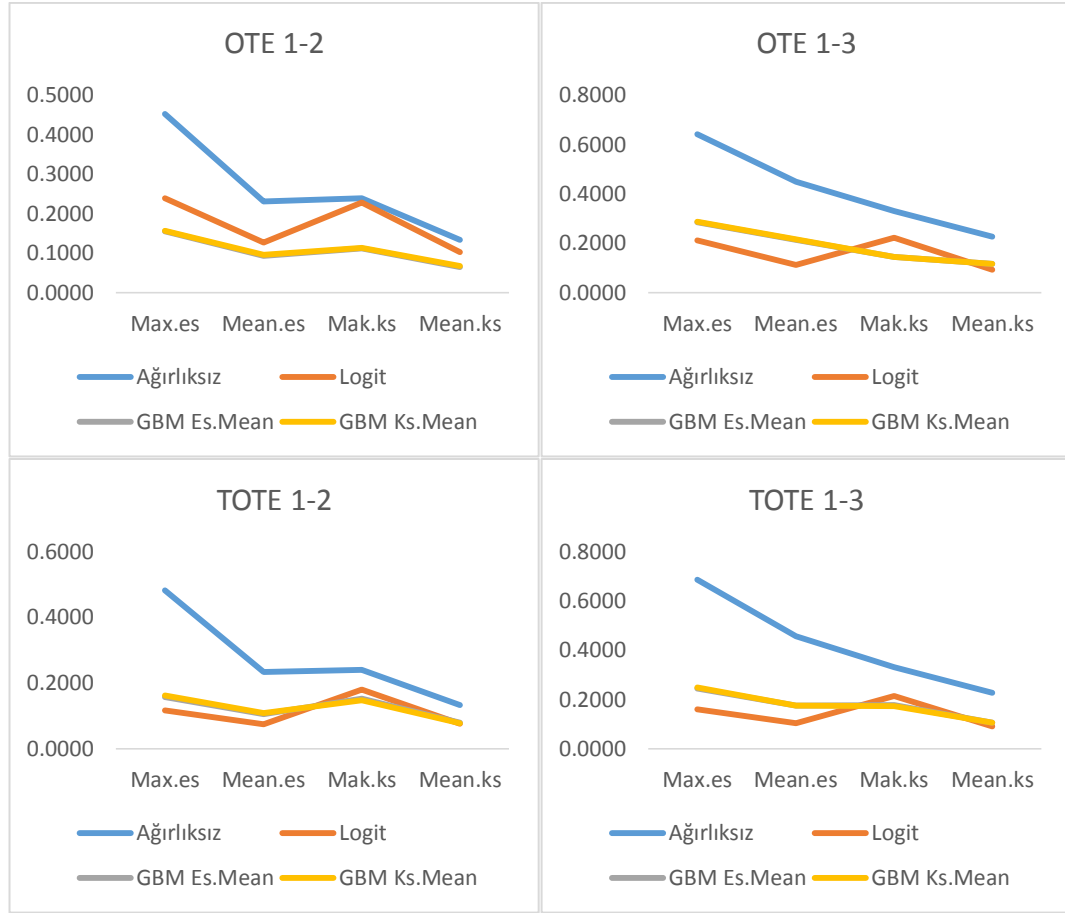


Şekil 5.4. Yöntemlere göre denge değerleri (S4, n=500)

Mak.es, means.es, mak.ks ve mean.ks değerlerinde ağırlıklandırma sonrasında düşüş olmaktadır. Bu düşüş GBM’de 1 ve 2 grupları için en alt düzeyde iken MLR’de nispeten daha fazla gözlenmektedir. Diğer ikili karşılaştırmalarda MLR’de, max.es ve mean.es için daha fazla düşüş gözlenirken, mak.ks ve mean.ks için GBM daha iyi sonuç üretmektedir. Genel olarak GBM’deki denge değerleri MLR’ye göre daha benzerdir (Tablo 5.4 ve Şekil 5.4).

Senaryo 5, n=500**Ana etkiler, dört tane iki yönlü etkileşim ve bir karesel terimin olduğu model****Tablo 5.5. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 5; n=500)**

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	177,276	94,584	177,276	94,584	0,4519	0,2308	0,2388	0,1327
	1-3	177,276	228,140	177,276	228,140	0,6424	0,4488	0,3297	0,2271
Ağırlıksız TOTE	1-2	177,276	94,584	177,276	94,584	0,4812	0,2333	0,2388	0,1327
	1-3	177,276	228,140	177,276	228,140	0,6853	0,4545	0,3297	0,2271
MLR OTE	1-2	177,276	94,584	65,688	73,152	0,2393	0,1262	0,2283	0,1017
	1-3	177,276	228,140	65,688	140,005	0,2112	0,1117	0,2208	0,0929
MLR TOTE	1-2	177,276	94,584	177,276	59,115	0,1160	0,0740	0,1801	0,0760
	1-3	177,276	228,140	177,276	56,767	0,1598	0,1024	0,2136	0,0900
GBM Mean.es OTE	1-2	177,276	94,584	126,806	77,716	0,1551	0,0927	0,1116	0,0642
	1-3	177,276	228,140	126,806	185,382	0,2844	0,2133	0,1436	0,1161
GBM Mean.ks OTE	1-2	177,276	94,584	126,738	77,308	0,1563	0,0950	0,1123	0,0672
	1-3	177,276	228,140	126,738	186,224	0,2878	0,2146	0,1447	0,1164
GBM Mean.es TOTE	1-2	177,276	94,584	177,276	51,848	0,1560	0,1050	0,1516	0,0779
	1-3	177,276	228,140	177,276	71,060	0,2444	0,1734	0,1757	0,1047
GBM Mean.ks TOTE	1-2	177,276	94,584	177,276	54,233	0,1615	0,1080	0,1466	0,0764
	1-3	177,276	228,140	177,276	73,170	0,2470	0,1751	0,1726	0,1040

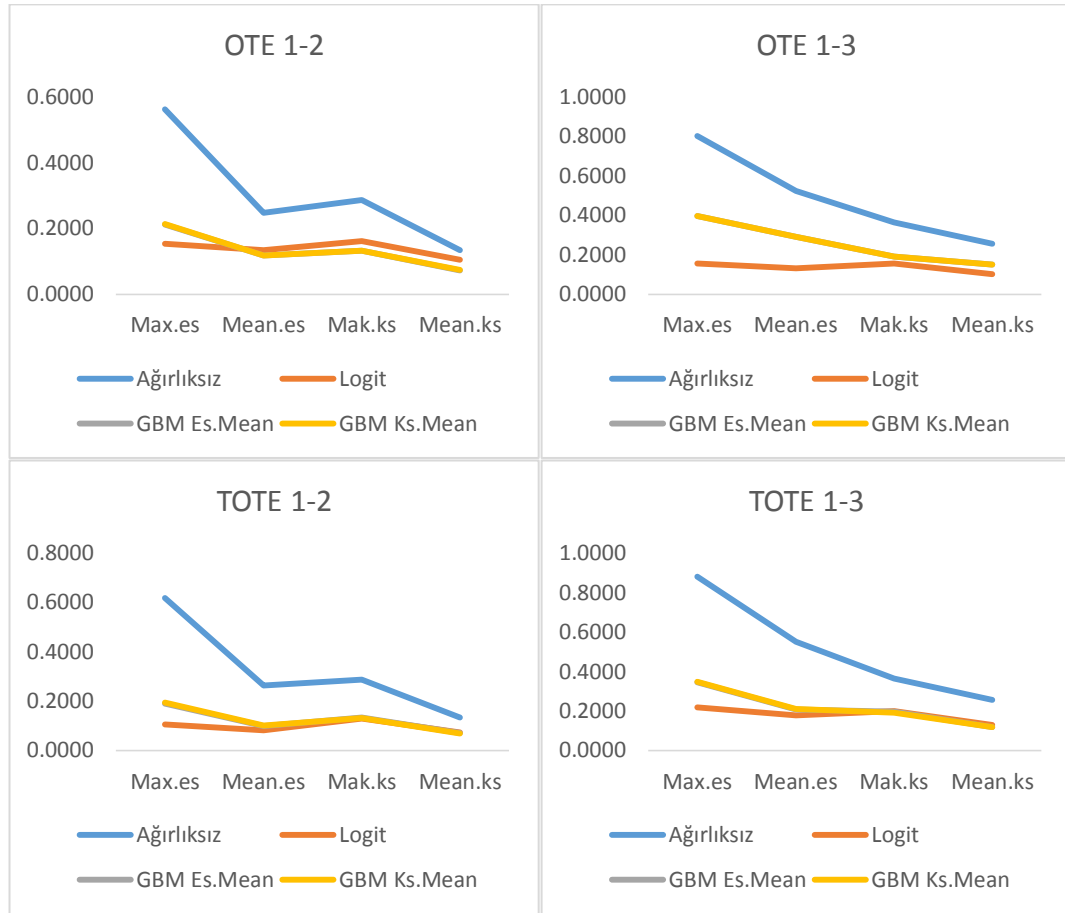


Şekil 5.5. Yöntemlere göre denge değerleri (S5, n=500)

Mak.es, means.es, mak.ks ve mean.ks değerleri ağırlıklandırma sonrasında düşüş olmaktadır. Bu düşüş GBM’de 1 ve 2 grupları için en alt düzeyde iken lojistikte nispeten daha fazla gözlenmektedir. Diğer ikili karşılaştırmalarda MLR, max.es ve mean.es için daha fazla düşüş gözlenirken, mak.ks ve mean.ks için GBM daha iyi sonuç üretmektedir. Genel olarak GBM’deki denge değerleri MLR’ye göre daha benzerdir (Tablo 5.5 ve Şekil 5.5).

Senaryo 6, n=500**Ana etkiler, on tane iki yönlü etkileşim terimin olduğu model****Tablo 5.6. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 6; n=500)**

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	167,195	104,802	167,195	104,802	0,5628	0,2479	0,2870	0,1340
	1-3	167,195	228,003	167,195	228,003	0,8015	0,5228	0,3642	0,2553
Ağırlıksız TOTE	1-2	167,195	104,802	167,195	104,802	0,6184	0,2630	0,2870	0,1340
	1-3	167,195	228,003	167,195	228,003	0,8808	0,5499	0,3642	0,2553
MLR OTE	1-2	167,195	104,802	55,944	64,826	0,1538	0,1335	0,1616	0,1043
	1-3	167,195	228,003	55,944	89,535	0,1561	0,1324	0,1555	0,1008
MLR TOTE	1-2	167,195	104,802	167,195	51,684	0,1042	0,0804	0,1284	0,0730
	1-3	167,195	228,003	167,195	26,169	0,2168	0,1781	0,1999	0,1297
GBM Mean.es OTE	1-2	167,195	104,802	118,260	85,007	0,2126	0,1175	0,1320	0,0727
	1-3	167,195	228,003	118,260	173,181	0,3953	0,2898	0,1910	0,1489
GBM Mean.ks OTE	1-2	167,195	104,802	118,894	84,531	0,2137	0,1181	0,1318	0,0744
	1-3	167,195	228,003	118,894	174,250	0,3970	0,2912	0,1913	0,1493
GBM Mean.es TOTE	1-2	167,195	104,802	167,195	54,848	0,1893	0,0988	0,1329	0,0694
	1-3	167,195	228,003	167,195	54,420	0,3438	0,2059	0,1949	0,1187
GBM Mean.ks TOTE	1-2	167,195	104,802	167,195	56,846	0,1947	0,1012	0,1308	0,0682
	1-3	167,195	228,003	167,195	56,819	0,3464	0,2085	0,1915	0,1176

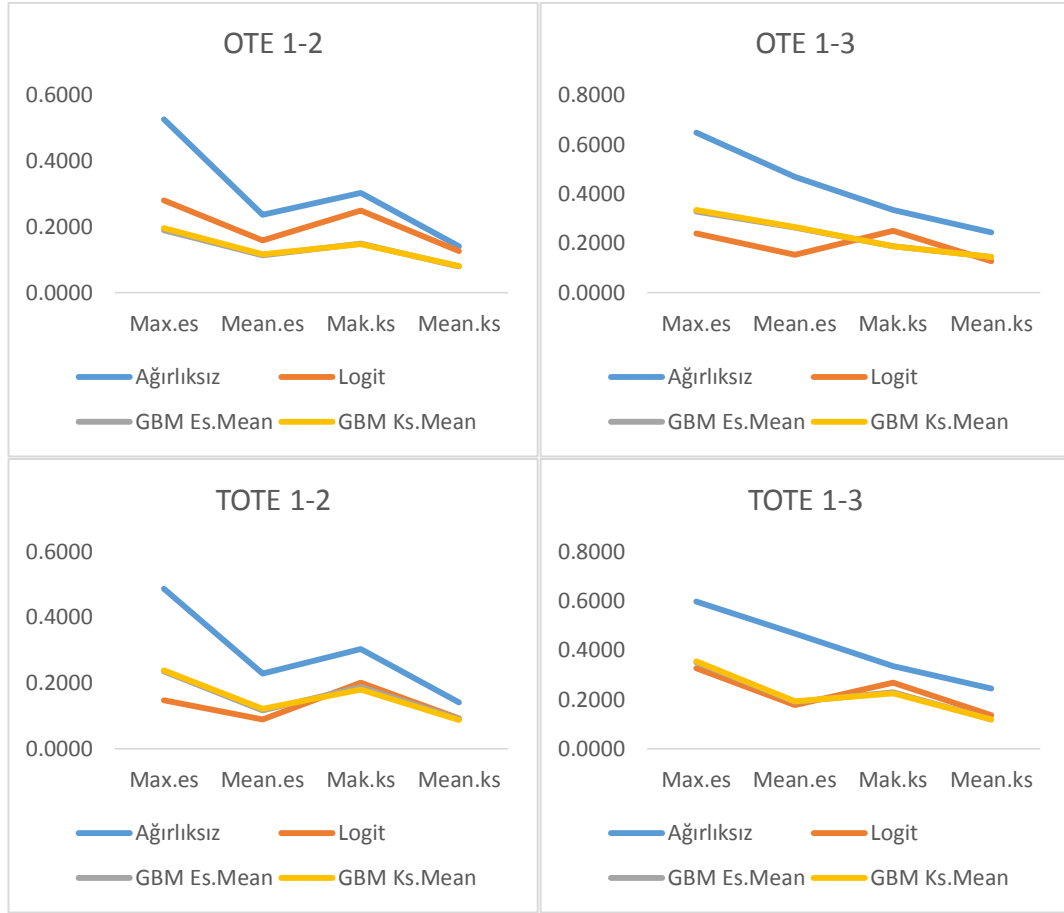


Şekil 5.6. Yöntemlere göre denge değerleri (S6, n=500)

Mak.es, means.es, mak.ks ve mean.ks değerleri ağırlıklandırma sonrasında düşüş olmaktadır. Bu düşüş GBM’de 1 ve 2 grupları için en alt düzeyde iken lojistikte nispeten daha fazla gözlenmektedir. Diğer ikili karşılaştırmalarda MLR’de, max.es ve mean.es için daha fazla düşüş gözlenirken, mak.ks ve mean.ks için GBM daha iyi sonuç üretmektedir. Genel olarak GBM’deki denge değerleri MLR’ye göre daha benzerdir (Tablo 5.6 ve Şekil 5.6).

Senaryo 7, n=500**Ana etkiler, on tane iki yönlü etkileşim, üç tane karesel terimin olduğu model****Tablo 5.7. MLR ve GBM’ye ilişkin özet denge değerleri (Senaryo 7; n=500)**

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	217,621	73,686	217,621	73,686	0,5268	0,2356	0,3020	0,1412
	1-3	217,621	208,693	217,621	208,693	0,6472	0,4685	0,3352	0,2425
Ağırlıksız TOTE	1-2	217,621	73,686	217,621	73,686	0,4853	0,2288	0,3020	0,1412
	1-3	217,621	208,693	217,621	208,693	0,5961	0,4671	0,3352	0,2425
MLR OTE	1-2	217,621	73,686	63,386	53,268	0,2804	0,1575	0,2490	0,1257
	1-3	217,621	208,693	63,386	90,214	0,2398	0,1515	0,2504	0,1267
MLR TOTE	1-2	217,621	73,686	217,621	46,776	0,1467	0,0891	0,2005	0,0914
	1-3	217,621	208,693	217,621	36,999	0,3257	0,1758	0,2670	0,1361
GBM Mean.es OTE	1-2	217,621	73,686	162,264	57,537	0,1886	0,1130	0,1481	0,0787
	1-3	217,621	208,693	162,264	157,064	0,3284	0,2630	0,1863	0,1437
GBM Mean.ks OTE	1-2	217,621	73,686	161,581	57,723	0,1964	0,1154	0,1475	0,0808
	1-3	217,621	208,693	161,581	158,043	0,3350	0,2652	0,1876	0,1440
GBM Mean.es TOTE	1-2	217,621	73,686	217,621	40,173	0,2341	0,1168	0,1872	0,0895
	1-3	217,621	208,693	217,621	56,314	0,3500	0,1885	0,2281	0,1187
GBM Mean.ks TOTE	1-2	217,621	73,686	217,621	42,027	0,2375	0,1205	0,1797	0,0874
	1-3	217,621	208,693	217,621	58,868	0,3535	0,1909	0,2245	0,1176



Şekil 5.7. Yöntemlere göre denge değerleri (S7, n=500)

Mak.es, means.es, mak.ks ve mean.ks değerleri ağırlıklandırma sonrasında düşüş olmaktadır. Bu düşüş GBM’de 1 ve 2 grupları için en alt düzeyde iken lojistikte nispeten daha fazla gözlenmektedir. Diğer ikili karşılaştırmalarda lojistik regresyonda, max.es ve mean.es için daha fazla düşüş gözlenirken, mak.ks ve mean.ks için GBM daha iyi sonuç üretmektedir. Genel olarak GBM’deki denge değerlerindeki sapma MLR’ye göre daha azdır (Tablo 5.7 ve Şekil 5.7).

Senaryolar için genel değerlendirme yapmak gerekirse, elde edilen standartlaştırılmış yanlılık ve kolmogorov smirnov istatistiği denge değerleri için bu istatistiklere ilişkin değerler ve ortalama değerleri GBM için daha benzerdir. Ayrıca tüm senaryolar için ağırlıklandırma öncesi denge skorları tüm senaryolar için yüksektir.

