YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

GÖZLEME DAYALI ÇALIŞMALARDA PROPENSİTY SKOR ve TIP BİLİMLERİ'NDE BİR UYGULAMA

İstatistikçi Elif Çiğdem ALTUNOK

FBE İstatistik Anabilim Dalında
Hazırlanan

YÜKSEK LİSANS TEZİ

Tez Danışmanı: Prof. Dr. Mehmet GENCELİ

İSTANBUL, 2006

İÇİNDEKİLER

Sayfa

SİMGE LİSTESİ

C_{0R}	Kontrol grubundaki örnek kovaryans matrisi
d	Fark
m_s	s tabakadaki tedavi birimlerinin sayısı
n_Λ	Aynı propensity skor değerine sahip birimlerin sayısı
n_s	s tabakadaki birim sayısı
Pr	Olasılık
s	Tabaka sayısı
u	Tedavi birimleri için eşleşen değişkenlerin değerleri
$x_{[j]}$	j. birim için ortak değişken
v	Kontrol birimleri için eşleşen değişkenlerin değerleri
$Z_{[j]}$	j. birim için tedaviye atanma
γ	Skaler bir değer
ε_i	Hata payı
$\lambda(x)$	Propensity Skor
Λ	Propensity skorun herhangi bir değeri
μ_{0M}	Eşleşmiş kontrol grubundaki x in beklenen değeri
$\pi_{[j]}$	j. birimin tedaviye atanma olasılığı
$\hat{\tau}$	Tedavi etkisi
Ω	Küme

KISALTMA LİSTESİ

EPBR	Equal-Percent Bias Reducing
PPskor	Propensity Skor
PS	Propensity Score

ŞEKİL LİSTESİ

Sayfa

ÇİZELGE LİSTESİ

Sayfa

ÖNSÖZ

Bu tezin hazırlanması sırasında öncelikle bana danışmanlık yaparak değerli bilgi ve desteğini esirgemeyen sayın Prof. Dr. Mehmet Genceli hocam olmak üzere, tezin başından sonuna kadar her konudaki yardımlarından ve desteğinden ötürü Marmara Üniversitesi Tıp Fakültesi Biyoistatistik anabilim dalı başkanı sayın Doç. Dr. Nural Bekiroğlu hocama, ayrıca çalışmalarında topladıkları verileri benimle paylaşarak, tezi pekiştirmemi sağlayan Marmara Üniversitesi Tıp Fakültesi Göğüs Cerrahisinden Anabilim dalı başkanı Prof. Dr. Mustafa Yüksel ile Yrd. Doç. Dr. Bedrettin Yıldızeli hocalarıma ve bu yoğun çalışma sırasında anlayış ve desteklerinden ötürü aileme çok teşekkür ederim.

ÖZET

Özellikle gözlemsel çalışmalarda, araştırmacının tedavi ve kontrol gruplarındaki birimleri rastlantısal olarak gruplara atama işleminde kontrolü yoktur. Bu nedenle, olgu-denetim gruplarına düşen bireylerin gerek demografik özelliklerinde gerekse ortak değişkenlerinde farklılıklar gözlemlenebilir ve bu farklılıklarda tedavi etkisinin sistematik hatalı tahminlerine neden olabilmektedir.

Bir dengeleme skoru olarak tanımlayabileceğimiz propensity skor, tedavinin gözlenen ortak değişkenlere göre koşullu olasılığı olarak ifade edilir ve gözlemsel çalışmalarda başlıca sistematik hatanın azaltılmasında, kesinliğin artmasında ve belirli ortak değişkenlerin etkilerini ortaya koymak amacıyla kullanılır. Propensity skor lojistik regresyon yardımı ile hesaplanır. Propensity skor bir kez tahmin edildikten sonra eşleştirme, tabakalara ayırma, regresyon düzeltmesi veya bu üçünün bileşiminin kullanılması yöntemleriyle sistematik hatanın azaltılması amaçlanır.

Çalışmamızda, Marmara Üniversitesi Hastanesi Göğüs Cerrahisi bölümünde 1996–2003 yılları arasında aynı doktor tarafından göğüs cerrahisi ameliyatı geçirmiş n=478 hasta kullanılmıştır. Ameliyat sonrası delirium tanısı alan ve almayan hastalara ait, ameliyattan öncesi 10 risk faktörü ile ve ameliyat sonrası 14 risk faktörüne lojistik regresyon uygulanmış ve sonuçlar elde edilmiştir. Propensity skor hesaplanmış, tabakalara ayırma yöntemi kullanılarak benzer propensity skora sahip tedavi ve kontrol bireyleriyle oluşan yeni örneğe istatistik analiz uygulanmış ve değerlendirilmiştir.