5.1.2 n=1000 için Senaryoların Denge Değerlendirmesi

Senaryo 1, n=1000

Ana etkilerin olduğu model

Tablo 5.8. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 1; n=1000)

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	278,425	204,769	278,425	204,769	0,3824	0,2020	0,1911	0,1072
	1-3	278,425	516,806	278,425	516,806	0,5999	0,4655	0,2911	0,2267
Ağırlıksız TOTE	1-2	278,425	204,769	278,425	204,769	0,3943	0,2105	0,1911	0,1072
	1-3	278,425	516,806	278,425	516,806	0,6294	0,4841	0,2911	0,2267
MLR OTE	1-2	278,425	204,769	120,869	149,514	0,0777	0,0696	0,0968	0,0601
	1-3	278,425	516,806	120,869	316,339	0,0670	0,0583	0,0843	0,0512
MLR TOTE	1-2	278,425	204,769	278,425	161,459	0,0436	0,0371	0,0766	0,0420
	1-3	278,425	516,806	278,425	192,141	0,0758	0,0633	0,0932	0,0553
GBM Mean.es OTE	1-2	278,425	204,769	190,580	165,830	0,1171	0,0718	0,0822	0,0463
	1-3	278,425	516,806	190,580	411,721	0,2168	0,1790	0,1084	0,0939
GBM Mean.ks OTE	1-2	278,425	204,769	192,011	166,075	0,1167	0,0717	0,0813	0,0478
	1-3	278,425	516,806	192,011	412,663	0,2186	0,1800	0,1093	0,0942
GBM Mean.es TOTE	1-2	278,425	204,769	278,425	124,505	0,0906	0,0584	0,0759	0,0437
	1-3	278,425	516,806	278,425	137,689	0,1334	0,1133	0,0916	0,0683
GBM Mean.ks TOTE	1-2	278,425	204,769	278,425	128,151	0,0912	0,0598	0,0733	0,0430
	1-3	278,425	516,806	278,425	140,785	0,1344	0,1144	0,0895	0,0678

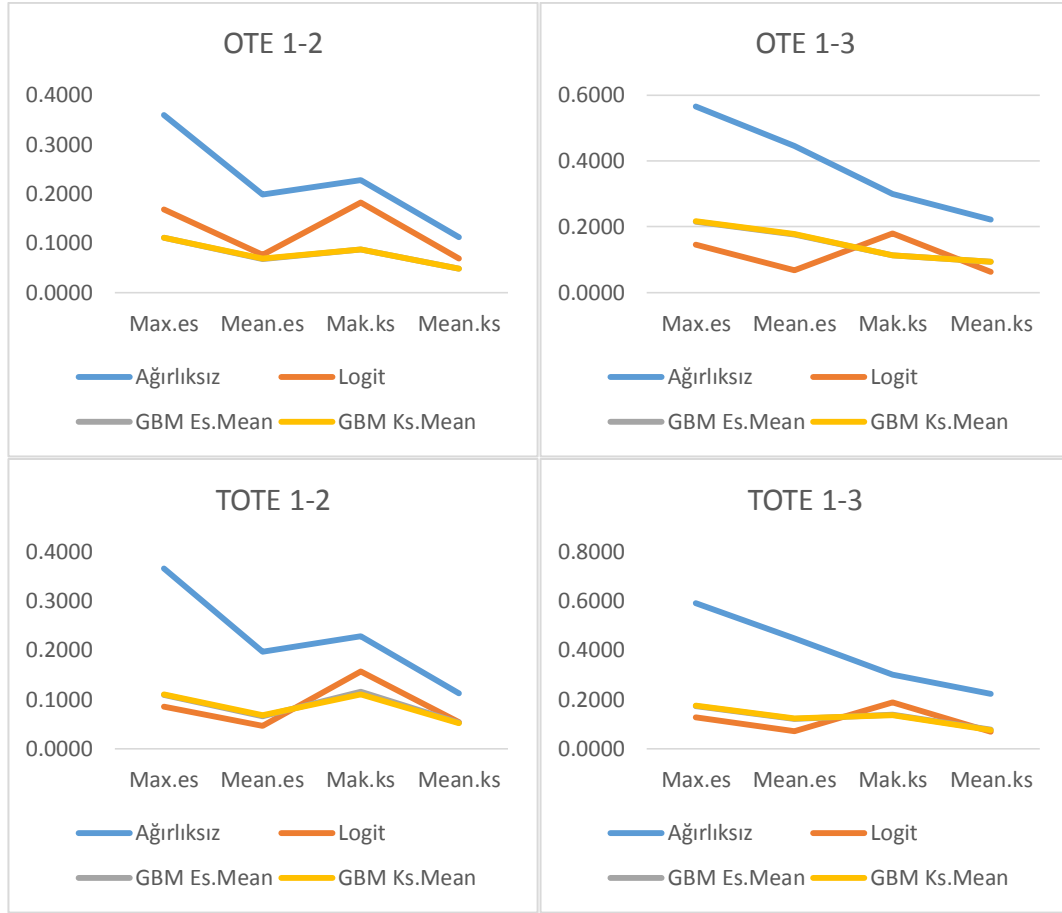


Şekil 5.8. Yöntemlere göre denge değerleri (S1, n=1000)

Mak.es, means.es, mak.ks ve mean.ks değerlerinin ağırlıklandırma sonrasında 500 örneklemlili senaryoya göre daha da düşük olduğu görülmektedir. Bu düşüş MLR için mak.es ve mean.es için genel olarak daha çoktur. Yine OTE 1-3 için MLR daha iyi denge yaptığı gözükmemektedir. Dağılımları KS istatistiği değerleriyle karşılaştırdığımızda GBM'den elde edilen denge skorları MLR'ye göre daha benzerdir (Tablo 5.8 ve Şekil 5.8).

Senaryo 2, n=1000**Ana etkiler ve bir karesel terimin olduğu model****Tablo 5.9. MLR ve GBM’ye ilişkin özet denge değerleri (Senaryo 2; n=1000)**

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	352,975	184,477	352,975	184,477	0,3597	0,1990	0,2277	0,1118
	1-3	352,975	462,548	352,975	462,548	0,5660	0,4459	0,3002	0,2217
Ağırlıksız TOTE	1-2	352,975	184,477	352,975	184,477	0,3649	0,1969	0,2277	0,1118
	1-3	352,975	462,548	352,975	462,548	0,5892	0,4458	0,3002	0,2217
MLR OTE	1-2	352,975	184,477	156,218	141,403	0,1683	0,0765	0,1819	0,0688
	1-3	352,975	462,548	156,218	275,336	0,1460	0,0667	0,1800	0,0627
MLR TOTE	1-2	352,975	184,477	352,975	147,049	0,0850	0,0456	0,1560	0,0540
	1-3	352,975	462,548	352,975	177,163	0,1275	0,0705	0,1877	0,0677
GBM Mean.es OTE	1-2	352,975	184,477	244,175	149,575	0,1105	0,0679	0,0867	0,0460
	1-3	352,975	462,548	244,175	356,374	0,2150	0,1761	0,1121	0,0930
GBM Mean.ks OTE	1-2	352,975	184,477	244,507	149,435	0,1113	0,0689	0,0870	0,0478
	1-3	352,975	462,548	244,507	357,885	0,2167	0,1771	0,1127	0,0933
GBM Mean.es TOTE	1-2	352,975	184,477	352,975	104,031	0,1089	0,0653	0,1151	0,0527
	1-3	352,975	462,548	352,975	126,860	0,1727	0,1199	0,1377	0,0762
GBM Mean.ks TOTE	1-2	352,975	184,477	352,975	108,413	0,1103	0,0673	0,1095	0,0515
	1-3	352,975	462,548	352,975	131,221	0,1733	0,1214	0,1345	0,0756

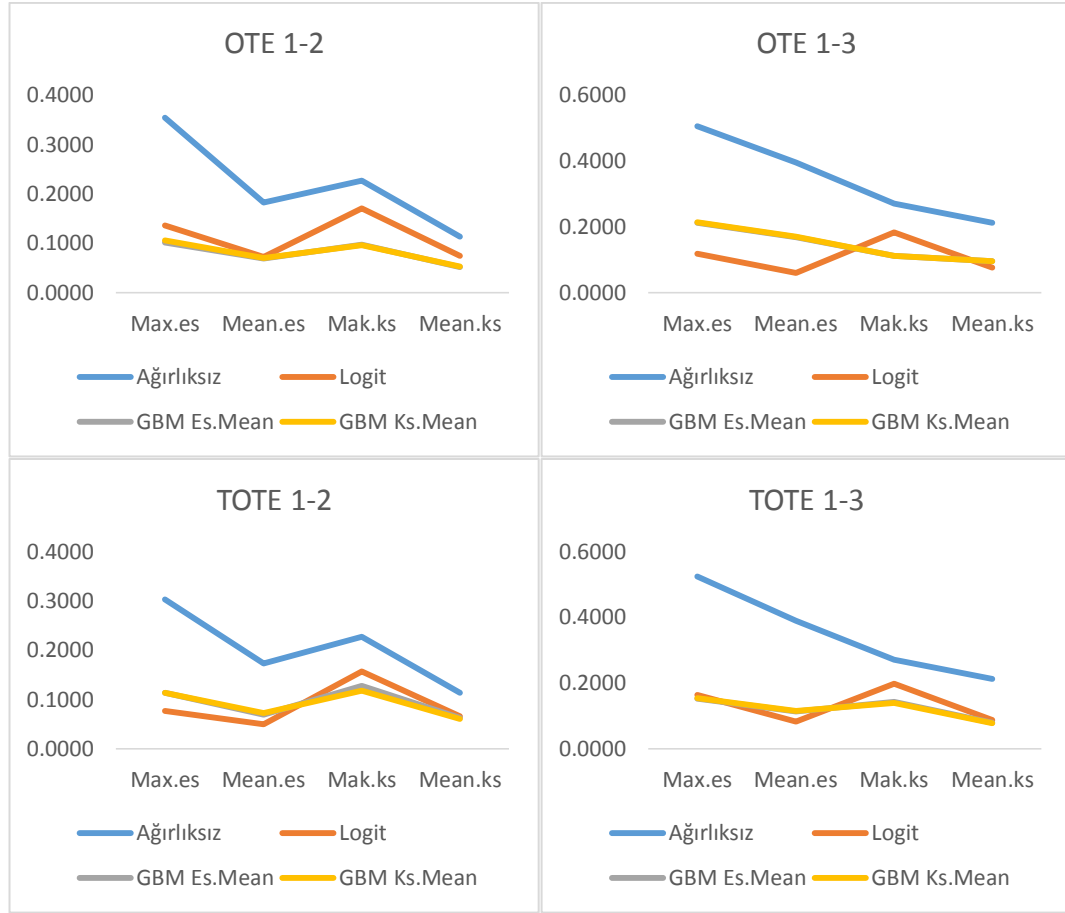


Şekil 5.9. Yöntemlere göre denge değerleri (S2, n=1000)

Ağırlıklandırma sonrası denge skorlarında düşüşler gözlenmektedir. Ayrıca bu düşüş yine 500 örneklemlili veriden daha çok olmaktadır. GBM'ye ilişkin denge değerlerindeki benzerlik MLR'ye göre daha fazladır (Tablo 5.9 ve Şekil 5.9).

Senaryo 3, n=1000**Ana etkiler ve üç karesel terimin olduğu model****Tablo 5.10. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 3; n=1000)**

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	404,212	139,251	404,212	139,251	0,3546	0,1822	0,2263	0,1134
	1-3	404,212	456,537	404,212	456,537	0,5049	0,3950	0,2697	0,2120
Ağırlıksız TOTE	1-2	404,212	139,251	404,212	139,251	0,3022	0,1731	0,2263	0,1134
	1-3	404,212	456,537	404,212	456,537	0,5231	0,3892	0,2697	0,2120
MLR OTE	1-2	404,212	139,251	209,697	114,963	0,1358	0,0724	0,1707	0,0738
	1-3	404,212	456,537	209,697	303,046	0,1173	0,0591	0,1832	0,0757
MLR TOTE	1-2	404,212	139,251	404,212	116,089	0,0766	0,0496	0,1569	0,0657
	1-3	404,212	456,537	404,212	212,056	0,1626	0,0826	0,1965	0,0865
GBM Mean.es OTE	1-2	404,212	139,251	279,323	108,879	0,1011	0,0685	0,0971	0,0511
	1-3	404,212	456,537	279,323	344,714	0,2113	0,1685	0,1105	0,0953
GBM Mean.ks OTE	1-2	404,212	139,251	278,837	110,303	0,1051	0,0703	0,0959	0,0528
	1-3	404,212	456,537	278,837	346,181	0,2133	0,1695	0,1111	0,0955
GBM Mean.es TOTE	1-2	404,212	139,251	404,212	77,694	0,1128	0,0687	0,1269	0,0621
	1-3	404,212	456,537	404,212	122,020	0,1518	0,1123	0,1418	0,0781
GBM Mean.ks TOTE	1-2	404,212	139,251	404,212	81,656	0,1133	0,0723	0,1175	0,0601
	1-3	404,212	456,537	404,212	126,526	0,1538	0,1140	0,1387	0,0774



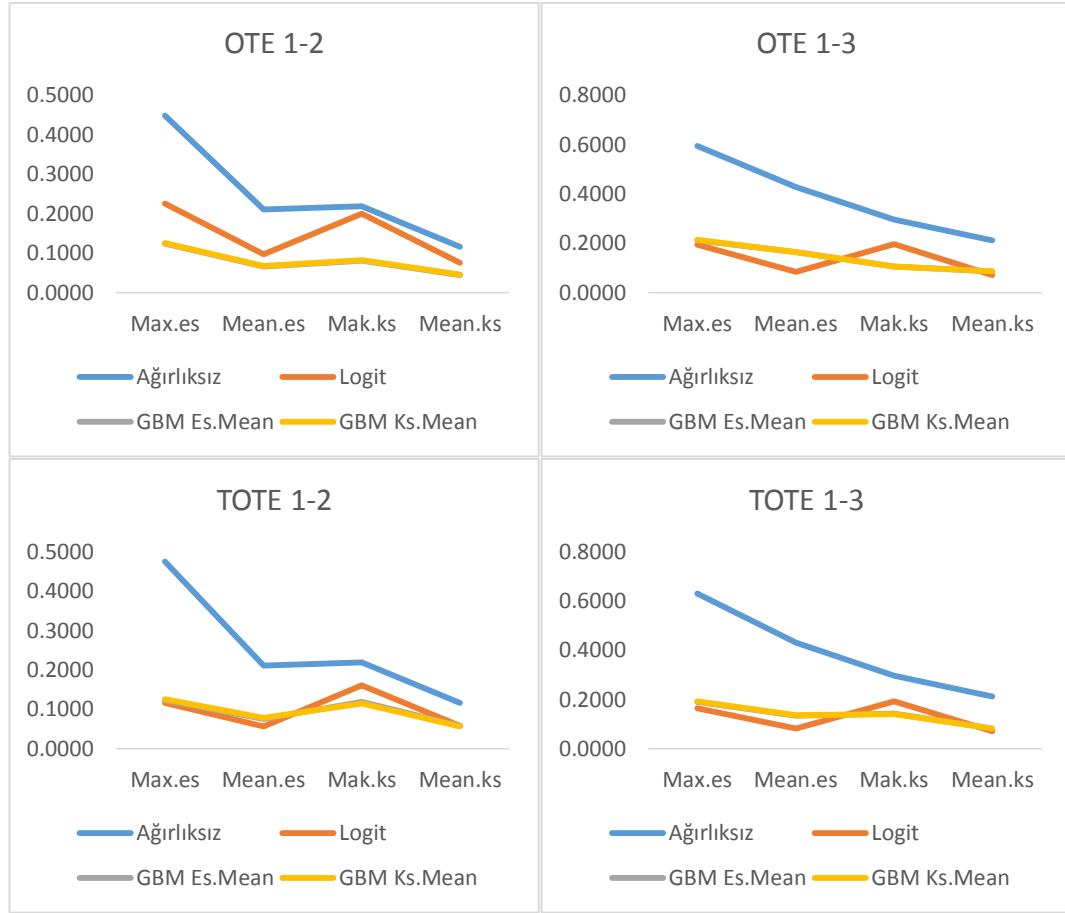
Şekil 5.10. Yöntemlere göre denge değerleri (S3, n=1000)

Ağırlıklandırma sonrası denge skorlarında düşüşler gözlenmektedir. Ayrıca bu düşüş yine 500 örneklemlili veriden daha çok olmaktadır. Özellikle OTE 1-3 grubu için MLR’de mak.es ve mean.es değerleri (0,1173 ve (0,0591) daha iyidir. GBM denge değerlerindeki benzerlik MLR’den daha fazladır (Tablo 5.10 ve Şekil 5.10).

Senaryo 4, n=1000**Ana etkiler ve üç tane iki yönlü etkileşim terimin olduğu model**

Tablo 5.11. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 4; n=1000)

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	369,169	194,926	369,169	194,926	0,4478	0,2097	0,2187	0,1155
	1-3	369,169	435,905	369,169	435,905	0,5935	0,4269	0,2946	0,2116
Ağırlıksız TOTE	1-2	369,169	194,926	369,169	194,926	0,4738	0,2102	0,2187	0,1155
	1-3	369,169	435,905	369,169	435,905	0,6282	0,4288	0,2946	0,2116
MLR OTE	1-2	369,169	194,926	146,419	159,661	0,2249	0,0972	0,1996	0,0757
	1-3	369,169	435,905	146,419	287,423	0,1926	0,0843	0,1950	0,0694
MLR TOTE	1-2	369,169	194,926	369,169	156,986	0,1151	0,0567	0,1606	0,0576
	1-3	369,169	435,905	369,169	189,025	0,1640	0,0817	0,1909	0,0699
GBM Mean.es OTE	1-2	369,169	194,926	259,799	160,752	0,1235	0,0652	0,0808	0,0439
	1-3	369,169	435,905	259,799	337,984	0,2105	0,1632	0,1052	0,0861
GBM Mean.ks OTE	1-2	369,169	194,926	259,476	160,924	0,1252	0,0667	0,0813	0,0457
	1-3	369,169	435,905	259,476	339,970	0,2123	0,1642	0,1058	0,0864
GBM Mean.es TOTE	1-2	369,169	194,926	369,169	103,969	0,1226	0,0756	0,1182	0,0568
	1-3	369,169	435,905	369,169	127,350	0,1892	0,1340	0,1409	0,0812
GBM Mean.ks TOTE	1-2	369,169	194,926	369,169	109,172	0,1249	0,0778	0,1142	0,0557
	1-3	369,169	435,905	369,169	131,399	0,1910	0,1353	0,1387	0,0806



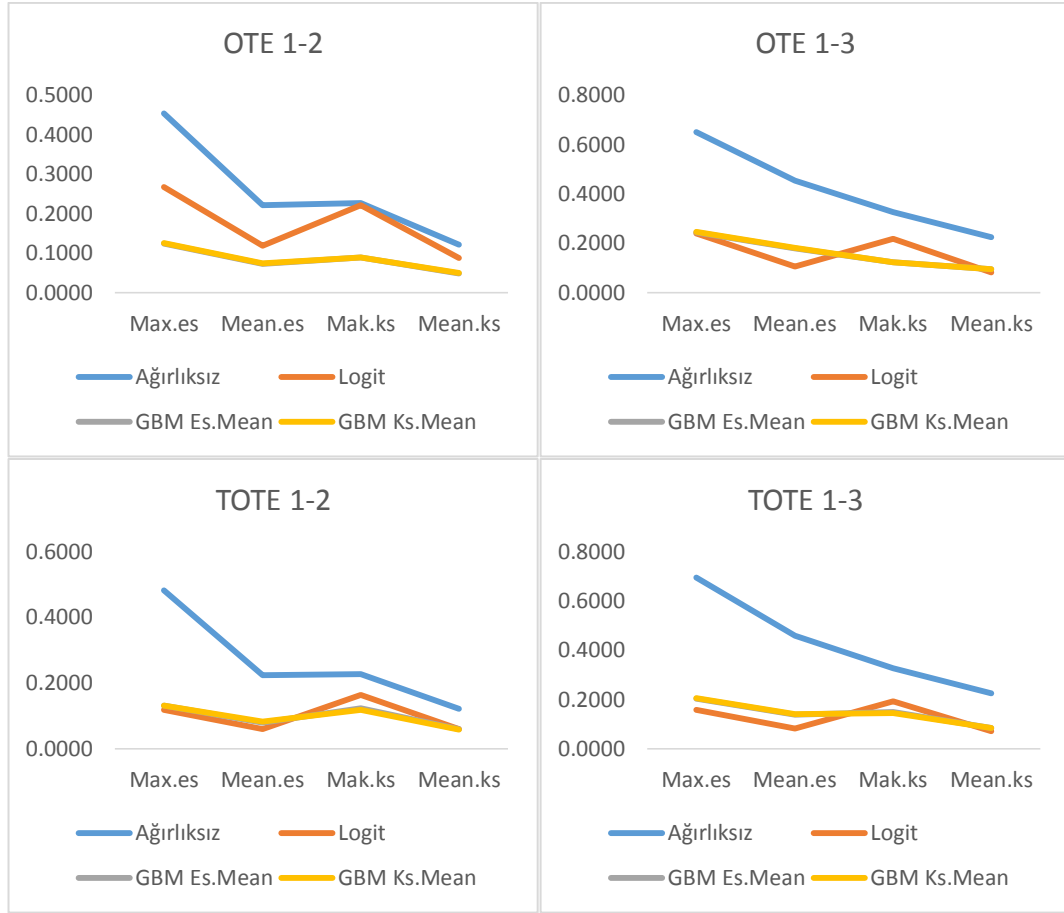
Şekil 5.11. Yöntemlere göre denge değerleri (S4, n=1000)

Ağırlıklandırma sonrası yine denge skorlarında düşüşler gözlenmektedir. Ayrıca bu düşüş yine 500 örneklemlili veriden daha çok olmaktadır. OTE 1-2 grubu için GBM’de mak.es ve mean.es (0,1252 ve 0,0667) çok daha iyi değerler vermektedir. GBM için denge değerlerindeki benzerlik MLR’den daha fazladır (Tablo 5.11 ve Şekil 5.11).

Senaryo 5, n=1000**Ana etkiler, dört tane iki yönlü etkileşim ve bir karesel terimin olduğu model**

Tablo 5.12. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 5; n=1000)

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	355,218	187,492	355,218	187,492	0,4542	0,2206	0,2270	0,1206
	1-3	355,218	457,290	355,218	457,290	0,6510	0,4532	0,3256	0,2240
Ağırlıksız TOTE	1-2	355,218	187,492	355,218	187,492	1,0003	0,7143	0,2270	0,1206
	1-3	355,218	457,290	355,218	457,290	1,0003	0,7143	0,3256	0,2240
MLR OTE	1-2	355,218	187,492	113,892	147,316	0,2666	0,1182	0,2210	0,0875
	1-3	355,218	457,290	113,892	281,142	0,2386	0,1058	0,2167	0,0814
MLR TOTE	1-2	355,218	187,492	355,218	148,756	0,1175	0,0589	0,1627	0,0592
	1-3	355,218	457,290	355,218	183,557	0,1577	0,0819	0,1925	0,0709
GBM Mean.es OTE	1-2	355,218	187,492	236,562	152,448	0,1243	0,0721	0,0880	0,0483
	1-3	355,218	457,290	236,562	349,749	0,2431	0,1794	0,1216	0,0946
GBM Mean.ks OTE	1-2	355,218	187,492	236,655	152,934	0,1253	0,0735	0,0883	0,0498
	1-3	355,218	457,290	236,655	351,630	0,2451	0,1804	0,1226	0,0948
GBM Mean.es TOTE	1-2	355,218	187,492	355,218	97,219	1,0003	0,7143	0,2040	0,1525
	1-3	355,218	457,290	355,218	117,179	1,0003	0,7143	0,0880	0,0497
GBM Mean.ks TOTE	1-2	355,218	187,492	355,218	102,679	1,0003	0,7143	0,1216	0,0946
	1-3	355,218	457,290	355,218	120,980	1,0003	0,7143	0,0826	0,0706



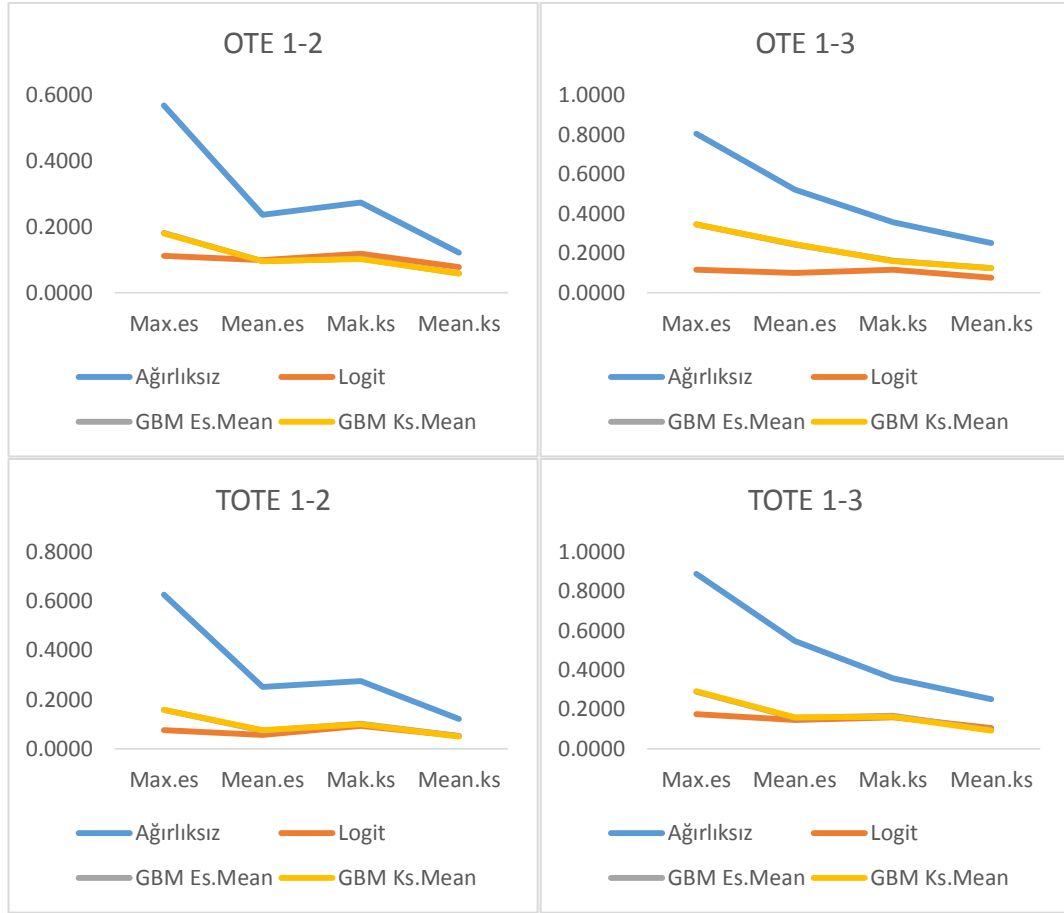
Şekil 5.12. Yöntemlere göre denge değerleri (S5, n=1000)

Ağırlıklandırma sonrası yine denge skorlarında düşüşler gözlenmektedir. OTE 1-2 grubu için GBM’de mak.es ve mean.es (0,1253 ve 0,0735) çok daha iyi değerler vermektedir. Aynı grupta MLR ağırlıklandırmadan sonra çok az bir düşüş sağlamaktadır. GBM denge değerlerindeki benzerlik MLR’den daha fazladır (Tablo 5.12 ve Şekil 5.12).

Senaryo 6, n=1000**Ana etkiler, on tane iki yönlü etkileşim terimin olduğu model**

Tablo 5.13. MLR ve GBM’ye ilişkin özet denge değerleri (Senaryo 6; n=1000)

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	333,435	209,368	333,435	209,368	0,5686	0,2358	0,2729	0,1209
	1-3	333,435	457,197	333,435	457,197	0,8059	0,5205	0,3553	0,2489
Ağırlıksız TOTE	1-2	333,435	209,368	333,435	209,368	0,6242	0,2500	0,2729	0,1209
	1-3	333,435	457,197	333,435	457,197	0,8847	0,5462	0,3553	0,2489
MLR OTE	1-2	333,435	209,368	99,280	130,600	0,1110	0,0981	0,1171	0,0767
	1-3	333,435	457,197	99,280	165,593	0,1161	0,0979	0,1143	0,0747
MLR TOTE	1-2	333,435	209,368	333,435	150,375	0,0745	0,0562	0,0910	0,0513
	1-3	333,435	457,197	333,435	106,322	0,1737	0,1453	0,1581	0,1042
GBM Mean.es OTE	1-2	333,435	209,368	214,730	167,479	0,1812	0,0957	0,1037	0,0572
	1-3	333,435	457,197	214,730	306,156	0,3455	0,2430	0,1597	0,1219
GBM Mean.ks OTE	1-2	333,435	209,368	216,885	167,176	0,1796	0,0957	0,1022	0,0574
	1-3	333,435	457,197	216,885	308,222	0,3461	0,2442	0,1595	0,1222
GBM Mean.es TOTE	1-2	333,435	209,368	333,435	107,758	0,1563	0,0727	0,1016	0,0502
	1-3	333,435	457,197	333,435	77,698	0,2878	0,1564	0,1658	0,0929
GBM Mean.ks TOTE	1-2	333,435	209,368	333,435	111,421	0,1564	0,0743	0,0988	0,0493
	1-3	333,435	457,197	333,435	83,153	0,2918	0,1595	0,1616	0,0916



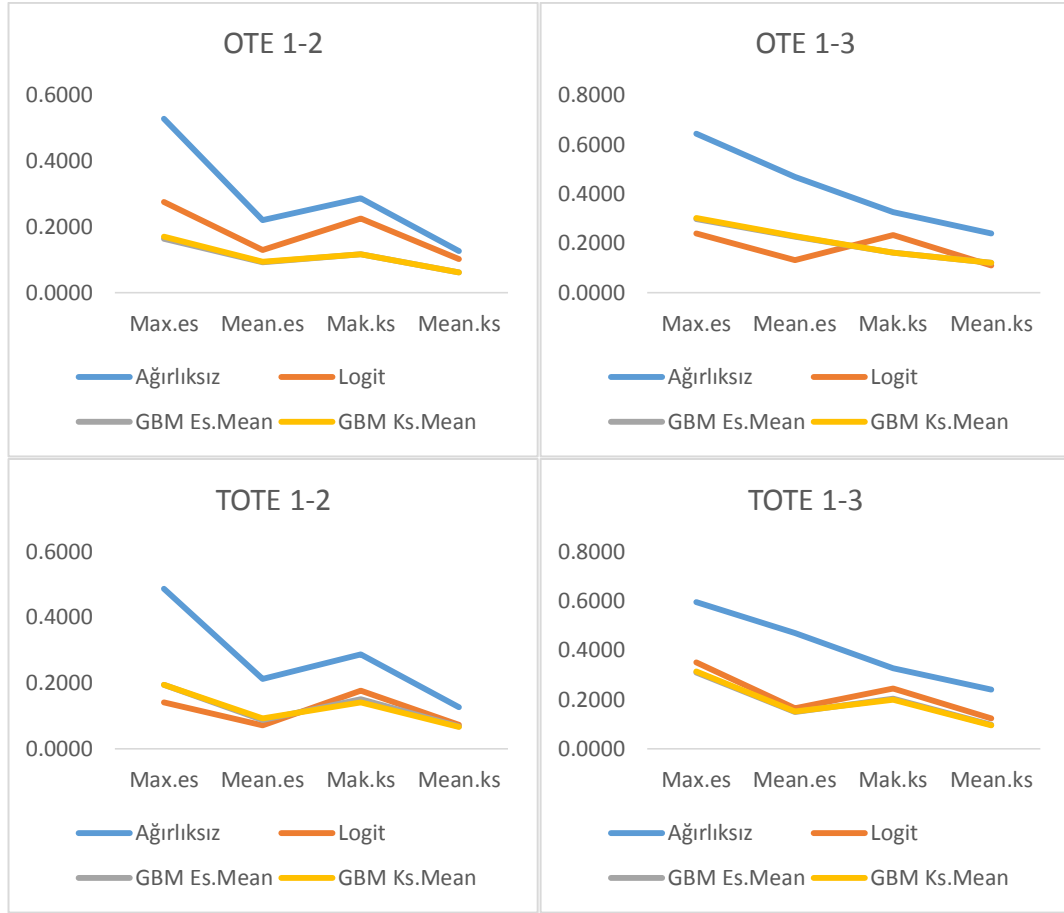
Şekil 5.13. Yöntemlere göre denge değerleri (S6, n=1000)

Ağırlıklandırma sonrası yine denge skorlarında düşüşler gözlenmektedir. Bu senaryoda MLR’de denge değerlerindeki benzerliğin GBM’den daha fazla olduğu gözükmemektedir (Tablo 5.13ve Şekil 5.13).

Senaryo 7, n=1000**Ana etkiler, on tane iki yönlü etkileşim, üç tane karesel terimin olduğu model**

Tablo 5.14. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 7; n=1000)

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	436,718	146,575	436,718	146,575	0,5287	0,2200	0,2864	0,1260
	1-3	436,718	416,707	436,718	416,707	0,6446	0,4693	0,3246	0,2381
Ağırlıksız TOTE	1-2	436,718	146,575	436,718	146,575	0,4865	0,2123	0,2864	0,1260
	1-3	436,718	416,707	436,718	416,707	0,5931	0,4674	0,3246	0,2381
MLR OTE	1-2	436,718	146,575	113,632	108,552	0,2743	0,1287	0,2246	0,1021
	1-3	436,718	416,707	113,632	166,839	0,2402	0,1311	0,2320	0,1094
MLR TOTE	1-2	436,718	146,575	436,718	113,216	0,1397	0,0699	0,1757	0,0726
	1-3	436,718	416,707	436,718	124,162	0,3488	0,1637	0,2433	0,1219
GBM Mean.es OTE	1-2	436,718	146,575	297,598	111,962	0,1629	0,0910	0,1161	0,0606
	1-3	436,718	416,707	297,598	275,510	0,2985	0,2272	0,1610	0,1207
GBM Mean.ks OTE	1-2	436,718	146,575	296,660	113,313	0,1695	0,0927	0,1153	0,0613
	1-3	436,718	416,707	296,660	278,016	0,3026	0,2287	0,1614	0,1210
GBM Mean.es TOTE	1-2	436,718	146,575	436,718	74,217	0,1933	0,0871	0,1508	0,0680
	1-3	436,718	416,707	436,718	83,942	0,3090	0,1484	0,2025	0,0959
GBM Mean.ks TOTE	1-2	436,718	146,575	436,718	79,090	0,1936	0,0910	0,1405	0,0658
	1-3	436,718	416,707	436,718	89,644	0,3136	0,1512	0,1983	0,0947



Şekil 5.14. Yöntemlere göre denge değerleri (S7, n=1000)

Ağırlıklandırma sonrası düşüşler gözlenmektedir. GBM'nin denge değerleri MLR'ye göre daha benzerdir (Tablo 5.14 ve Şekil 5.14).

Genel olarak tüm senaryolar için ağırlıklandırma sonrası düşüşler dengenin ortak değişken düzeltilmesinin çalıştığını göstermektedir. n=1000 örneklem büyüklüğü için türetilen veri setlerinden elde edilen tüm senaryo sonuçlarına göre genel olarak standartlaştırılmış yanlılık ve kolmogorov smirnov istatistik değerlerinde azalma n=500'e göre daha fazla olduğu görülmektedir.

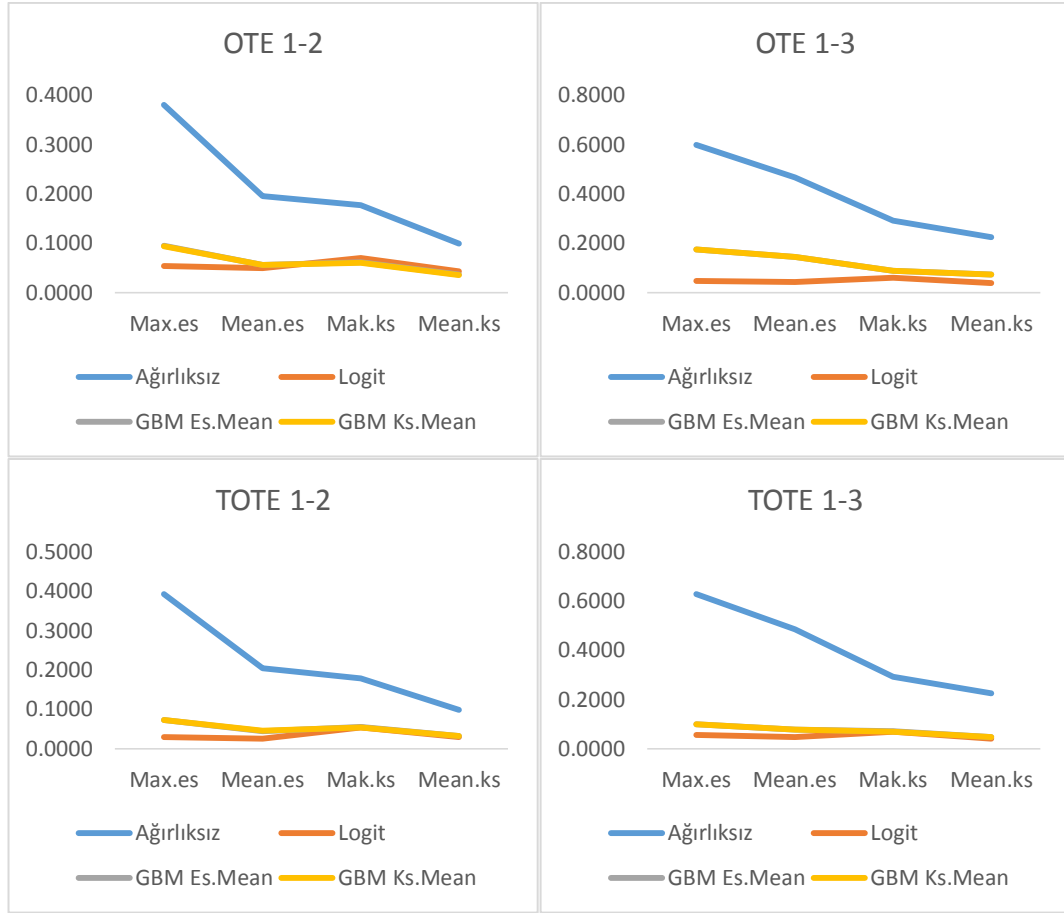
5.1.3 n=2000 için Senaryoların Denge Değerlendirmesi

Senaryo 1, n=2000

Ana etkilerin olduğu model

Tablo 5.15. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 1; n=2000)

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	555,43	408,83	555,43	408,83	0,3803	0,1958	0,1774	0,0987
	1-3	555,43	1035,74	555,43	1035,74	0,5978	0,4658	0,2907	0,2229
Ağırlıksız TOTE	1-2	555,43	408,83	555,43	408,83	0,3919	0,2037	0,1774	0,0987
	1-3	555,43	1035,74	555,43	1035,74	0,6258	0,4839	0,2907	0,2229
MLR OTE	1-2	555,43	408,83	234,32	303,44	0,0530	0,0488	0,0693	0,0429
	1-3	555,43	1035,74	234,32	629,46	0,0471	0,0418	0,0604	0,0369
MLR TOTE	1-2	555,43	408,83	555,43	328,28	0,0291	0,0249	0,0534	0,0290
	1-3	555,43	1035,74	555,43	387,38	0,0548	0,0464	0,0679	0,0403
GBM Mean.es OTE	1-2	555,43	408,83	361,21	331,75	0,0948	0,0557	0,0611	0,0349
	1-3	555,43	1035,74	361,21	780,41	0,1733	0,1433	0,0867	0,0733
GBM Mean.ks OTE	1-2	555,43	408,83	365,26	333,98	0,0938	0,0553	0,0603	0,0355
	1-3	555,43	1035,74	365,26	781,86	0,1744	0,1441	0,0872	0,0735
GBM Mean.es TOTE	1-2	555,43	408,83	555,43	258,57	0,0725	0,0439	0,0552	0,0315
	1-3	555,43	1035,74	555,43	232,89	0,0982	0,0770	0,0695	0,0474
GBM Mean.ks TOTE	1-2	555,43	408,83	555,43	263,92	0,0725	0,0446	0,0536	0,0311
	1-3	555,43	1035,74	555,43	237,18	0,0984	0,0776	0,0683	0,0471

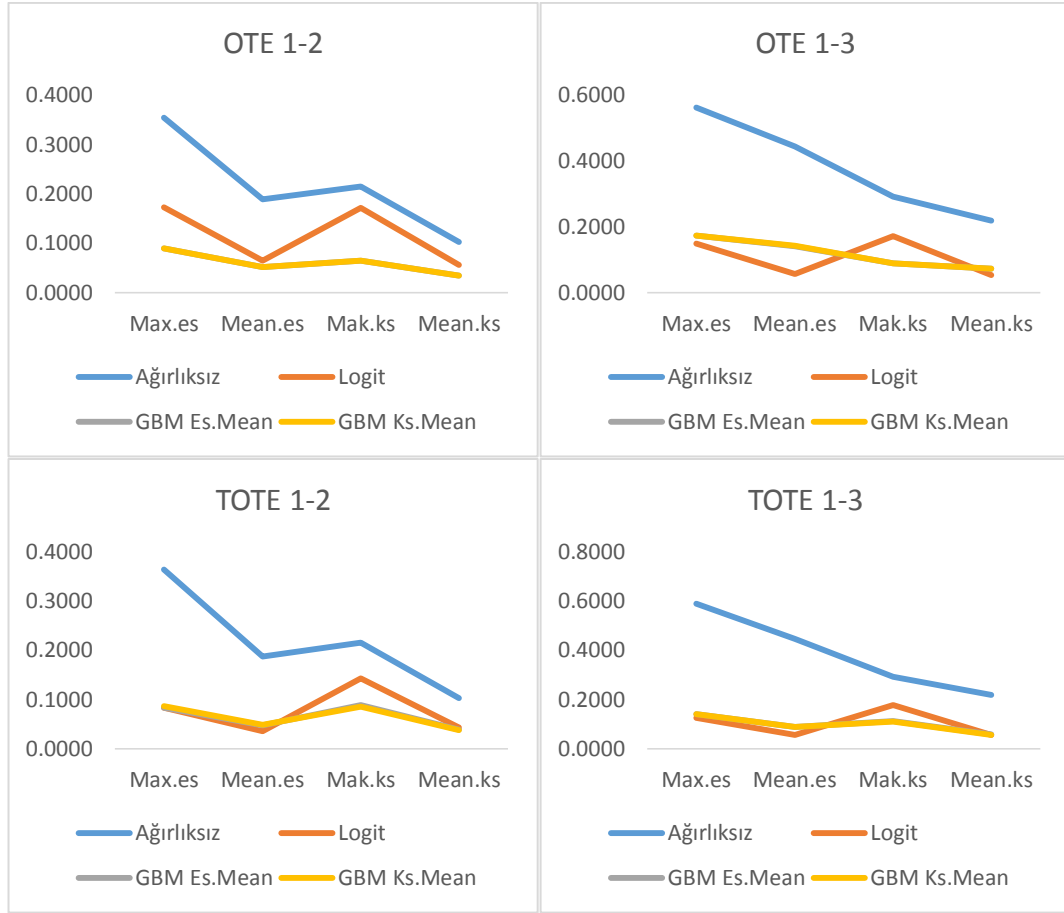


Şekil 5.15. Yöntemlere göre denge değerleri (S1, n=2000)

Mak.es, means.es, mak.ks ve Mean.ks değerleri ağırlıklandırma sonrasında 500 ve 1000 örneklemlili senaryoya göre daha düşük olduğu görülmektedir. Bu düşüş MLR için mak.es ve mean.es için genel olarak daha çoktur. Yine OTE 1-3 için MLR daha iyi denge skoru vermektedir. Dağılımları karşılaştırmaya yönelik KS istatistiği değerleri için GBM'den elde edilen denge skorları MLR'ye göre daha iyidir (Tablo 5.15 ve Şekil 5.15).