Propensity skor ile tabaklara ayırma yöntemi kullanılarak birbiri ile benzer dağılıma sahip tedavi ve kontrol bireyleri seçilmiş böylece yeni örneklemdeki tedavi ve kontrol grupları hemen hemen aynı karakteristiklere sahip olmuş ve sistematik hata azalmıştır. Sonuç olarak, propensity skor öncesi örneklem ile propensity skor sonrası tedavi ve kontrol bireylerine uygulanan lojistik regresyon sonucunda risk faktörlerinin anlamlılıklarında değişiklik gözlemlenmiştir.

Anahtar kelimeler: Propensity skor, gözleme dayalı çalışmalar, sistematik hatanın azaltılması

ABSTRACT

In observational and/or nonrandomized studies, investigators have no control on the random allocation of subjects to treatment and control groups. Because of this reason, some differences may be observed on the covariates and the demographic characteristics in patients who belong to case- control groups, therefore bias in the estimates might have accurred by these differences.

Propensity score, which can be determined as a balancing score, is the conditional probability of the observed covariates of case relative to control groups. Propensity score is used primarily to reduce bias in retrospective studies, increase the precision of the estimates as in the prospective studies and determine the effects of some covariates. Propensity score is calculated by using logistic regression analysis. Once the propensity score is estimated, it is aimed to reduce sampling bias by means of matching, stratification (quantiles), regression adjustment or combination of three methods.

This study included 478 patients who were operated by a single thoracic surgeon in Marmara University Hospital between 1996 and 2003. With regard to 10 preoperative risk factors and 14 postoperative risk factors based on clinical data, logistic regression was performed regarding 18 patients with postoperative delirium (POD) after thoracic surgery and 460 patients without POD. Then propensity score was calculated and patients who had similar propensity score in case and control groups were matched by using stratification methods, therefore some patients from case and control groups were excluded to form the new sample. After this new sampling, case-control groups had similar characteristics in risk factors and bias in retrospective studies was reduced.

As a result, the significance in risk factors, when logistic regression was used for the sample of case-control groups' patients before and after adjusting the propensity score, was observed to be different.

Keywords: Propensity score, observational studies, reduction bias

1. GİRİŞ

İstatistik yöntemleri günümüzde ölçmeye ve verilerin değerlendirilmesine dayanan çok çeşitli bilim dallarında kullanılmaktadır. İstatistiğin kullanıldığı en önemli alanlardan birini de Tıp Bilimleri oluşturmaktadır. Tıp Bilimlerinde araştırmalar genelde gözlem sonucu elde edilen verilerin değerlendirilmesi ile ilerleme kaydetmektedir. Gelişen bilgisayar programları ile de bugüne kadar çözümü zor hatta olanaksız birçok istatistik yöntem artık çözülebilir hale gelmiştir. Propensity skor da son on yılda gözleme dayalı çalışmalarda kullanılan önemli bir yöntem olup özelliği sistematik hatayı azaltması hatta düzeltebilmesidir. Tez konusu olarak seçilen bu konu ana hatları ile açıklanmaya çalışıldıktan sonra uygulama olarak Marmara Üniversitesi Hastanesinde yapılan bir araştırmanın verileri kullanılarak gözleme bağlı yöntemlerden propensity skor kullanılmıştır. Uygulama için propensity skor yönteminin tercih edilmesinin başka bir nedeni de diğer yöntemlerden farklı olarak anılan yöntemde ortak değişkenlerin sayısı bakımından herhangi bir sınırlama getirilmemesidir. Tezde propensity skor sonuçlarının rassal düzene yakın sonuçlar verdiği de açıklanmaya çalışılmaktadır.