Senaryo 2, n=2000**Ana etkiler ve bir karesel terimin olduğu model****Tablo 5.16. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 2; n=2000)**

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	707,392	367,993	707,392	367,993	0,3542	0,1892	0,2145	0,1020
	1-3	707,392	924,615	707,392	924,615	0,5616	0,4445	0,2917	0,2175
Ağırlıksız TOTE	1-2	707,392	367,993	707,392	367,993	0,3629	0,1865	0,2145	0,1020
	1-3	707,392	924,615	707,392	924,615	0,5864	0,4438	0,2917	0,2175
MLR OTE	1-2	707,392	367,993	301,382	285,090	0,1721	0,0649	0,1711	0,0562
	1-3	707,392	924,615	301,382	549,898	0,1488	0,0563	0,1717	0,0519
MLR TOTE	1-2	707,392	367,993	707,392	298,300	0,0832	0,0355	0,1426	0,0430
	1-3	707,392	924,615	707,392	353,734	0,1250	0,0563	0,1760	0,0548
GBM Mean.es OTE	1-2	707,392	367,993	464,606	297,788	0,0898	0,0509	0,0647	0,0334
	1-3	707,392	924,615	464,606	666,237	0,1721	0,1410	0,0889	0,0727
GBM Mean.ks OTE	1-2	707,392	367,993	466,152	300,244	0,0892	0,0513	0,0647	0,0343
	1-3	707,392	924,615	466,152	669,270	0,1730	0,1417	0,0890	0,0729
GBM Mean.es TOTE	1-2	707,392	367,993	707,392	207,809	0,0832	0,0471	0,0887	0,0384
	1-3	707,392	924,615	707,392	219,058	0,1394	0,0877	0,1112	0,0565
GBM Mean.ks TOTE	1-2	707,392	367,993	707,392	216,322	0,0856	0,0486	0,0851	0,0376
	1-3	707,392	924,615	707,392	224,989	0,1399	0,0886	0,1092	0,0561

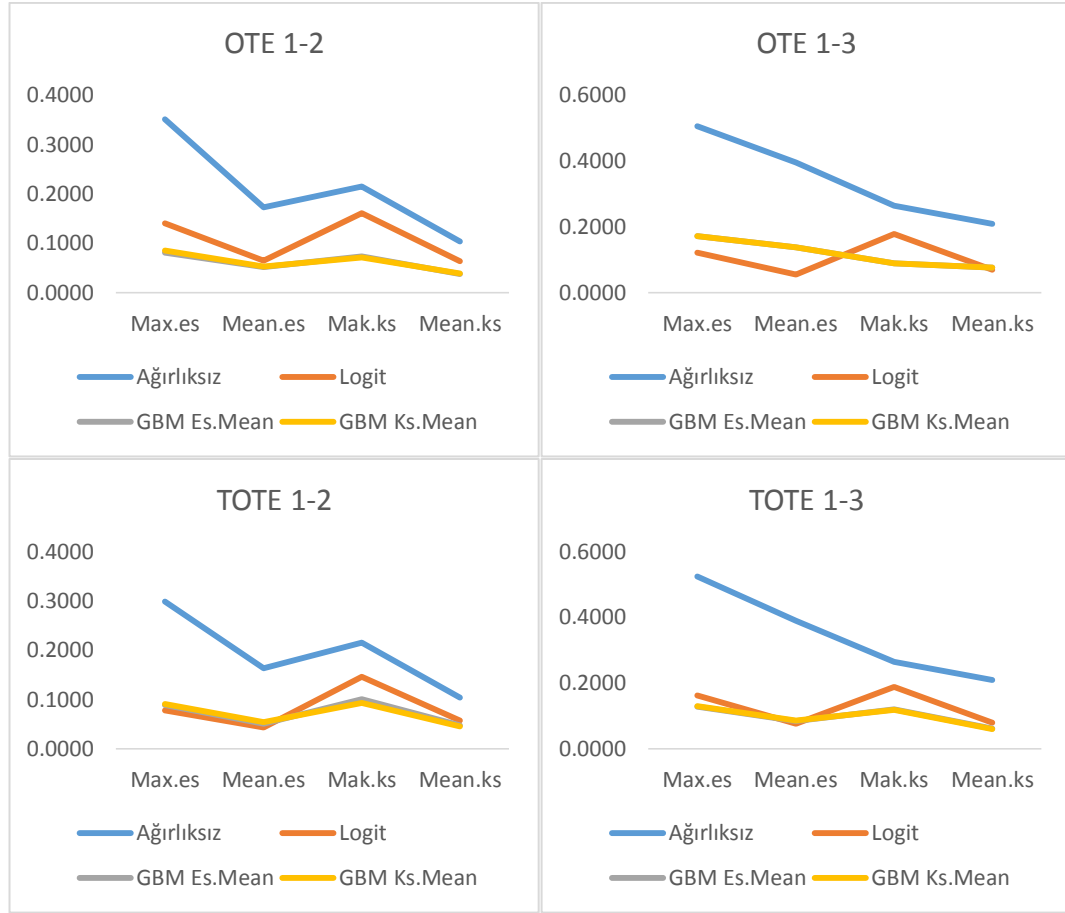


Şekil 5.16. Yöntemlere göre denge değerleri (S2,n=2000)

Mak.es, means.es, mak.ks ve Mean.ks değerleri ağırlıklandırma sonrasında 500 ve 1000 örneklemlili senaryoya göre daha da düşük olduğu görülmektedir. Bu düşüş MLR için mak.es ve mean.es için genel olarak daha çoktur. Genel olarak GBM'den elde edilen denge skorları MLR'ye göre daha iyidir (Tablo 5.16 ve Şekil 5.16).

Senaryo 3, n=2000**Ana etkiler ve üç karesel terimin olduğu model****Tablo 5.17. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 3; n=2000)**

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	808,192	278,741	808,192	278,741	0,3508	0,1727	0,2148	0,1034
	1-3	808,192	913,067	808,192	913,067	0,5058	0,3947	0,2638	0,2093
Ağırlıksız TOTE	1-2	808,192	278,741	808,192	278,741	0,2981	0,1630	0,2148	0,1034
	1-3	808,192	913,067	808,192	913,067	0,5238	0,3886	0,2638	0,2093
MLR OTE	1-2	808,192	278,741	408,541	233,580	0,1403	0,0642	0,1608	0,0630
	1-3	808,192	913,067	408,541	610,079	0,1213	0,0539	0,1774	0,0696
MLR TOTE	1-2	808,192	278,741	808,192	236,353	0,0773	0,0424	0,1456	0,0563
	1-3	808,192	913,067	808,192	422,897	0,1608	0,0753	0,1874	0,0787
GBM Mean.es OTE	1-2	808,192	278,741	516,317	213,430	0,0809	0,0514	0,0728	0,0376
	1-3	808,192	913,067	516,317	630,129	0,1706	0,1364	0,0888	0,0758
GBM Mean.ks OTE	1-2	808,192	278,741	516,519	218,537	0,0851	0,0523	0,0714	0,0382
	1-3	808,192	913,067	516,519	632,856	0,1719	0,1372	0,0892	0,0760
GBM Mean.es TOTE	1-2	808,192	278,741	808,192	147,095	0,0882	0,0507	0,0998	0,0472
	1-3	808,192	913,067	808,192	197,232	0,1273	0,0833	0,1198	0,0601
GBM Mean.ks TOTE	1-2	808,192	278,741	808,192	155,255	0,0904	0,0540	0,0923	0,0454
	1-3	808,192	913,067	808,192	204,364	0,1294	0,0846	0,1171	0,0595



Şekil 5.17. Yöntemlere göre denge değerleri (S3, n=2000)

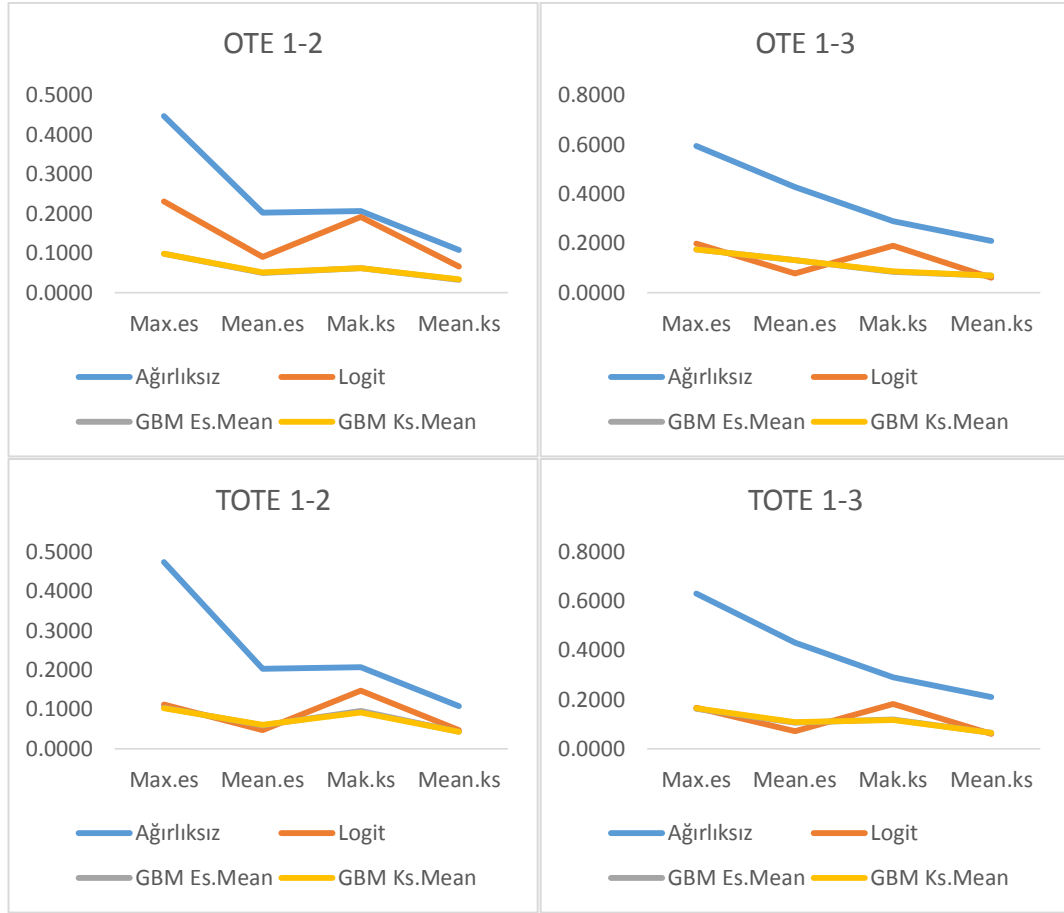
Mak.es, means.es, mak.ks ve Mean.ks değerleri ağırlıklandırma sonrasında 500 ve 1000 örneklemlili senaryoya göre daha da düşük olduğu görülmektedir. Bu düşüş MLR için mak.es ve mean.es için genel olarak daha çöktür. Genel olarak GBM'den elde edilen denge skorları MLR'ye göre daha iyidir (Tablo 5.17 ve Şekil 5.17).

Senaryo 4, n=2000

Ana etkiler ve üç tane iki yönlü etkileşim terimin olduğu model

Tablo 5.18. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 4; n=2000)

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	740,140	388,356	740,140	388,356	0,4464	0,2024	0,2065	0,1070
	1-3	740,140	871,504	740,140	871,504	0,5930	0,4267	0,2878	0,2081
Ağırlıksız TOTE	1-2	740,140	388,356	740,140	388,356	0,4724	0,2027	0,2065	0,1070
	1-3	740,140	871,504	740,140	871,504	0,6275	0,4285	0,2878	0,2081
MLR OTE	1-2	740,140	388,356	278,919	323,472	0,2307	0,0896	0,1919	0,0653
	1-3	740,140	871,504	278,919	572,264	0,1986	0,0771	0,1892	0,0602
MLR TOTE	1-2	740,140	388,356	740,140	319,362	0,1117	0,0464	0,1462	0,0462
	1-3	740,140	871,504	740,140	381,280	0,1650	0,0713	0,1815	0,0589
GBM Mean.es OTE	1-2	740,140	388,356	497,406	321,592	0,0975	0,0492	0,0618	0,0324
	1-3	740,140	871,504	497,406	635,287	0,1731	0,1305	0,0843	0,0672
GBM Mean.ks OTE	1-2	740,140	388,356	497,053	323,277	0,0981	0,0502	0,0621	0,0333
	1-3	740,140	871,504	497,053	640,379	0,1741	0,1313	0,0847	0,0674
GBM Mean.es TOTE	1-2	740,140	388,356	740,140	202,132	0,1033	0,0582	0,0950	0,0433
	1-3	740,140	871,504	740,140	218,224	0,1615	0,1054	0,1181	0,0636
GBM Mean.ks TOTE	1-2	740,140	388,356	740,140	211,920	0,1026	0,0598	0,0920	0,0425
	1-3	740,140	871,504	740,140	225,806	0,1632	0,1064	0,1158	0,0631



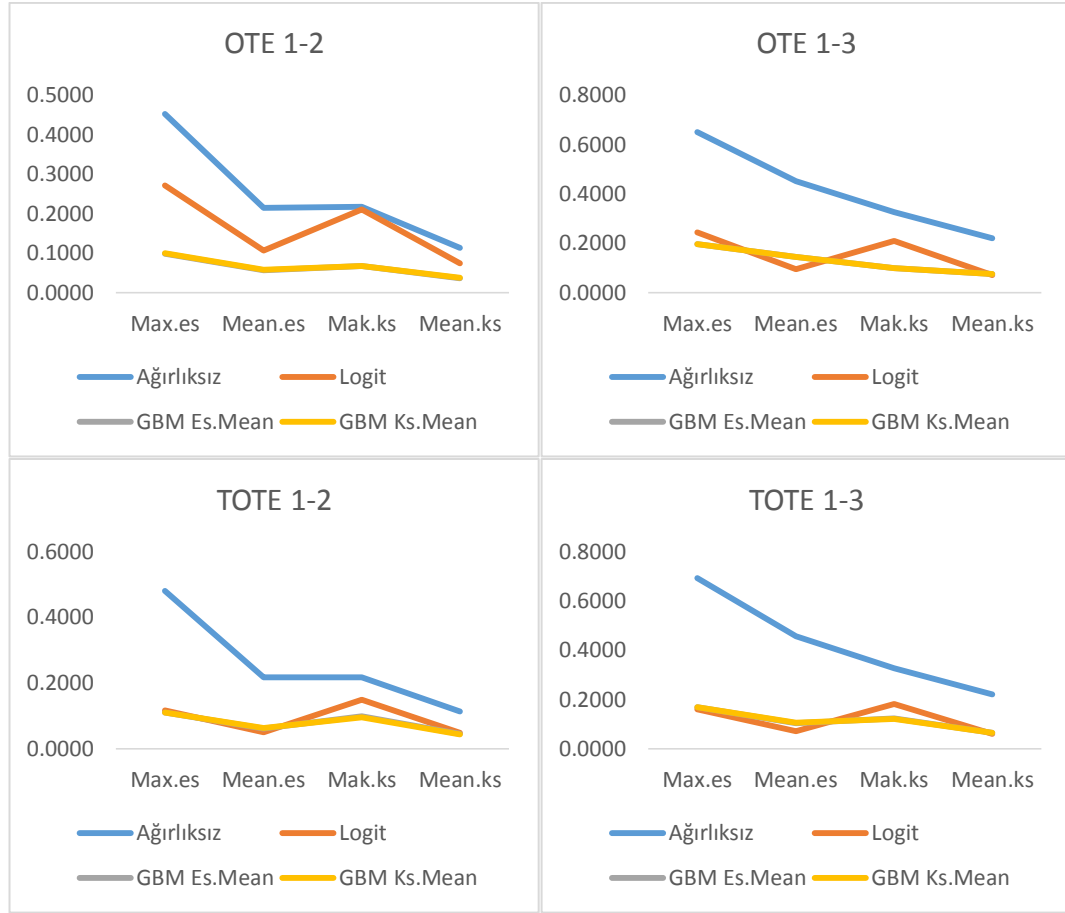
Şekil 5.18. Yöntemlere göre denge değerleri (S4, n=2000)

Mak.es, means.es, mak.ks ve Mean.ks değerleri ağırlıklandırma sonrasında 500 ve 1000 örneklemlili senaryoya göre daha da düşük olduğu görülmektedir. Bu düşüş MLR için mak.es ve mean.es için genel olarak daha çoktur. Genel olarak GBM'den elde edilen denge skorları MLR'ye göre daha iyidir (Tablo 5.18 ve Şekil 5.18).

Senaryo 5, n=2000**Ana etkiler, dört tane iki yönlü etkileşim ve bir karesel terimin olduğu model**

Tablo 5.19. MLR ve GBM’ye ilişkin özet denge değerleri (Senaryo 5; n=2000)

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	710,031	376,482	710,031	376,482	0,4528	0,2146	0,2170	0,1126
	1-3	710,031	913,487	710,031	913,487	0,6495	0,4511	0,3248	0,2200
Ağırlıksız TOTE	1-2	710,031	376,482	710,031	376,482	0,4798	0,2162	0,2170	0,1126
	1-3	710,031	913,487	710,031	913,487	0,6922	0,4556	0,3248	0,2200
MLR OTE	1-2	710,031	376,482	212,743	302,854	0,2707	0,1058	0,2105	0,0739
	1-3	710,031	913,487	212,743	560,519	0,2430	0,0951	0,2088	0,0694
MLR TOTE	1-2	710,031	376,482	710,031	303,189	0,1159	0,0490	0,1486	0,0478
	1-3	710,031	913,487	710,031	372,494	0,1589	0,0706	0,1814	0,0592
GBM Mean.es OTE	1-2	710,031	376,482	446,929	307,732	0,0987	0,0560	0,0670	0,0362
	1-3	710,031	913,487	446,929	651,008	0,1948	0,1433	0,0974	0,0738
GBM Mean.ks OTE	1-2	710,031	376,482	446,624	310,328	0,0996	0,0569	0,0676	0,0369
	1-3	710,031	913,487	446,624	655,997	0,1961	0,1440	0,0981	0,0740
GBM Mean.es TOTE	1-2	710,031	376,482	710,031	188,969	0,1103	0,0606	0,0988	0,0447
	1-3	710,031	913,487	710,031	193,776	0,1672	0,1037	0,1224	0,0635
GBM Mean.ks TOTE	1-2	710,031	376,482	710,031	199,032	0,1098	0,0624	0,0953	0,0438
	1-3	710,031	913,487	710,031	200,759	0,1686	0,1049	0,1201	0,0629



Şekil 5.19. Yöntemlere göre denge değerleri (S5, n=2000)

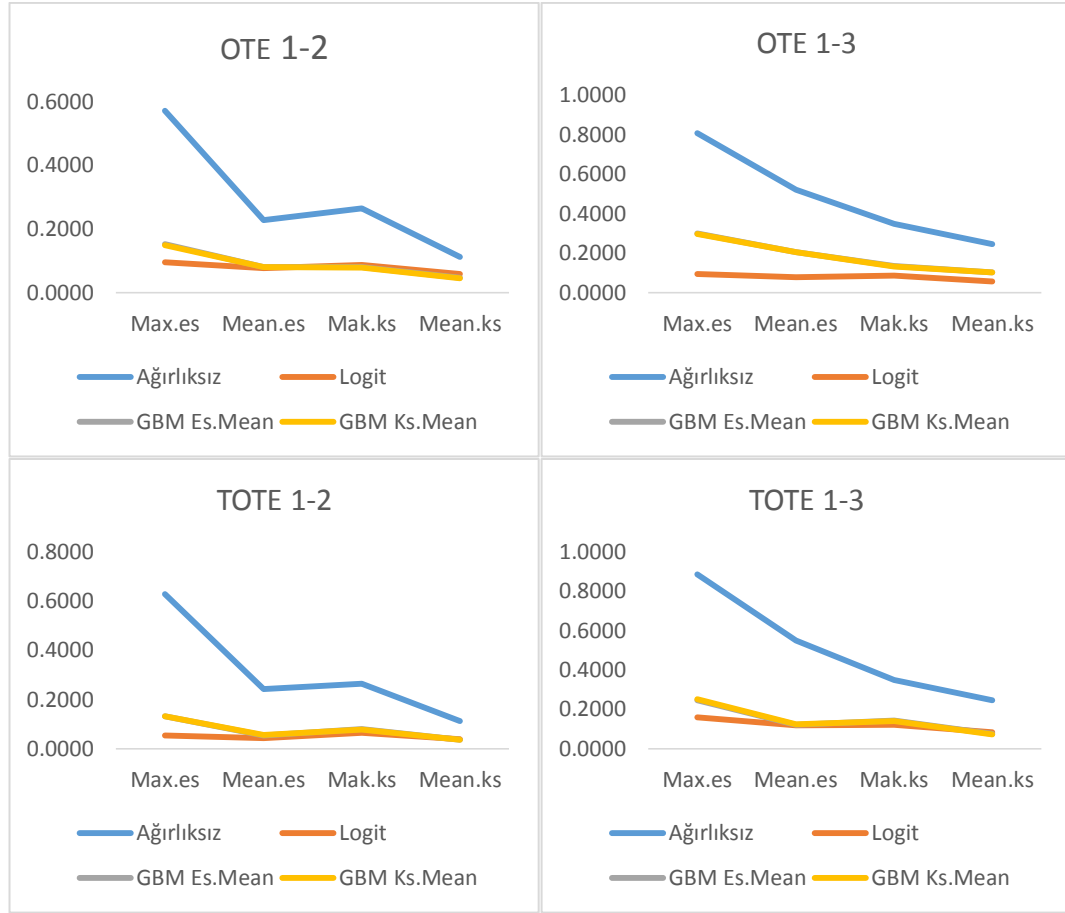
Mak.es, means.es, mak.ks ve Mean.ks değerleri ağırlıklandırma sonrasında 500 ve 1000 örneklemlili senaryoya göre daha da düşük olduğu görülmektedir. Bu düşüş MLR için mak.es ve mean.es için genel olarak daha çoktur. Genel olarak GBM'den elde edilen denge skorları MLR'ye göre daha iyidir (Tablo 5.19 ve Şekil 5.19).

Senaryo 6, n=2000

Ana etkiler, on tane iki yönlü etkileşim terimin olduğu model

Tablo 5.20. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 6; n=2000)

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	666,903	419,389	666,903	419,389	0,5708	0,2272	0,2637	0,1118
	1-3	666,903	913,708	666,903	913,708	0,8064	0,5217	0,3480	0,2460
Ağırlıksız TOTE	1-2	666,903	419,389	666,903	419,389	0,6254	0,2409	0,2637	0,1118
	1-3	666,903	913,708	666,903	913,708	0,8838	0,5468	0,3480	0,2460
MLR OTE	1-2	666,903	419,389	179,370	261,870	0,0951	0,0757	0,0867	0,0576
	1-3	666,903	913,708	179,370	303,862	0,0932	0,0761	0,0865	0,0564
MLR TOTE	1-2	666,903	419,389	666,903	305,240	0,0535	0,0415	0,0645	0,0371
	1-3	666,903	913,708	666,903	207,682	0,1592	0,1178	0,1194	0,0823
GBM Mean.es OTE	1-2	666,903	419,389	396,222	332,579	0,1518	0,0799	0,0799	0,0457
	1-3	666,903	913,708	396,222	533,483	0,2979	0,2041	0,1327	0,1003
GBM Mean.ks OTE	1-2	666,903	419,389	400,981	333,900	0,1485	0,0796	0,0782	0,0453
	1-3	666,903	913,708	400,981	536,544	0,2972	0,2050	0,1321	0,1006
GBM Mean.es TOTE	1-2	666,903	419,389	666,903	215,502	0,1302	0,0535	0,0784	0,0360
	1-3	666,903	913,708	666,903	115,783	0,2460	0,1208	0,1418	0,0734
GBM Mean.ks TOTE	1-2	666,903	419,389	666,903	221,155	0,1301	0,0545	0,0767	0,0356
	1-3	666,903	913,708	666,903	125,581	0,2498	0,1236	0,1380	0,0721



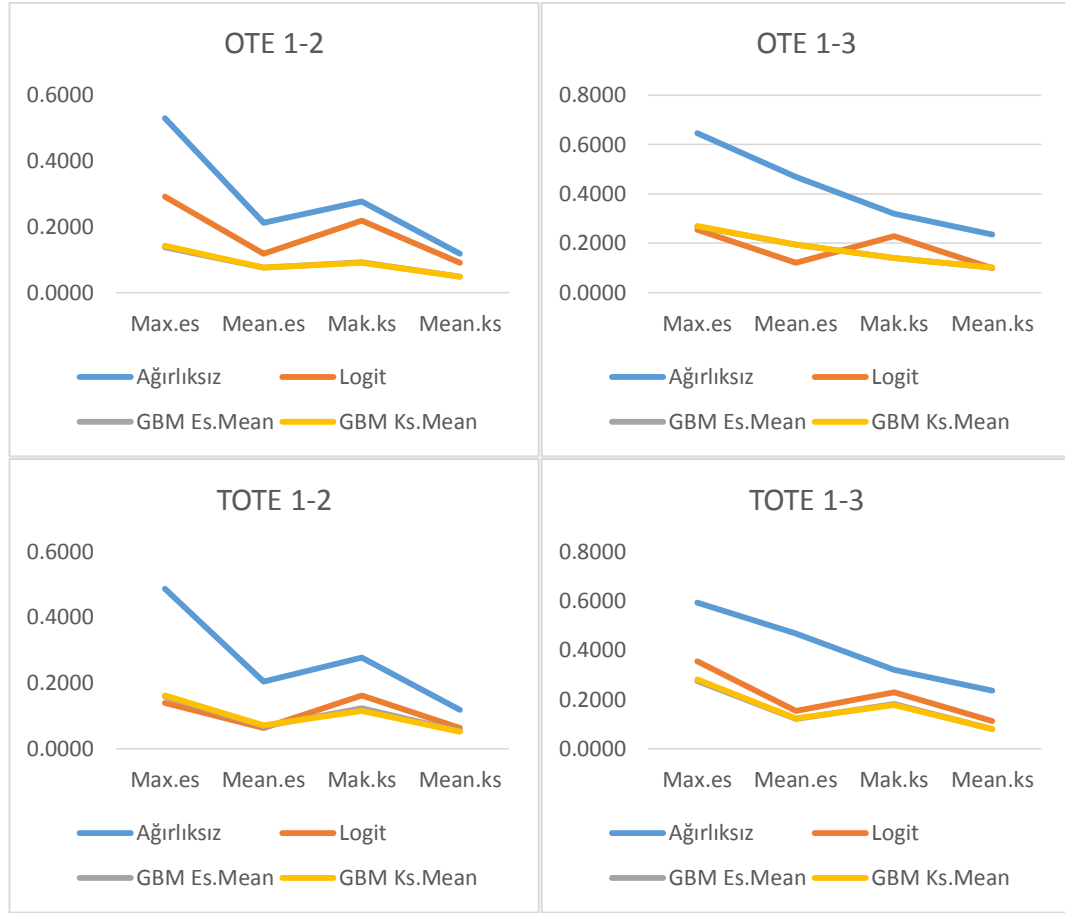
Şekil 5.20. Yöntemlere göre denge değerleri (S6, n=2000)

Mak.es, means.es, mak.ks ve Mean.ks değerleri ağırlıklandırma sonrasında 500 ve 1000 örneklemlili senaryoya göre daha da düşük olduğu görülmektedir. Bu düşüş MLR için mak.es ve mean.es için genel olarak daha çoktur. Genel olarak GBM'den elde edilen denge skorları MLR'ye göre daha iyidir (Tablo 5.20 ve Şekil 5.20).

Senaryo 7, n=2000**Ana etkiler, on tane iki yönlü etkileşim, üç tane karesel terimin olduğu model**

Tablo 5.21. MLR ve GBM'ye ilişkin özet denge değerleri (Senaryo 7; n=2000)

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	874,713	293,202	874,713	293,202	0,5296	0,2119	0,2773	0,1174
	1-3	874,713	832,085	874,713	832,085	0,6454	0,4691	0,3196	0,2353
Ağırlıksız TOTE	1-2	874,713	293,202	874,713	293,202	0,4863	0,2034	0,2773	0,1174
	1-3	874,713	832,085	874,713	832,085	0,5925	0,4667	0,3196	0,2353
MLR OTE	1-2	874,713	293,202	187,476	221,061	0,2907	0,1169	0,2183	0,0895
	1-3	874,713	832,085	187,476	315,840	0,2542	0,1205	0,2282	0,0992
MLR TOTE	1-2	874,713	293,202	874,713	229,931	0,1384	0,0620	0,1609	0,0628
	1-3	874,713	832,085	874,713	246,062	0,3541	0,1526	0,2276	0,1108
GBM Mean.es OTE	1-2	874,713	293,202	542,906	216,775	0,1369	0,0748	0,0909	0,0480
	1-3	874,713	832,085	542,906	479,550	0,2674	0,1929	0,1387	0,1005
GBM Mean.ks OTE	1-2	874,713	293,202	542,400	222,163	0,1416	0,0760	0,0899	0,0478
	1-3	874,713	832,085	542,400	483,286	0,2698	0,1939	0,1390	0,1007
GBM Mean.es TOTE	1-2	874,713	293,202	874,713	137,126	0,1590	0,0667	0,1218	0,0525
	1-3	874,713	832,085	874,713	128,811	0,2728	0,1207	0,1799	0,0791
GBM Mean.ks TOTE	1-2	874,713	293,202	874,713	148,076	0,1616	0,0700	0,1136	0,0505
	1-3	874,713	832,085	874,713	140,191	0,2800	0,1233	0,1764	0,0779



Şekil 5.21. Yöntemlere göre denge değerleri (S7, n=2000)

n=2000 örneklem için bulunan sonuçlara göre denge değerlerinde ortalama azalma n=500 ve n=1000 örneklemli türetilen veri setlerindeki denge değerleri ortalamalarından daha düşüktür. Yani örneklem sayısı arttıkça skorlardaki düşüş de artmaktadır (Tablo 5.21 ve Şekil 5.21).

Senaryolar kompleks hale geldikçe genel olarak GBM'nin MLR'den daha iyi sonuçlar ortaya koymaktadır. Ana etkilerin olduğu modelde ise MLR'nin daha iyi sonuçlar ortaya çıkardığı görülmektedir. Kompleks modeller için esnek bir kestirim yapan GBM yöntemi MLR'den daha başarılıdır.

5.2 Yanıt Değişken Üzerine Olan Etkilerin Değerlendirilmesi

GBM için twang paketi yardımıyla elde edilen kontrollerin (grafiksel ve tablo) ardından, mnps fonksiyonu ile türetilen propensity skorlarını kullanarak yanıt değişken analiz edilecektir. Kullanılan iki durma kuralının (Mean.es, Mean.ks) benzer sonuçlar verdiği göz önüne alınırsa, tek bir durma kuralı alınarak analize devam edilebilmektedir. Mean.es ağırlıklarını kullanarak veriyi analiz etmek için survey paketi kullanılmaktadır. GBM için ağırlıkları veri setine eklemek için get.weights fonksiyonu kullanılmaktadır.

MLR'de elde edilen ağırlıklar kullanılarak yanıt üzerine etki incelenmektedir. Elde edilecek modeller hem OTE hem TOTE ye göre değerlendirilmektedir. Tablo 5.22-5.24'deki S1-S7 ifadeleri 1'den 7'ye kadar olan senaryoları temsil etmektedir.

5.2.1 n=500 için Etkilerin Değerlendirilmesi

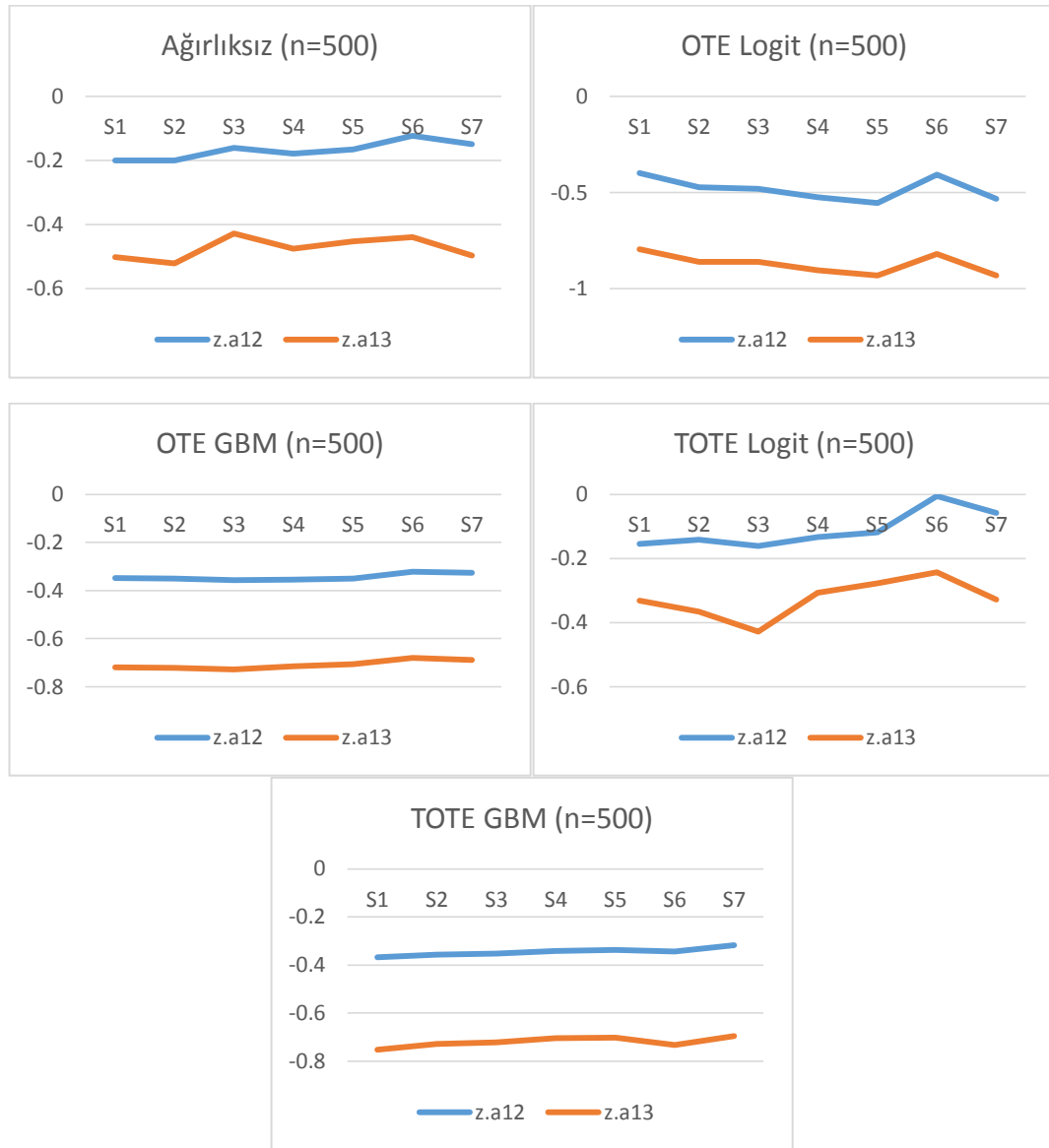
Tablo 5.22. Senaryolara ilişkin tedavi etkileri (n=500)

Kestirim	Tedavi etkisi								Propensity skor kestirimi
	Kategori	S1	S2	S3	S4	S5	S6	S7	
Ağırlıksız	z.a12	-0,2018	-0,2038	-0,2118	-0,1807	-0,1616	-0,1254	-0,1556	Ağırlıksız
	z.a13	-0,5046	-0,5204	-0,5604	-0,4756	-0,4537	-0,4390	-0,4998	
OTE	z.a12	-0,3975	-0,4691	-0,4779	-0,5237	-0,5449	-0,3997	-0,5290	MLR
	z.a13	-0,7977	-0,8552	-0,8619	-0,8999	-0,9223	-0,8097	-0,9189	
	z.a12	-0,3264	-0,3329	-0,3369	-0,3417	-0,3324	-0,2905	-0,3081	GBM Mean.es
	z.a13	-0,6913	-0,6938	-0,7093	-0,6891	-0,6787	-0,6402	-0,6657	
TOTE	z.a12	-0,1568	-0,1494	-0,1577	-0,1335	-0,1117	-0,0093	-0,0666	MLR
	z.a13	-0,3313	-0,3640	-0,4318	-0,3047	-0,2770	-0,2367*	-0,3343	
	z.a12	-0,3524	-0,3328	-0,3164	-0,3176	-0,3090	-0,3176	-0,2856	GBM Mean.es
	z.a13	-0,7275	-0,6971	-0,6879	-0,6714	-0,6657	-0,7000	-0,6618	

İtalik ve koyu olarak verilen sayıların tedavi etkilerinin önemsiz iken ağırlıklandırmadan sonra önemli hale geldiğini ifade etmektedir.

**: Ağırlıklandırmadan önce önemli iken önemsiz hale gelmektedir.*

z.a12 (tedavi değişkeni): 1 referans kategoriye göre 2 tedavi grubunun yanıt üzerindeki etkisi



Şekil 5.22. n=500 için tedavi etkileri

Tüm senaryolar için ağırlıklandırmadan önceki z.a12 değişkeninin yanıt üzerinde etkisi önemsiz iken z.a13 değişkeninin etkisi önemlidir. Senaryo 6 için elde edilen sonuçlara göre ağırlıklandırmadan önce z.a13 değişkeninin yanıt üzerindeki etkisi önemli iken lojistik regresyon kullanarak ağırlıklandırdıktan elde edilen TOTE kestirimine göre yanıt üzerindeki etkisi önemsiz hale gelmektedir. Senaryo 7 için elde edilen sonuçlara göre, GBM kullanarak elde edilen TOTE kestirimine göre diğer senaryolarda z.a12 değişkeninin etkisi önemli iken burada yine önemsizliğini korumaktadır.

5.2.2 n=1000 için Etkilerin Değerlendirilmesi

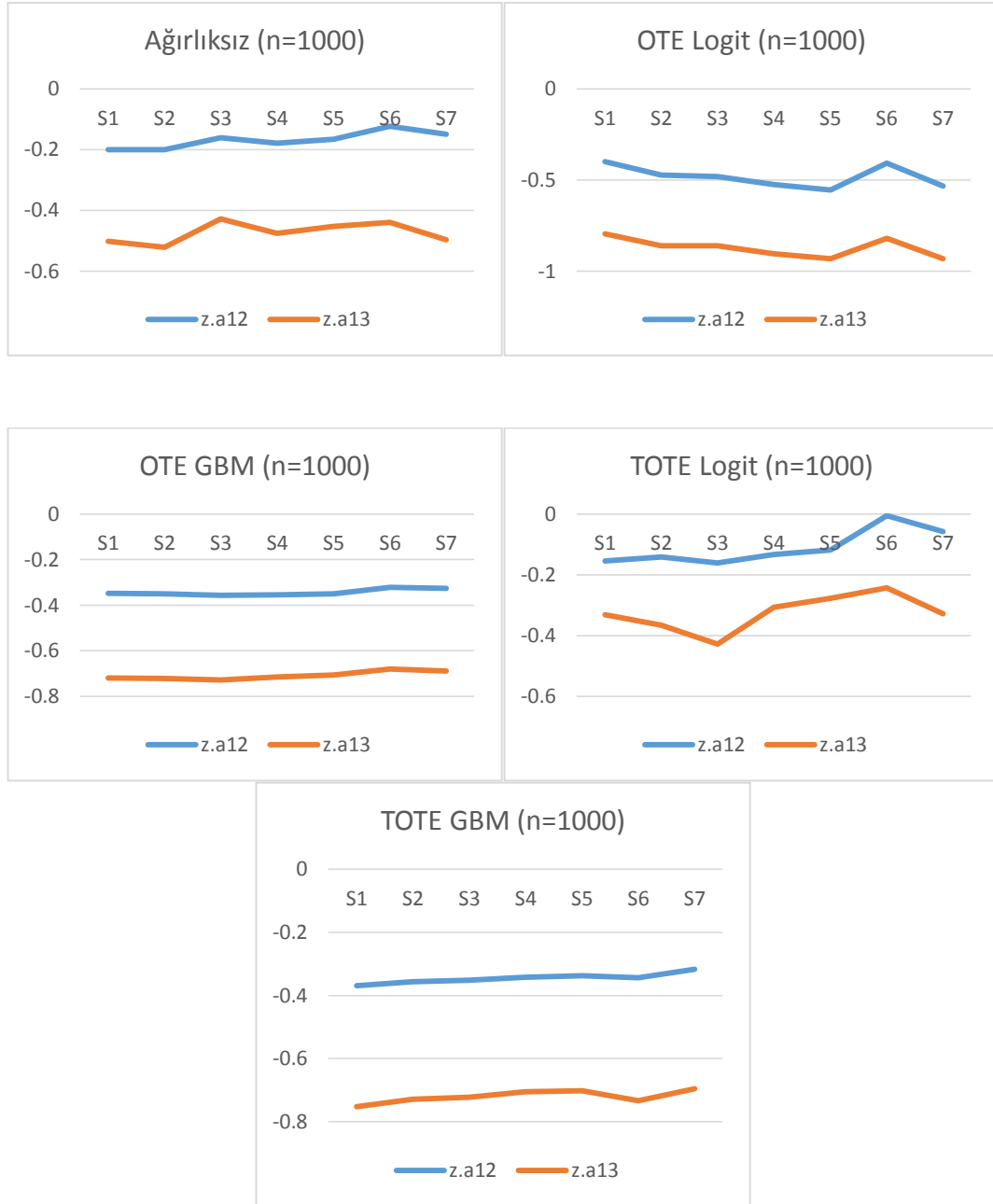
Tablo 5.23. Senaryolara ilişkin tedavi etkileri (n=1000)

Kestirim	Tedavi etkisi								Propensity skor kestirimi
	Kategori	S1	S2	S3	S4	S5	S6	S7	
Ağırlıksız	z.a12	-0,2026	-0,1978	-0,2118	-0,1794	-0,1646	-0,1224	-0,1509	Ağırlıksız
	z.a13	-0,5035	-0,5196	-0,5604	-0,4748	-0,4537	-0,4397	-0,4969	
OTE	z.a12	-0,4005	-0,4715	-0,4779	-0,5243	-0,5508	-0,4031	-0,5297	MLR
	z.a13	-0,7993	-0,8601	-0,8619	-0,9021	-0,9299	-0,8165	-0,9284	
	z.a12	-0,3374	-0,3414	-0,3369	-0,3450	-0,3384	-0,3068	-0,3158	GBM Mean.es
	z.a13	-0,7059	-0,7099	-0,7093	-0,7010	-0,6914	-0,6606	-0,6779	
TOTE	z.a12	-0,1558*	-0,1408*	-0,1577*	-0,1331*	-0,1179	-0,0058	-0,0630	MLR
	z.a13	-0,3334	-0,3652	-0,4318	-0,3055	-0,2746	-0,2410	-0,3271	
	z.a12	-0,3631	-0,3455	-0,3164	-0,3295	-0,3274	-0,3337	-0,3005	GBM Mean.es
	z.a13	-0,7396	-0,7126	-0,6879	-0,6895	-0,6848	-0,7191	-0,6814	

İtalik ve koyu olarak verilen sayıların tedavi etkilerinin önemsiz iken ağırlıklandırmadan sonra önemli hale geldiğini ifade etmektedir.