2. GÖZLEME DAYALI ÇALIŞMALAR

2.1 Gözleme Dayalı Çalışmaların Ortaya Çıkış Nedenleri

Herhangi bir X bağımsız değişkenler kümesi ve herhangi bir Y bağımlı değişkeni arasındaki sebep-sonuç ilişkisini ortaya koyabilecek çeşitli modeller veya fonksiyonel ilişkiler söz konusudur. Oluşturulabilecek modeller arasında en basit ve en esnek olanları ise doğrusal modellerdir. Bağımlı değişkenin gözlenen değeri, bağımsız değişken değerlerinin tartılı toplamı artı hata payını içermesi halinde ise doğrusal modelden söz edilir.[*]

Ancak bazı ilişkiler basit bir doğrusal denklem ile açıklanamamaktadır. Bunun başlıca nedenleri olarak:

a) Bağımlı Değişkenin Dağılımı: Bağımlı değişkenin dikotom olduğu durumlarda, bağımlı değişken için yapılan öngörü değeri bağımlı değişkenin sahip olduğu dağılımı gerçekleyecektir. Bunun dışındaki öngörü değerleri ise mantık açısından olası değildir. Örneğin, bir araştırmacı üç mümkün sonuçtan birini tahmin etmek isteyebilir. Sözgelimi bireylerde yaş ile hastalık belirtileri arasındaki ilişki doğrusal bir ilişki ile ifade edilememektedir. Bu durumda, bağımlı değişken sadece 3 ayrı değer alabilir ve bağımlı değişkenin böyle dağılımına çok terimli dağılım denilir.

b) İlişki Fonksiyonu: Belirli bir ilişkinin her zaman çoklu regresyon modelleri ile açıklanamayacağının ikinci bir nedeni ise, örneğin insan yaşı ile hastalık belirtileri arasındaki ilişkide olduğu gibi bağımlı ve bağımsız değişkenler arasındaki ilişki gibi ilişkilerin doğrusal olmamasıdır.

sıralanabilir.

Regresyon analizi de doğrusal modelin bir alt kümesi olup, doğrusal modelden farklı olarak değişkenler genelde niceldir. Regresyon bağımlı, bağımsız değişken ayırımı yaparak bir bağımlı, bir veya birden fazla bağımsız değişken arasındaki ortalama ilişkiyi sebep-sonuç ilişkisine dayanarak ortaya koyan sistemdir. Doğrusal regresyon analizi, genelde, basit doğrusal model için kullanılan bir terimdir. Bu bağlamda doğrusal model varsayımlarının

---

[*] Hays, W. L., (1991), Statistics, Harcourt Brace Collage Publishers Fourth Edition, New York.

eklenmesi yoluyla doğrusal modelden doğrusal regresyon modeline geçilebilir. Modelin varsayımları da şu şekilde sıralanabilir:

— Ana kütle hata payları rassal değişken olup normal dağılmaktadır.

— Ana kütle hata paylarının beklenen değeri 0'dır.

—Ana kütle hata paylarına ilişkin varyans her alt ana kütle için aynıdır.

—Ana kütle hata payları ardışık değerleri birbirinden bağımsızdır.

$$Y_i = \beta_0 + \beta_i x_i + \varepsilon_i \qquad (2.1)$$

Yukarıdaki regresyon modelinde (2.1) β_0 ve β_i regresyon parametreleri bilinmediğinden parametreler veri kümesinden tahmin edilmelidir. Modeldeki açıklayıcı değişkenlerin ve/veya modelin biçiminin doğru seçildiği bilinmediğinden verilerin de analiz edilmeleri gerekmektedir.

Regresyon analizinde kullanılacak veriler deneysel veya deneysel olmayan niteliktedirler. Deneysel olmayan nitelikteki verilere örnek olarak gözleme dayalı veriler verilebilir.

Gözleme Dayalı Veriler: Gözleme dayalı veriler deneysel olmayan çalışmalardan elde edilen verilerdir. Deneysel olmayan çalışmalarda açıklayıcı değişkenler ve/veya bağımlı değişkenler üzerinde araştırmacının kontrolü bulunmamaktadır. Örneğin; bir firmada çalışanların yaşları (x) ile bir önceki yıldaki hastalıklı gün sayısı (Y) arasındaki ilişki araştırılmak istenildiğinde regresyon analizi için gerekli veriler birimlerin kayıtlarından sağlanmaktadır.

Gözleme dayalı çalışmalardaki başlıca kısıt, neden-sonuç ilişkisi hakkında genelde doğru bilgi sağlanamadığıdır.

Gözleme dayalı veriler temel alınarak oluşturulacak regresyon analizinde açıklayıcı değişkenler daha ayrıntılı araştırmalıdır, böylece neden-sonuç ilişkisi daha iyi açıklanabilir.