**:Değişkenin etkisi önemli iken önemsiz hale gelmektedir.*

z.a12 (tedavi değişkeni): 1 referans kategoriye göre 2 tedavi grubunun yanıt üzerindeki etkisi



Şekil 5.23. n=1000 için tedavi etkileri

Senaryo 1-4 için ağırlıklandırmadan önceki tedavi etkileri önemlidir. Ağırlıklandırmadan önce senaryo 5,6 ve 7 için ise z.a12 değişkenin yanıt üzerinde etkisi önemsizdir. Senaryo 5, 6 ve 7’de lojistik regresyon TOTE kestirimi hariç tüm yöntemlerin z.a12 değişkeni için yanıt üzerinde etkisi önemlidir.

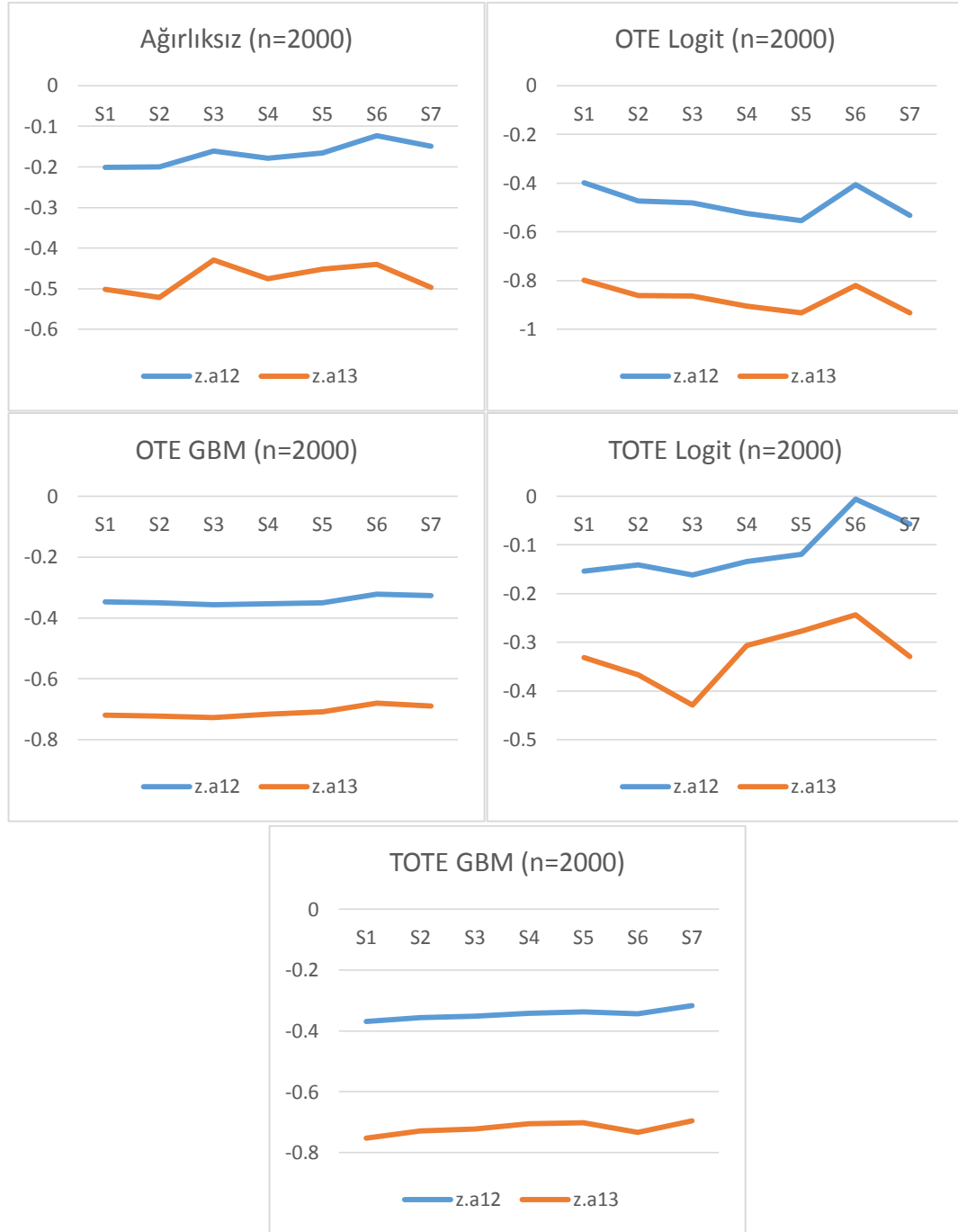
5.2.3 n=2000 için Etkilerin Değerlendirilmesi

Tablo 5.24. Senaryolara ilişkin tedavi etkileri (n=2000)

Kestirim	Tedavi etkisi								Propensity skor kestirimi
	Kategori	S1	S2	S3	S4	S5	S6	S7	
Ağırlıksız	z.a12	-0,2010	-0,1998	-0,1614	-0,1792	-0,1655	-0,1233	-0,1490	Ağırlıksız
	z.a13	-0,5016	-0,5211	-0,4292	-0,4759	-0,4523	-0,4403	-0,4967	
OTE	z.a12	-0,3989	-0,4734	-0,4809	-0,5246	-0,5550	-0,4071	-0,5331	MLR
	z.a13	-0,7974	-0,8620	-0,8625	-0,9043	-0,9316	-0,8209	-0,9324	
	z.a12	-0,3474	-0,3503	-0,3564	-0,3539	-0,3497	-0,3226	-0,3267	GBM Mean.es
	z.a13	-0,7194	-0,7224	-0,7276	-0,7156	-0,7077	-0,6799	-0,6894	
TOTE	z.a12	-0,1543*	-0,1414*	-0,1614	-0,1339*	-0,1192*	-0,0057*	-0,0577	MLR
	z.a13	-0,3311	-0,3669	-0,4292	-0,3065	-0,2769	-0,2433	-0,3290	
	z.a12	-0,3684	-0,3571	-0,3525	-0,3420	-0,3380	-0,3447	-0,3178	GBM Mean.es
	z.a13	-0,7529	-0,7282	-0,7225	-0,7044	-0,7023	-0,7336	-0,6958	

İtalik ve koyu verilen sayıların tedavi etkilerinin önemsiz iken ağırlıklandırmadan sonra önemli hale geldiğini ifade etmektedir.

**:Değişkenin etkisi önemli iken önemsiz hale gelmektedir.*



Şekil 5.24. n=2000 için tedavi etkileri

Senaryo 1,2,4,5,6 için ağırlıklandırmadan önceki tedavi etkileri önemlidir. Ağırlıklandırmadan önce senaryo 3 ve 7 için ise z.a12 değişkenin yanıt üzerinde etkisi önemsizdir.

5.3 Gerçek Veri Seti Uygulaması

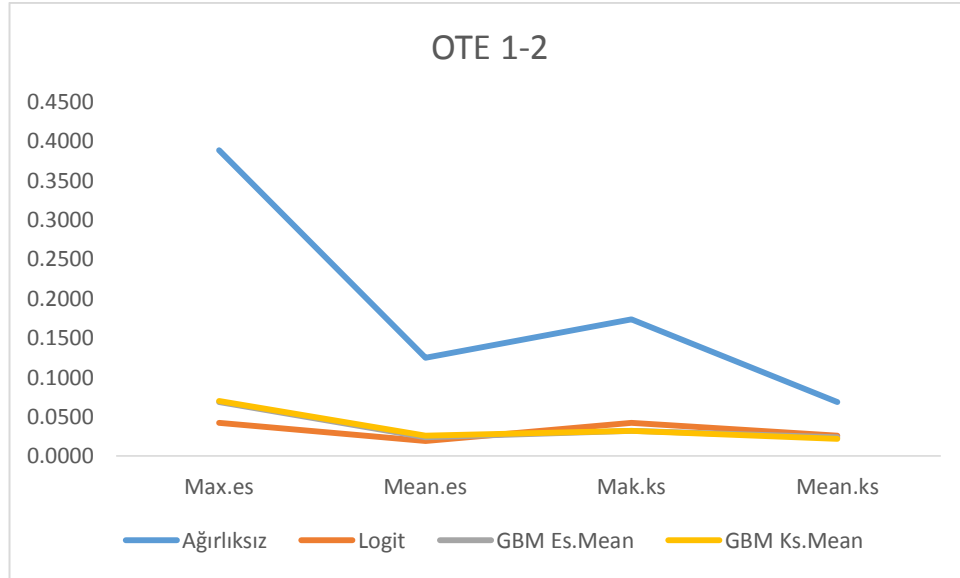
Gaziosmanpaşa Üniversitesi Sağlık Araştırma ve Uygulama merkezine 2010-2015 yılları arasında başvuran ve aynı anda Vitamin B12, MCHC (Mean Corpuscular Hemoglobin Concentration-Ortalama Alyuvar Hemoglobin Derişimi), MCV (Mean Corpuscular Volume-Ortalama Alyuvar Çapı), HGB (Hemoglobin), FOLAT (Folik Asit), HCT (Hematokrit) laboratuvar tetkiklerinin çalışıldığı 5448 birey örneklem olarak alınmaktadır.

Çalışmada, bireylerin yaşadıkları yerin üç gruba (Tokat Merkez, Tokat İlçeleri ve Tokat Dışı) ayrıldığı ve tedavi değişkeni olarak düşünülen yer değişkeninin, yanıt değişken (vitamin B12) üzerinde etkisi araştırılmaktadır. Çalışmada vitamin B12 yanıt değişkeni üzerinde etkili olduğu düşünülen diğer laboratuvar tetkikleri ile beraber yaş, cinsiyet ve medeni durum değişkenleri de ortak değişken olarak alınmaktadır. Kısaca düşünülen model Şekil 4.2’de gösterilmektedir. Ayrıca uygulama verisinde kullanılan değişkenlerin yaşadığı yere göre tanımlayıcı istatistikleri ve hipotez testleri Tablo 4.1’de ve vitamin B12 üzerinde etkisi olduğu düşünülen değişkenlere ilişkin doğrusal regresyon modeli sonuçları Tablo 4.2’de verilmektedir.

5.3.1 Uygulama Verisi için MLR ve GBM ile Denge Değerlendirmesi

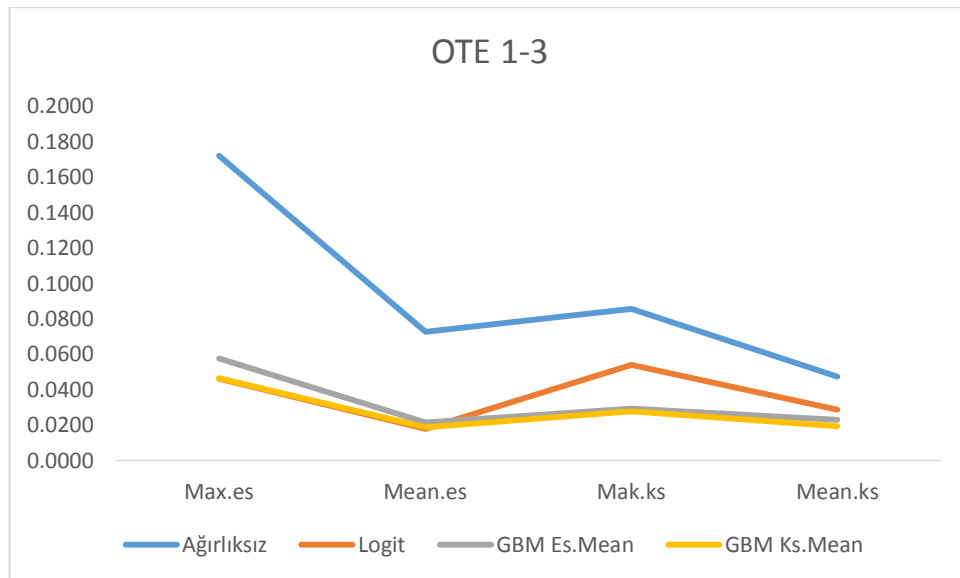
Tablo 5.25. MLR ve GBM'ye ilişkin özet denge değerleri (n=5448)

Yöntem	Grup	Tedavi n	Kontrol n	ESS Tedavi n	ESS Kontrol n	Mak.es	Mean.es	Mak.ks	Mean.ks
Ağırlıksız OTE	1-2	1354	1806	1354,00	1806,00	0,3884	0,1246	0,1732	0,0684
	1-3	1354	2288	1354,00	2288,00	0,1720	0,0727	0,0856	0,0472
Ağırlıksız TOTE	1-2	1354	1806	1354,00	1806,00	0,5528	0,1659	0,1739	0,0777
	1-3	1354	2288	1354,00	2288,00	0,6584	0,2443	0,3379	0,1268
MLR OTE	1-2	1354	1806	1219,79	1744,67	0,0420	0,0188	0,0420	0,0259
	1-3	1354	2288	1219,79	2271,06	0,0460	0,0178	0,0540	0,0286
MLR TOTE	1-2	1354	1806	1354,00	1530,51	0,0150	0,0065	0,0400	0,0216
	1-3	1354	2288	1354,00	2101,59	0,0200	0,0059	0,0590	0,0270
GBM Mean.es .OTE	1-2	1354	1806	1253,42	1697,86	0,0683	0,0240	0,0319	0,0236
	1-3	1354	2288	1253,42	2203,14	0,0577	0,0214	0,0292	0,0230
GBM Mean.ks OTE	1-2	1354	1806	1220,38	1689,01	0,0698	0,0258	0,0321	0,0216
	1-3	1354	2288	1220,38	2159,65	0,0465	0,0188	0,0277	0,0192
GBM Mean.es. TOTE	1-2	1354	1806	1354,00	1343,43	0,0162	0,0068	0,0179	0,0125
	1-3	1354	2288	1354,00	1640,01	0,0321	0,0198	0,0260	0,0191
GBM Mean.ks. TOTE	1-2	1354	1806	1354,00	1398,12	0,0184	0,0072	0,0186	0,0119
	1-3	1354	2288	1354,00	1727,65	0,0391	0,0200	0,0249	0,0178



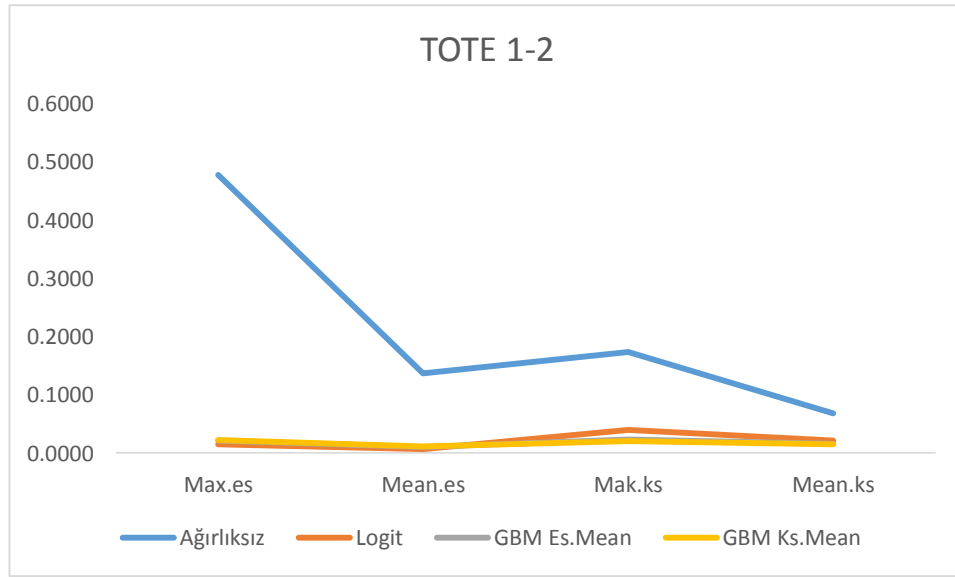
Şekil 5.25. 1'e karşı 2 tedavi grubunun OTE değerlendirmesi

Tablo 5.25 ve Şekil 5.25 incelendiğinde 1'e karşı 2. tedavi grubunun dengesinin değerlendirmesinde ağırlıklandırmadan önceki Mak.es değeri 0.3884 iken ağırlıklandırmadan sonra hem lojistik regresyon hem de GBM için 0.20'nin altında bir değere inmektedir. Maksimum ve ortalama standartlaştırılmış yanlılık değerleri incelendiğinde lojistik regresyonda daha iyi performans gösterirken, maksimum KS ve ortalama KS değerleri için GBM düşük değerler vermektedir.



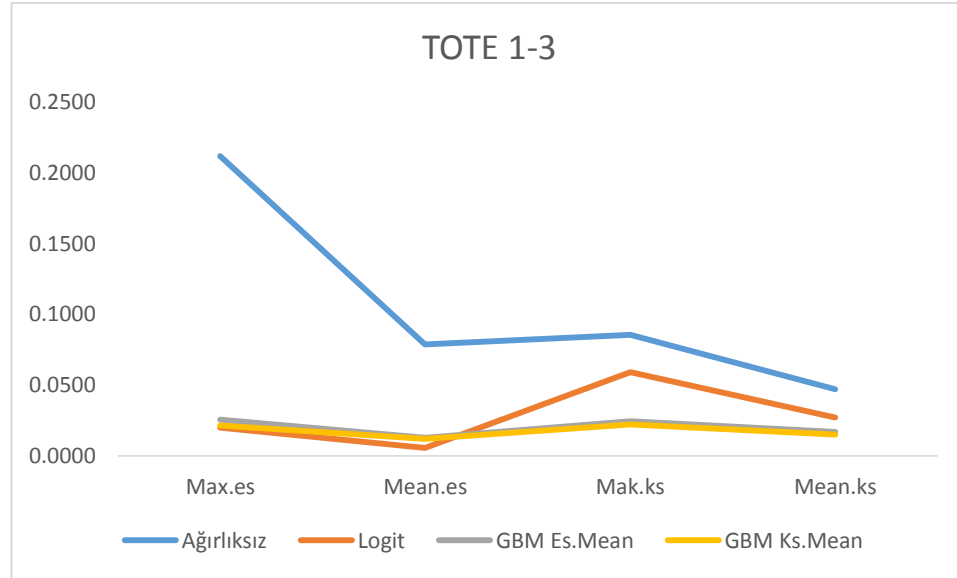
Şekil 5.26. 1'e karşı 3 tedavi grubunun OTE değerlendirmesi

Tablo 5.25 ve Şekil 5.26 incelendiğinde 1'e karşı 3 tedavi grubunun dengesinin değerlendirmesinde ağırlıklandırmadan önceki Mak.es değeri 0.1720 iken ağırlıklandırmadan sonra hem lojistik regresyon hem de GBM için 0.10'nun altında değere inmektedir. Maksimum ve ortalama standartlaştırılmış yanlılık değerleri incelendiğinde MLR, GBM ile yakın sonuçlar verirken, maksimum KS ve ortalama KS değerleri için GBM düşük değerler vermektedir.



Şekil 5.27. 1'e karşı 2 tedavi grubunun TOTE değerlendirmesi

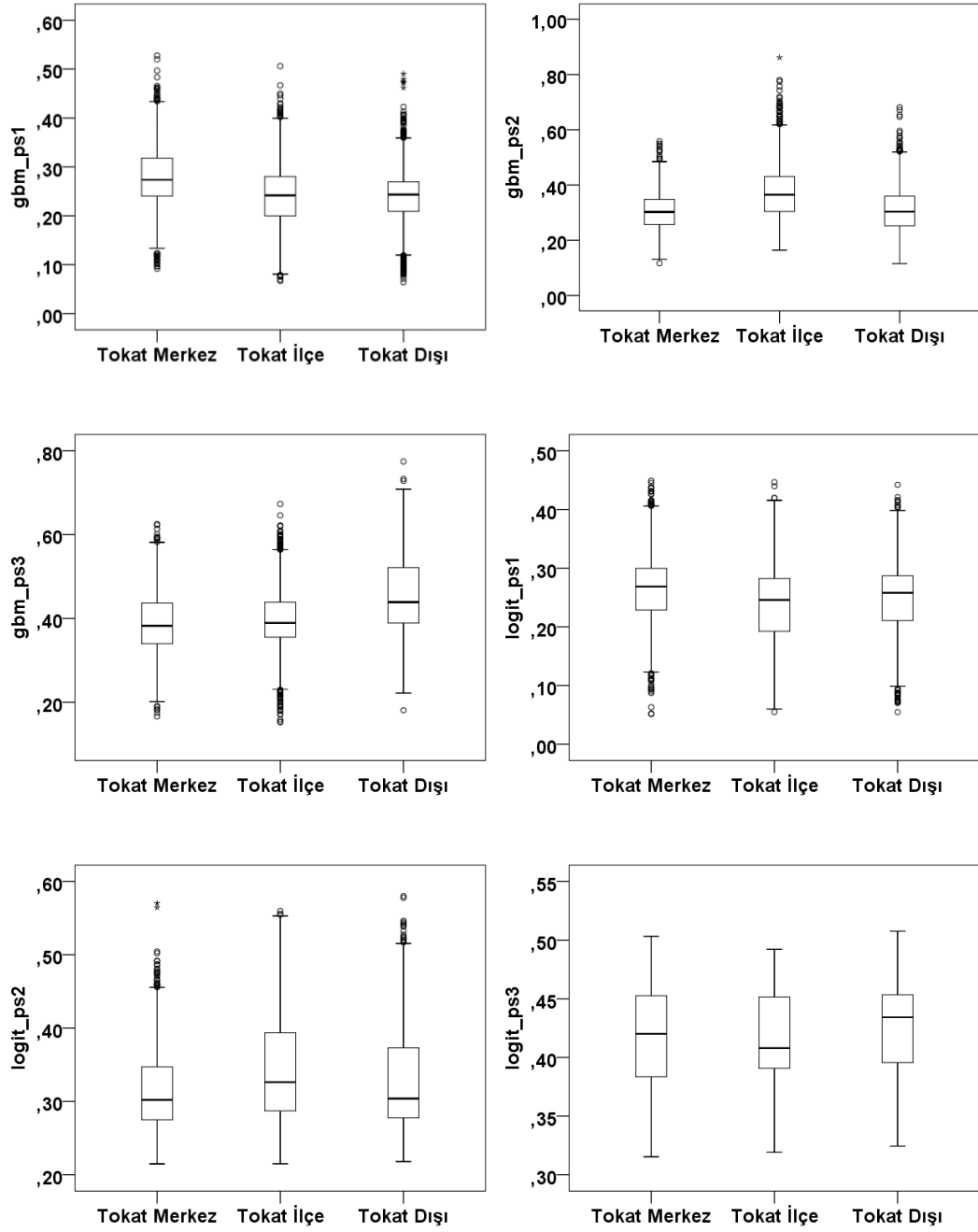
TOTE değerlendirmesi için Tablo 5.25 ve Şekil 5.27 incelendiğinde 1'e karşı 2. tedavi grubunun dengesinin değerlendirmesinde ağırlıklandırmadan önceki Mak.es değeri 0.5528 iken ağırlıklandırmadan sonra tüm değerler her iki yöntem için 0.10'nun altında bir değere inmektedir. Maksimum ve ortalama standartlaştırılmış yanlılık değerleri incelendiğinde MLR, GBM ile yakın sonuçlar verirken, maksimum KS ve ortalama KS değerleri için GBM düşük değerler vermektedir.



Şekil 5.28. 1'e karşı 3 tedavi grubunun TOTE değerlendirmesi

TOTE değerlendirmesi için Tablo 5.25 ve Şekil 5.28 incelendiğinde 1'e karşı 3. tedavi grubunun dengesinin değerlendirilmesinde ağırlıklandırmadan önceki Mak.es değeri 0.6584 iken ağırlıklandırmadan sonra tüm değerler her iki yöntem için 0.10'nun altında bir değere inmektedir. Maksimum ve ortalama standartlaştırılmış yanlılık değerleri incelendiğinde MLR, GBM ile yakın sonuçlar verirken, maksimum KS ve ortalama KS değerleri için GBM düşük değerler vermektedir.

5.3.2 Örtüşme (Overlap) Varsayımının MLR ve GBM için Görsel Değerlendirmesi



Şekil 5.29. Tedavi gruplarına göre GBM ve MLR ile kestirilen propensity skorlarına ilişkin kutu çizgi grafiği

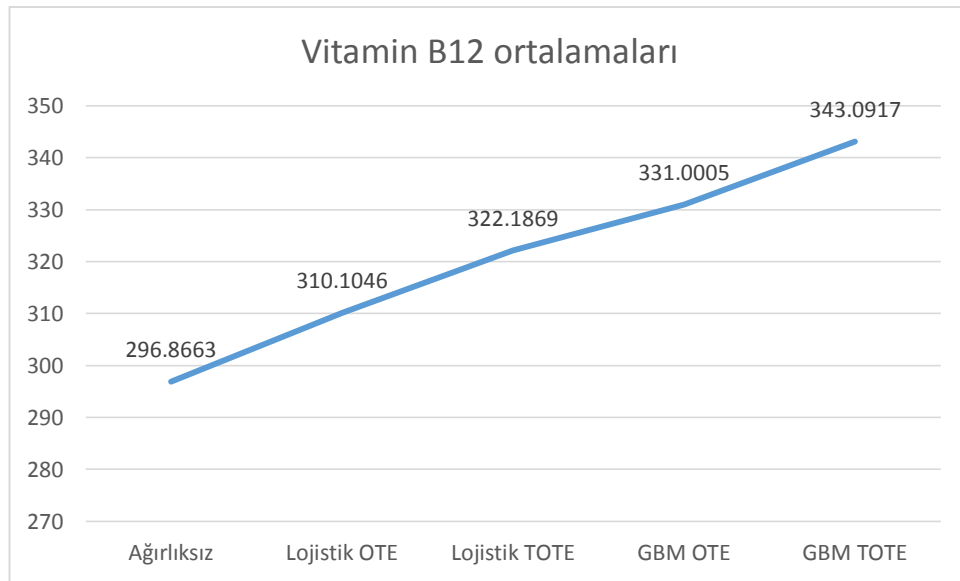
Şekil 5.29'dan görüldüğü gibi hem GBM'den hem de MLR'den elde edilen propensity skor değerlerine ilişkin dağılım gruplara göre benzer özellik göstermektedir. Dolayısıyla propensity skor tekniklerini uygulayabilmek için gerekli olan örtüşme ya da pozitiflik varsayımının yerine geldiği söylenebilir.

Tablo 5.26. Ağırlıklandırmadan önce ve sonra tüm değişkenlerin yanıt üzerinde tedavinin etkisi

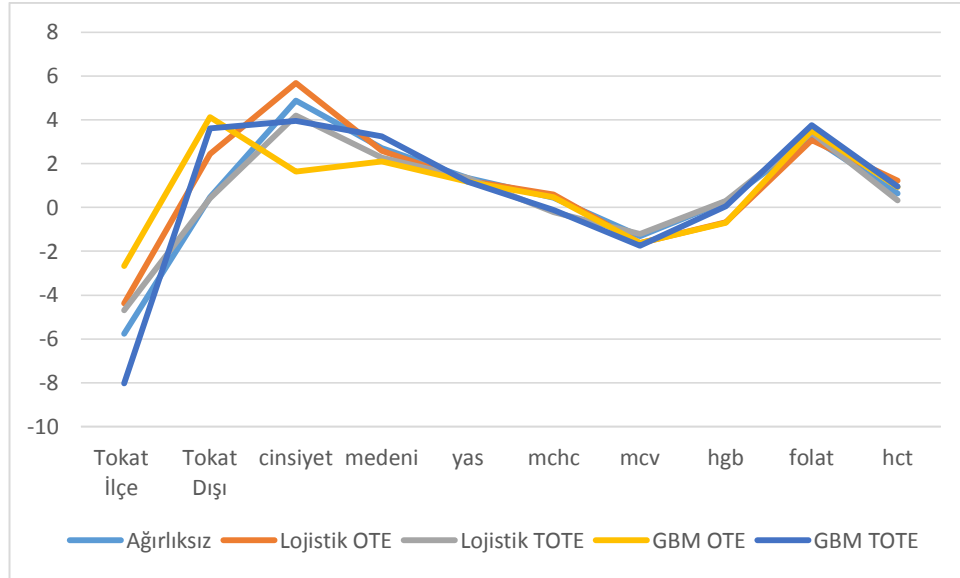
	Model	Katsayı	Standart hata	t	p
Ağırlıksız	Sabit	296,8663	45,0293	6,5927	<0,0001
	Tedavi Tokat İlçe	-5,7540	5,2146	-1,1034	0,2699
	Tokat Dışı	0,4603	4,9380	0,0932	0,9257
	Cinsiyet	4,8841	5,0149	0,9739	0,3301
	Medeni	2,7157	5,2135	0,5209	0,6025
	YAŞ	1,3496	0,1629	8,2828	<0,0001
	MCHC	0,4412	1,4604	0,3021	0,7626
	MCV	-1,3048	0,3158	-4,1315	<0,0001
	HGB	0,2633	1,7285	0,1523	0,8789
	FOLAT	3,2052	0,6203	5,1670	<0,0001
	HCT	0,6356	0,6467	0,9827	0,3258
Lojistik OTE	Sabit	310,1046	45,8779	6,7593	<0,0001
	Tedavi Tokat İlçe	-4,3739	5,2403	-0,8347	0,4039
	Tokat Dışı	2,4414	5,0582	0,4827	0,6294
	Cinsiyet	5,6726	5,9513	0,9532	0,3405
	Medeni	2,5939	5,4403	0,4768	0,6335
	YAŞ	1,2339	0,1923	6,4160	<0,0001
	MCHC	0,5836	1,5656	0,3728	0,7093
	MCV	-1,6227	0,4482	-3,6203	0,0003
	HGB	-0,6862	1,6829	-0,4077	0,6835
	FOLAT	3,0757	0,8341	3,6876	0,0002
	HCT	1,2147	0,7892	1,5393	0,1238
Lojistik TOTE	Sabit	322,1869	48,7588	6,6078	<0,0001
	Tedavi Tokat İlçe	-4,6953	5,4571	-0,8604	0,3896
	Tokat Dışı	0,4091	4,7644	0,0859	0,9316
	Cinsiyet	4,1911	5,7937	0,7234	0,4695
	Medeni	2,2776	5,7126	0,3987	0,6901
	YAŞ	1,3482	0,1917	7,0310	<0,0001
	MCHC	-0,2032	1,6202	-0,1254	0,9002
	MCV	-1,2014	0,3903	-3,0786	0,0021
	HGB	0,3028	1,7823	0,1699	0,8651
	FOLAT	3,4217	0,7223	4,7372	<0,0001
	HCT	0,3113	0,7302	0,4263	0,6699
GBM OTE	Sabit	331,0005	46,9358	7,0522	<0,0001
	Tedavi Tokat İlçe	-2,6817	5,2510	-0,5107	0,6096
	Tokat Dışı	4,1308	5,0128	0,8241	0,4099
	Cinsiyet	1,6337	5,8314	0,2802	0,7794
	Medeni	2,0999	5,4620	0,3845	0,7007
	YAŞ	1,1829	0,1968	6,0101	<0,0001

	MCHC	0,4591	1,5818	0,2902	0,7716	
	MCV	-1,6296	0,4160	-3,9176	0,0001	
	HGB	-0,6908	1,6987	-0,4067	0,6842	
	FOLAT	3,5020	0,7367	4,7534	<0,0001	
	HCT	0,9431	0,7523	1,2536	0,2101	
GBM TOTE	Sabit	343,0917	47,5414	7,2167	<0,0001	
	Tedavi	Tokat İlçe	-8,0331	5,1652	-1,5552	0,1200
		Tokat Dışı	3,6044	5,0174	0,7184	0,4726
	Cinsiyet	3,9393	5,8822	0,6697	0,5031	
	Medeni	3,2431	5,6585	0,5731	0,5666	
	YAŞ	1,1712	0,2071	5,6548	<0,0001	
	MCHC	-0,1080	1,6254	-0,0664	0,9470	
	MCV	-1,7431	0,3934	-4,4312	<0,0001	
	HGB	0,0432	1,5863	0,0272	0,9783	
	FOLAT	3,7643	0,7669	4,9084	<0,0001	
	HCT	0,9636	0,6998	1,3771	0,1685	

Yer tedavi (Tokat Merkez-Tokat ilçe-Tokat dışı) değişkeni göstermelik değişken olarak kodlanmaktadır. Tokat merkez referans olarak alınmaktadır.



Şekil 5.30. Ağırlıklandırma sonrası sabite ilişkin kestirimler (Tokat merkez referans)



Şekil 5.31. Ağırlıklandırma sonrası öncesi ve sonrası etkiler

Tablo 5.26 ve Şekil 5.31 incelendiğinde ağırlıklandırma sonrası kurulan modellerden elde edilen sabitler yani vitamin B12 kestirimleri ağırlıksız model için daha düşük iken (296,87) ortalamanın en yüksek sabit değerler GBM OTE (331) ve GBM TOTE'ye (343,09) aittir. Ağırlıklandırma sonrasında değişkenlerin anlamlılıklarına yönelik herhangi bir değişiklik bulunmamaktadır. Fakat düzeltme sonrasında I. tip hata olasılık değerlerinde büyümeler gözlenmektedir. Şekil 5.31'de ise vitamin B12 üzerinden tedavi grupları arasındaki farkı göstermektedir. Ağırlıksız model ile diğer yöntemlerin OTE ve TOTE'ye göre yakın değerler gösterdiği görülmektedir.

6 TARTIŞMA

Araştırmacılar tedavi etkilerini belirlerken, maliyet ve etik nedenlerden ya da geleneksel istatistiksel yöntemlerin kısıtlılıklarından dolayı alternatif yöntemlere gereksinim duyabilirler. Tedavi gruplarına atanan bireyler için rasgeleleştirmenin sağlanamadığı durumlarda, tedavi gruplarındaki bireylerin özelliklerinden kaynaklanan nedenlerden dolayı benzer gruplar oluşturamama gibi bir durumla da karşılaşılabilir. Dolayısıyla tedavi etkinliğini belirlerken ortaya çıkacak olan farklılığın bireylerin özelliklerinden değil de tedaviden kaynaklandığı yanlış bir sonuca varılabilir. Aynı şekilde tedavi etkinliğini belirlerken ortaya çıkması gereken bir farklılığın çıkmaması da tersi bir durum olarak görülebilir. Bu bireylerin gruplara atanmasındaki denge probleminden kaynaklanmaktadır.

Araştırmacıların rasgeleleştirmeyi sağlayamadığı çalışma düzenlerinde ya da gözlemsel çalışmalarda bireylerin özelliklerinden ve açıklanan diğer nedenlerden dolayı tedavi etkinliğini belirlemedeki yanlışlığı ortadan kaldırması için ortak değişkenler nedeniyle gruplar arasındaki farklılıkları küçültecek şekilde propensity skorlarının kullanılması önerilmektedir.

Gruplar arasındaki farklılığı gidermek için propensity skorlarını kullanan yöntemlerden eşleştirme, tabakalama ve propensity skoru ile regresyon düzeltmesi üzerine alan yazında pek çok çalışma mevcuttur. Son yöntem olan propensity skoru ile ağırlıklandırma ya da ters ağırlıklandırma olarak geçen yöntemin ülkemizde uygulamasının olmadığı görülmüştür. Ağırlıklandırma yöntemiyle çoklu tedavi için GBM ve MLR'nin performanslarını hem OTE hem de TOTE için değerlendiren yayın bulunmamaktadır.

Tedavi değişkeninin iki kategorili olduğu durumda propensity skor ağırlıklandırmasında yapay zeka algoritmalarının, lojistik regresyondan yanlışlığın azaltılması açısından daha iyi sonuçlar verdiği çalışmalar bulunmaktadır (6). Yapay zeka algoritmalarından olan GBM, iki kategorili tedavi değişkenin yanıt değişken üzerindeki etkinliğini belirlemede kullanılan oldukça esnek bir yöntemdir. Esnek olması kompleks ve doğrusal ilişkiler için içerisinde ortak değişkenleri kullanan

çoklu regresyon ağaçlarının olduğu modelleri kapsayan yinelemeli bir yöntem olması demektir (18).

Bu tez çalışmasında üç kategorili tedavi değişkeninin olduğu durumda propensity skor ağırlıklandırmasıyla yapay gruplar oluşturularak dengenin ve tedavinin yanıt üzerindeki etkisinin belirlenmesi amaçlanmıştır. Alan yazında gerçek veri seti ile GBM'nin üç kategorili tedavi değişkeni olduğu durumda da denge konusunda çok kategorili lojistik regresyondan daha iyi performans gösterdiğine ilişkin çalışmaları destekleyen sonuçlara ulaşılmaktadır (7). Genel olarak GBM'den elde edilen denge sonuçları daha tutarlıdır. Ana etkilerin olduğu senaryo altında çoklu lojistik regresyon daha başarılı denge ortaya koymaktadır. Genel olarak senaryolar kompleks hale geldikçe, GBM'nin denge konusunda daha iyi olduğu görülmektedir. Bu da daha önceki çalışmalarla uygunluk göstermektedir (43).

Örneklem genişliği arttıkça genel olarak her iki standartlaştırılmış yanlılık ve Kolmogorov Smirnov denge istatistiği değerlerinde de azalmalar görülmektedir. Tüm senaryolarda ağırlıklandırmadan önceki denge skorları ağırlıklandırıldıktan sonraki denge değerlerinden daha büyüktür.

Gözlemsel çalışmalarda tedavi etkinliğini belirlemek için iki kestirici kullanıldı. Bunlardan ilki ortalama tedavi etkisi (OTE) yani çalışmadaki her bir bireyin tüm tedavileri aldığı varsayımı altında kestirilen etkidir. Burada gözlenemeyen yanıtlar ilgili birey için karşıolgusallığı ifade etmektedir. Diğer ise tedaviyi alan bireyler arasında ortalama tedavi etkisi (TOTE) yani belirli bir tedaviyi alan bireylerin diğer bir tedaviyi aldığı durum altındaki kestirilecek ortalama tedavi etkisidir. Her iki kestirici için amaç tedavi gruplarına karşılık gelen yanıtlardaki farklılığı belirlemektir. Bu çerçevede çalışmamızda TOTE kestirimlerinden elde edilen denge değerleri OTE kestirimlerine göre daha yakın ve düşük değerlere sahiptir. Denge değerlerindeki bu fark arttıkça bireylerin gruplara atamadaki rasgeleliğin bozulduğunun bir göstergesidir.

Tedavilerin yanıt üzerindeki etkilerine ilişkin elde edilen Tablo 5.22-24 ve Şekil 5.22-24 incelendiğinde değişik senaryolar altında MLR ve GBM'ye ilişkin ağırlıksız, OTE ve TOTE sonuçlarına göre tedavinin yanıt üzerindeki etkisi önemli

iken önemli hale gelmektedir. Aynı şekilde ağırlıklandırmadan önce tedavinin etkinliğinin önemsiz iken önemli hale geldiği senaryolar da bulunmaktadır.

Gerçek veri seti uygulaması sonrası geleneksel istatistiksel yöntemlerden çoklu doğrusal regresyon yöntemi ile cinsiyet, medeni hal, yaş, MCHC,MCV,HGB,FOLAT, HCT değişkenlerinin ve tedavi değişkeninin (yer) vitamin B12 üzerinde etkisinin araştırıldığı çalışmada yaş, MCV ve FOLAT değişkenlerinin etkisi önemli bulunmaktadır. Ağırlıklandırma sonrasında değişkenlerin önemliliklerinde GBM ve MLR'ye göre bir değişiklik gözlenmemektedir. Aynı şekilde tedavilerin sabitlerine ilişkin kestirimlerde vitamin B12 ortalaması ağırlıksız düşük iken (296,87) ortalamanın en yüksek değerleri GBM OTE (331) ve GBM TOTE'ye (343,09) aittir. MLR'de ise OTE (310,10) ve MLR TOTE'ye (322,18) olarak kestirilmektedir.