Deneysel nitelikteki veriye örnek olarak deneysel veri ve tam rastgele tasarım verilebilir.

Deneysel Veri: Deneysel çalışmalarda bağımsız değişkenler üzerinde kontrol sağlanmaktadır. Deneyde işleme girecek olan birimlere deneysel veriler olarak adlandırılmakta olup ve bu birimler işleme rastgele atanmaktadırlar. Sözgelimi bir firma, analizcilerin üretkenliği ile geçirdikleri eğitim süreleri arasındaki ilişki üzerinde çalışmak istemektedir. Bu çalışmada dokuz analizci kullanılmaktadır, bunlardan rastgele seçilmiş olan üçü iki hafta eğitilecek,

diğer üçü üç hafta, diğer bir üçü ise beş hafta eğitilecektir. Gelecek on hafta boyunca analizcilerin üretkenliği ölçülecektir. Eğitim süresi olan açıklayıcı değişken üzerinde kontrol sağlandığından elde edilen veriler deneysel verilerdir. Deneysel veri bize neden-sonuç ilişkisi hakkında, gözleme dayalı verilerden bağıl olarak daha çok bilgi verir, bunun nedeni de burada üretkenlik üzerinde analizcinin yeteneğinin de etkisinin olabileceği gibi, rastgeleliğin cevap değişkeni üzerinde etki edebilecek diğer değişkenler üzerinde denge oluşturulmasıdır.

Tam Rastgele Tasarım: Bu biçimdeki tasarım ile birimlerin araştırmaya atanmaları tamamen rastgele yapılır. Bu tam rastgeleliğin anlamı, her bir deneysel birimin herhangi bir işleme (treatment) girme şansı eşit olduğu anlamındadır.

2.2 Gözleme Dayalı Çalışmaların Ana Hatları

G. Cochran ilk olarak '' Observational Studies'' (Gözleme Dayalı Çalışmalar) başlığı altında gözleme dayalı çalışmaların istatistik yöntemlerini yayınlamıştır. 1965 yılında yazdığı bir raporda Cochran, ''Sigara içiminin erkeklerde akciğer kanserine neden olduğunun ve sigara içmenin kansere etki şiddetinin diğer bütün faktörlerden daha fazla olduğunu belirtmiş, kadınlar için elde edilen verilerden ise sigara içimi aynı doğrultuda daha az etkili olduğu'' sonucuna varmıştır.[*] Laboratuar hayvanları ile bazı deneyler yapılmış olmasına rağmen, sigara içimi ve insan sağlığı arasındaki ilişki için doğrudan kanıt gözleme dayalı veya deneysel olmayan çalışmalardan elde edilmiştir.

Bir deneyde, tedaviye (işleme) atananlar, deneyi yapan kişi tarafından kontrol altında tutulmaktadırlar. Deneyi yapan kişi farklı işleme gireceklerin ayırt edilebilir olmasını sağlar. Gözleme dayalı çalışmada ise, bir veya birkaç nedenden ötürü kontrol yoktur. Bunun nedeni örneğin, sigara içmenin veya radon gazının zararlı olduğu ve insanlara bunların deneyler için verilememesinden kaynaklanabilir. Birçok epidemiyolojik çalışma ve anket araştırmaları bu tip çalışmalara örnek oluşturmaktadır. Olaylar üzerinde, araştırıcının denetimi olmayarak (epidemiyolojik ender olaylar gibi) veya az denetimi olarak (anket gibi) yapmış olduğu araştırmalarda kullanılan genel bir terimdir. Bu tip çalışmalarda, risk etmenleriyle sonuç ölçümleri arasındaki ilişkinin incelenmesinde araştırıcının herhangi bir müdahalesi

[*] Cochran W. G., (1965),'' The Planning of Observational Studies of Human Populations'', Journal of the Royal Statistical Society, Series A, 128:134–155.

olmamaktadır.

2.2.1 Gözleme Dayalı Bazı Çalışmalar

Sigara içimi ve sağlık çalışmasında olduğu gibi gözlem yapılan çalışmalar önemli doğruları ortaya koymuştur ancak gözleme dayalı çalışmalar aynı zamanda araştırmacıları hatalı sonuçlara da ulaştırabilmektedir.

Gözleme dayalı çalışma ile bir deney çalışmasının karşılaştırılması için aşağıdaki örnek verilebilir.