7 SONUÇ VE ÖNERİLER

Bu tez çalışmasında, gözlemsel ya da yarı deneysel çalışmalarda gruplar arasındaki farklılığı en aza indirmek ya da gidermek için propensity skor ağırlıklandırma yardımıyla genelleştirilmiş boosted model ve çok kategorili lojistik regresyon yönteminin tedavi etkinliğini belirlemedeki performanslarının belirlenmesi ve tedavilerin etkinliklerinin karşılaştırılması üzerine durulmaktadır. Tıpta çeşitli nedenlerle çalışma düzenlerinin sağlıklı olarak belirlenememesi ihtimalini de göz önünde bulundurarak çalışmalardaki yanlılıkları gidermek büyük önem arz etmektedir. Bu bağlamda propensity skor ağırlıklandırma yöntemiyle bireysel nedenlerle grup farklılıklarını en aza indirip, tedavi etkinliğini belirlemek uygulamada sıklıkla kullanılabilecek bir yöntem olabilir. Dolayısıyla klasik istatistiksel yöntemlerdeki kısıtlılıklar nedeniyle yapılamayan analizler, propensity skor ağırlıklandırma ile yapılabilir.

Uygulamadaki veri seti için ağırlıklandırma sonrası önemliliklerinde bir değişiklik gözlenmese de benzetim çalışmaları sonrasında tedavinin yanıt üzerindeki etkisinin değiştiği gösterilmiştir.

Nedensel etki kestiriminde hangi kestiricinin kullanılması gerektiği iyi belirlenmesi gereken bir noktadır. Bir başka ifadeyle, OTE ve TOTE kestiricilerinden hangisinin kullanılacağına karar vermek önemlidir. Eğer tedaviye uygun tüm bireylerin kullanılacağı bir çalışmada bireylerin engel oluşturmayacak bir çalışma düzeninde OTE kestirimini kullanmak daha uygundur. Diğer taraftan tedaviye katılımın gönüllülük esasına dayalı olduğu bireylerin bu katılım için herhangi bir engelinin olmadığı durumda TOTE kestirimini kullanmak daha uygundur.

MLR ve diğer yapay zeka algoritmalarının kullanıldığı yöntemler kullanılarak tedavi etkinliği değerlendirilebilir. Kurulan modeller karmaşık hale geldikçe, GBM'nin MLR'den daha başarılı olduğundan, GBM'yi tercih etmek daha doğru olacaktır. Tedavi değişkeni için ikiden fazla kategori için bu analiz yapılabildiği gibi ortak değişkenlerin de ikiden fazla kategorili olan ortak değişken alınarak bu yöntemin nasıl performans gösterdiği belirlenebilir. Eksikliği görülen ve tedavi

grupları arasındaki farklılığı gidermeye çalışan propensity skoru ile ağırlıklandırma yöntemi, geniş bir uygulama alanı bulacaktır. R, SAS, Stata 13 ve sonrası programlar yardımıyla bu analizi gerçekleştirmek mümkündür.

KAYNAKLAR

1. Ellenberg, S.S., Fleming, T.R., DeMets, D.L. (2003). *Data monitoring committees in clinical trials: a practical perspective*: John Wiley & Sons.
2. Doll, R. (1998). Controlled trials: the 1948 watershed. *BMJ: British Medical Journal*, 317 (7167), 1217.
3. Hannan, E.L. (2008). Randomized clinical trials and observational studies: guidelines for assessing respective strengths and limitations. *JACC: Cardiovascular Interventions*, 1 (3), 211-217.
4. Alpar, R. (2010). Spor, Sağlık ve Eğitim Bilimlerinde Uygulamalı İstatistik ve Geçerlik-Güvenirlik: Birinci Baskı, Ankara: Detay Yayıncılık.
5. Austin, P.C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, 46 (3), 399-424.
6. Lee, B.K., Lessler, J., Stuart, E.A. (2010). Improving propensity score weighting using machine learning. *Statistics in medicine*, 29 (3), 337-346.
7. McCaffrey, D.F., Griffin, B.A., Almirall, D., Slaughter, M.E., Ramchand, R., Burgette, L.F. (2013). A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Statistics in medicine*, 32 (19), 3388-3414.
8. Gürtekin, B. (2011). ***Propensity Skor Yönteminin Meme Kanseri Yaş ile Dokudaki ER, PR, CEB-B2 Değişiminin Analizinde Kullanılması***. Sağlık Bilimleri Enstitüsü, Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, Biyoistatistik ve Demografi Programı, İstanbul Üniversitesi, İstanbul.
9. Kandemir, A.Ş. (2008). ***Gözleme dayalı çalışmalarda propensity skor ve sosyal bilimlerde bir uygulama***. Sosyal Bilimler Enstitüsü, Ekonometri Anabilim Dalı, İstatistik Bilim Dalı, Marmara Üniversitesi, İstanbul.
10. Altınok, E.Ç. (2006). ***Gözleme dayalı çalışmalarda propensity skor ve tıp bilimlerinde bir uygulama***. Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Yıldız Teknik Üniversitesi, İstanbul.
11. Rosenberger, W.F., Lachin, J.M. (2002). *Randomization in clinical trials: theory and practice*: John Wiley & Sons.

12. Sümbüloğlu, V., Alpar, R., Özdemir, P. (1998). Değişkenler Arası İlişkilerin İncelenmesi. *İç Hastalıkları Dergisi*, 5 (6), 416.
13. Aalen, O.O., Frigessi, A. (2007). What can statistics contribute to a causal understanding? *Scandinavian Journal of Statistics*, 34 (1), 155-168.
14. Rosenbaum, P.R., Rubin, D.B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70 (1), 41-55.
15. Rubin, D.B. (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of statistics*, 34-58.
16. StataCorp, L. (2013). *Stata Treatment-Effects Reference Manual*.
17. Söderbom, M. (2009) Applied Econometrics Lecture 11: Treatment Effects Part I.
18. McCaffrey, D.F., Ridgeway, G., Morral, A.R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological methods*, 9 (4), 403.
19. Alpar, R. (2011). *Uygulamalı Çok Değişkenli İstatistiksel Yöntemler (3. Baskı)*. Ankara: Detay.
20. Li, F., Zaslavsky, A.M., Landrum, M.B. (2013). Propensity score weighting with multilevel data. *Statistics in medicine*, 32 (19), 3373-3387.
21. Rosenbaum, P.R., Rubin, D.B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, 79 (387), 516-524.
22. Vansteelandt, S., Daniel, R. (2014). On regression adjustment for the propensity score. *Statistics in medicine*, 33 (23), 4053-4072.
23. Stone, C.A., Tang, Y. (2013). Comparing propensity score methods in balancing covariates and recovering impact in small sample educational program evaluations. *Practical Assessment, Research & Evaluation*, 18 (13), 1-12.
24. Cochran, W.G. (1968). The effectiveness of adjustment by subclassification in removing bias in observational studies. *Biometrics*, 295-313.
25. D'Agostino, R.B. (1998). Tutorial in biostatistics: propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Stat Med*, 17 (19), 2265-2281.

26. Austin, P.C., Mamdani, M.M. (2006). A comparison of propensity score methods: A case-study estimating the effectiveness of post-AMI statin use. *Statistics in medicine*, 25 (12), 2084-2106.
27. Williamson, E.J., Forbes, A. (2014). Introduction to propensity scores. *Respirology*, 19 (5), 625-635.
28. Cameron, A.C., Trivedi, P.K. (2005). Microeconometrics: methods and applications: Cambridge university press.
29. Arpino, B., Mealli, F. (2011). The specification of the propensity score in multilevel observational studies. *Computational Statistics & Data Analysis*, 55 (4), 1770-1780.
30. Elith, J., Leathwick, J.R., Hastie, T. (2008). A working guide to boosted regression trees. *Journal of Animal Ecology*, 77 (4), 802-813.
31. McCaffrey, D.F., Burgette, L.F., Griffin, B.A., Martin, C. (2015). Propensity Scores for Multiple Treatments.
32. Burgette, L., Griffin, B.A., McCaffrey, D. (2014). Propensity scores for multiple treatments: A tutorial for the mnps function in the twang package.
33. Ripley, W.N.V.a.B.D. (2002). Modern Applied Statistics with S (c. Fourth). New York: Springer.
34. Weisberg, J.F.a.S. (2011). An {R} Companion to Applied Regression (Second bs.). Thousand Oaks {CA}: Sage.
35. Eddelbuettel, D., Lucas, A., Tuszynski, J., Bengtsson, H., Urbanek, S., Frasca, M. ve diğ erleri. (2014). digest: Create Cryptographic Hash Digests of R Objects.
36. Lumley, T. (2004). Analysis of complex survey samples. *Journal of Statistical Software*, 9 (1), 1-19.
37. Analytics, R., Weston, S. (2011). doSNOW: Foreach parallel adaptor for the snow package. *R package*.
38. Analytics, R., Weston, S. (2013). foreach: Foreach looping construct for R. *R package version*, 1 (1).
39. Team, R.C., DebRoy, S., Bivand, R. (2013). foreign: Read Data Stored by Minitab, S, SAS, SPSS, Stata, Systat, Weka, dBase, 0.8-57.

40. Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*: Springer Science & Business Media.
41. Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev: Lawrence Erlbaum Associates, Inc.
42. Li, M. (2013). Using the Propensity Score Method to Estimate Causal Effects A Review and Practical Guide. *Organizational Research Methods*, 16 (2), 188-226.
43. Setoguchi, S., Schneeweiss, S., Brookhart, M.A., Glynn, R.J., Cook, E.F. (2008). Evaluating uses of data mining techniques in propensity score estimation: a simulation study. *Pharmacoepidemiology and drug safety*, 17 (6), 546.

EKLER

EK 1:

T.C.
GAZİOSMANPAŞA ÜNİVERSİTESİ
TIP FAKÜLTESİ DEKANLIĞI
Klinik Araştırmalar Etik Kurulu

Sayı : 83116987 - 372
Konu : Etik Kurul Kararı
Toplantı Tarihi : 25.08.2015
Toplantı No : 2015/13
Proje No : 15-KAEK-160

26.08.2015

Sayın. Yrd.Doç.Dr. Yalçın ÖNDER

Etik Kurulumuzun 25.08.2015 tarihli toplantısında görüşülen 15-KAEK-160 numaralı
“Çoklu Tedavi Etkinliğinin Karşılaştırılmasında Propensity Skor Ağırlıklandırması”
başlıklı çalışmanızın yapılmasında sakınca olmadığına karar verilmiştir.

Bilgilerinizi rica ederim.

Doç. Dr. Resul YILMAZ
Başkan



EK 2: Veri Türetim Senaryoları

1. Sadece ana etkilerin olduğu toplamsallık ve doğrusallığın olduğu model

$$P(T = t|X_i) = \frac{\exp\{\beta_{0t} + \beta_{1t}X_1 + \beta_{2t}X_2 + \beta_{3t}X_3 + \beta_{4t}X_4 + \beta_{5t}X_5 + \beta_{6t}X_6 + \beta_{7t}X_7\}}{1 + \sum_{h=1}^{M-1} \exp\{\beta_{0h} + \beta_{1h}X_1 + \beta_{2h}X_2 + \beta_{3h}X_3 + \beta_{4h}X_4 + \beta_{5h}X_5 + \beta_{6h}X_6 + \beta_{7h}X_7\}}$$
$$t = 1, \dots, M - 1$$

2. Bir karesel terimin olduğu model (hafif doğrusal olmayan model)

$$P(T = t|X_i) = \frac{\exp\{\beta_{0t} + \beta_{1t}X_1 + \beta_{2t}X_2 + \beta_{3t}X_3 + \beta_{4t}X_4 + \beta_{5t}X_5 + \beta_{6t}X_6 + \beta_{7t}X_7 + \beta_{2t}X_2X_2\}}{1 + \sum_{h=1}^{M-1} \exp\{\beta_{0h} + \beta_{1h}X_1 + \beta_{2h}X_2 + \beta_{3h}X_3 + \beta_{4h}X_4 + \beta_{5h}X_5 + \beta_{6h}X_6 + \beta_{7h}X_7 + \beta_{2h}X_2X_2\}}$$
$$t = 1, \dots, M - 1$$

3. Üç karesel terimin olduğu model (ılımlı doğrusal olmayan model)

$$P(T = t|X_i) = \frac{\exp\left\{\begin{aligned} &\beta_{0t} + \beta_{1t}X_1 + \beta_{2t}X_2 + \beta_{3t}X_3 + \beta_{4t}X_4 + \beta_{5t}X_5 + \beta_{6t}X_6 + \beta_{7t}X_7 \\ &+ \beta_{2t}X_2X_2 + \beta_{4t}X_4X_4 + \beta_{7t}X_7X_7 \end{aligned}\right\}}{1 + \sum_{h=1}^{M-1} \exp\left\{\begin{aligned} &\beta_{0h} + \beta_{1h}X_1 + \beta_{2h}X_2 + \beta_{3h}X_3 + \beta_{4h}X_4 + \beta_{5h}X_5 + \beta_{6h}X_6 + \beta_{7h}X_7 \\ &+ \beta_{2h}X_2X_2 + \beta_{4h}X_4X_4 + \beta_{7h}X_7X_7 \end{aligned}\right\}}$$
$$t = 1, \dots, M - 1$$

4. Üç tane iki yönlü etkileşim terimli hafif toplamsal olmayan model

$$P(T = t|X_i) = \frac{\exp\{\beta_{0t} + \beta_{1t}X_1 + \beta_{2t}X_2 + \beta_{3t}X_3 + \beta_{4t}X_4 + \beta_{5t}X_5 + \beta_{6t}X_6 + \beta_{7t}X_7 + 0.5\beta_{1t}X_1X_3 + 0.7\beta_{2t}X_2X_4 + 0.5\beta_{4t}X_4X_5 + 0.5\beta_{5t}X_5X_6\}}{1 + \sum_{h=1}^{M-1} \exp\{\beta_{0h} + \beta_{1h}X_1 + \beta_{2h}X_2 + \beta_{3h}X_3 + \beta_{4h}X_4 + \beta_{5h}X_5 + \beta_{6h}X_6 + \beta_{7h}X_7 + 0.5\beta_{1h}X_1X_3 + 0.7\beta_{2h}X_2X_4 + 0.5\beta_{4h}X_4X_5 + 0.5\beta_{5h}X_5X_6\}}$$
$$t = 1, \dots, M - 1$$

5. Dört tane iki yönlü etkileşim terimli ve bir karesel terimin olduğu hafif toplamsal ve doğrusal olmayan model

$$P(T = t|X_i) = \frac{\exp\left\{\beta_{0t} + \beta_{1t}X_1 + \beta_{2t}X_2 + \beta_{3t}X_3 + \beta_{4t}X_4 + \beta_{5t}X_5 + \beta_{6t}X_6 + \beta_{7t}X_7 + \beta_{2t}X_2X_2 + 0.5\beta_{1t}X_1X_3 + 0.7\beta_{2t}X_2X_4 + 0.5\beta_{4t}X_4X_5 + 0.5\beta_{5t}X_5X_6\right\}}{1 + \sum_{h=1}^{M-1} \exp\left\{\beta_{0h} + \beta_{1h}X_1 + \beta_{2h}X_2 + \beta_{3h}X_3 + \beta_{4h}X_4 + \beta_{5h}X_5 + \beta_{6h}X_6 + \beta_{7h}X_7 + \beta_{2h}X_2X_2 + 0.5\beta_{1h}X_1X_3 + 0.7\beta_{2h}X_2X_4 + 0.5\beta_{4h}X_4X_5 + 0.5\beta_{5h}X_5X_6\right\}}$$

$$t = 1, \dots, M - 1$$

6. 10 tane iki yönlü etkileşim terimli ılımlı toplamsal olmayan model

$$P(T = t|X_i) = \frac{\exp\left\{\beta_{0t} + \beta_{1t}X_1 + \beta_{2t}X_2 + \beta_{3t}X_3 + \beta_{4t}X_4 + \beta_{5t}X_5 + \beta_{6t}X_6 + \beta_{7t}X_7 + 0.5\beta_{1t}X_1X_3 + 0.7\beta_{2t}X_2X_4 + 0.5\beta_{3t}X_3X_5 + 0.7\beta_{4t}X_4X_6 + 0.5\beta_{5t}X_5X_7 + 0.5\beta_{1t}X_1X_6 + 0.7\beta_{2t}X_2X_3 + 0.5\beta_{3t}X_3X_4 + 0.5\beta_{4t}X_4X_5 + 0.5\beta_{5t}X_5X_6\right\}}{1 + \sum_{h=1}^{M-1} \exp\left\{\beta_{0h} + \beta_{1h}X_1 + \beta_{2h}X_2 + \beta_{3h}X_3 + \beta_{4h}X_4 + \beta_{5h}X_5 + \beta_{6h}X_6 + \beta_{7h}X_7 + 0.5\beta_{1h}X_1X_3 + 0.7\beta_{2h}X_2X_4 + 0.5\beta_{3h}X_3X_5 + 0.7\beta_{4h}X_4X_6 + 0.5\beta_{5h}X_5X_7 + 0.5\beta_{1h}X_1X_6 + 0.7\beta_{2h}X_2X_3 + 0.5\beta_{3h}X_3X_4 + 0.5\beta_{4h}X_4X_5 + 0.5\beta_{5h}X_5X_6\right\}}$$

$$t = 1, \dots, M - 1$$

7. 10 tane iki yönlü etkileşim terimli ve üç tane iki yönlü etkileşim terimli ılımlı toplamsal ve doğrusal olmayan model

$$P(T = t|X_i) = \frac{\exp\left\{\beta_{0t} + \beta_{1t}X_1 + \beta_{2t}X_2 + \beta_{3t}X_3 + \beta_{4t}X_4 + \beta_{5t}X_5 + \beta_{6t}X_6 + \beta_{7t}X_7 + \beta_{2t}X_2X_2 + \beta_{4t}X_4X_4 + \beta_{7t}X_7X_7 + 0.5\beta_{1t}X_1X_3 + 0.7\beta_{2t}X_2X_4 + 0.5\beta_{3t}X_3X_5 + 0.7\beta_{4t}X_4X_6 + 0.5\beta_{5t}X_5X_7 + 0.5\beta_{1t}X_1X_6 + 0.7\beta_{2t}X_2X_3 + 0.5\beta_{3t}X_3X_4 + 0.5\beta_{4t}X_4X_5 + 0.5\beta_{5t}X_5X_6\right\}}{1 + \sum_{h=1}^{M-1} \exp\left\{\beta_{0h} + \beta_{1h}X_1 + \beta_{2h}X_2 + \beta_{3h}X_3 + \beta_{4h}X_4 + \beta_{5h}X_5 + \beta_{6h}X_6 + \beta_{7h}X_7 + \beta_{2h}X_2X_2 + \beta_{4h}X_4X_4 + \beta_{7h}X_7X_7 + 0.5\beta_{1h}X_1X_3 + 0.7\beta_{2h}X_2X_4 + 0.5\beta_{3h}X_3X_5 + 0.7\beta_{4h}X_4X_6 + 0.5\beta_{5h}X_5X_7 + 0.5\beta_{1h}X_1X_6 + 0.7\beta_{2h}X_2X_3 + 0.5\beta_{3h}X_3X_4 + 0.5\beta_{4h}X_4X_5 + 0.5\beta_{5h}X_5X_6\right\}}$$

$$t = 1, \dots, M - 1$$

Yanıt modeli ise aşağıdaki gibi elde edilecektir.

$$Y = \alpha_0 + \alpha_1X_1 + \alpha_2X_2 + \alpha_3X_3 + \alpha_4X_4 + \alpha_5X_8 + \alpha_6X_9 + \alpha_7X_{10} + \gamma_1T$$

EK 3: Veri Türetiminde Kullanılan Katsayılar

Tedavi değişkeni için kullanılan katsayılar		Yanıt değişkeni için kullanılan katsayılar	
$\beta_{02} = -0,130$	$\beta_{03} = 1,051$	$\alpha_0 = -3,85$	$\gamma_7 = -0,4$
$\beta_{12} = -0,185$	$\beta_{13} = -1,455$	$\alpha_1 = 0,3$	
$\beta_{22} = -0,406$	$\beta_{23} = -0,496$	$\alpha_2 = -0,36$	
$\beta_{32} = -0,254$	$\beta_{33} = -1,088$	$\alpha_3 = -0,73$	
$\beta_{42} = -0,450$	$\beta_{43} = -0,691$	$\alpha_4 = -0,2$	
$\beta_{52} = 0,519$	$\beta_{53} = 1,192$	$\alpha_5 = 0,71$	
$\beta_{62} = 0,026$	$\beta_{63} = 0,836$	$\alpha_6 = -0,19$	
$\beta_{72} = -0,034$	$\beta_{73} = 0,461$	$\alpha_7 = 0,26$	