1976 da Cameron ve Pauling , '' Proceeding of The National of Sciences'' da, gözleme dayalı verilerle ilerlemiş kanser tedavisi için C vitamininin tedavi amaçlı olarak kullanımı kapsayan bir makale yayınlayarak ileri derecede kanser olan 100 hastaya C vitamini vererek hayatta kalmalarını yayınlamışlardır.[*]

Her hasta için, aynı yaş ve cinsiyetten, aynı tip kanser hastalığına sahip ve aynı tür tümör cinsine sahip 10 tane kontrol deneği seçilmiştir. Kontrollerin bu şekilde seçilmesi yöntemine ''eşleştirilmiş örnekleme'' (matched sampling) denilmektedir. Yöntem ile tedaviye girenler için tedaviye girmeden önce ölçülmüş özelliklerinin en benzeri bir eş seçilmektedir. Etkili bir şekilde kullanıldığında da genellikle ''eşleştirilmiş örnekleme'' bazı yollardan hala birbirinden farklı olsalar bile tedavi ve kontrol gruplarının birbirinden ayrılmasını sağlamaktadır.

Cameron ve Pauling, standart tedaviler tarafından tedavisi mümkün olmayan bireylerde, C vitamini alanlar ile kontrollerin ölüm sürelerinin birbirinden ayrıldığını, grup olarak C vitamini alan hastaların kontrollerden dört kat fazla yaşadığını saptamışlardır. Bu fark geleneksel istatistik testlerinde anlamlı çıkarak, p değeri < 0.001 bulunmuştur. Bu şansa bağlı olarak meydana gelmiş bir durum olmadığından Cameron ve Pauling bu durumun tesadüfen meydana gelmediğini öne sürerek '' … C vitamininin tedavide kullanımının yaşamda kalma süresini uzattığı güçlü bir kanıttır…'' şeklinde bir hükme vardılar.

Çalışma C vitamininin tedavi olarak kullanımına olan ilgiyi arttırdı. Sonrasında, kolon ve

---

[*] Cameron, E. ve Pauling, L., (1976), ''Supplemental Ascorbate In The Supportive Treatment of Cancer: Prolongation of Survival Times In Terminal Human Cancer'', Proceedings of the National Academy of Sciences (USA), 73: 3685–3689

rektumdan ileri kanser olan hastalar için C vitamini alanlar ve plasebo (içinde etki maddesi bulunmayan ilaç) alanları ayırt ederek rassal kontrollü bir deney yapıldı. Rassal bir deneyde birimler tedavi veya kontrole şans mekanizması temel alınarak atanır yani tedaviye girilmesi sadece şans faktörüne bağlıdır. Deneyin sonunda yapılan testlerde C vitamininin hayat süresini uzattığına dair bir gösterge bulunamamış, ayrılmış gruplarda hayatta kalma süreleri az miktarda, kesin olmayan ve anlamlı kabul edilmeyen şekilde uzun çıkmıştır. Nitekim günümüzde de C vitaminin kanseri tedavi ettiğine ilişkin bir kanıt bulunmamaktadır.

Gözleme dayalı çalışmada kullanılan kontroller birkaç önemli değişkenle eşleşmesine rağmen hayatta kalmak için önemli olan bazı yollar bakımından tedaviye giren hastalardan farklı olduğu açık ve seçiktir. Deney ile gözleme dayalı çalışma arasındaki önemli fark deneyde tedaviye rastgele atamaların olmasıdır. Deneyde, her bir hasta gruba rastlantısallık kullanılarak kontrol ve tedavi gruplarına ayrılırlar. Rassal deneylerde ise, bireylerin farklı tedavilere rassal olarak atanmaları, farklı tedavilere atananlar arasındaki gözlenen ve gözlenmeyen ortak değişkenlerde sistematik hatanın mümkün olduğu kadar azalmasını sağlamaktadır. Rastgele atamalarda kötü şans faktörü oluşabilmektedir, esas olarak tedavi ve kontrol grupları önemli yollardan farklı olabilir ancak kötü şans faktörünün potansiyel etkisini ve onu tedavinin bir etkisinden ayırt etmek zor değildir. Bilinen istatistik testler ve güven aralığı bu farkları ayırt edebilmekte, daha doğrusu burada ifade edilmek istenen gözleme dayalı çalışmada oluşan farkın ‘şans’ faktörü yüzünden oluşmadığıdır.

Gözleme dayalı çalışmada birimler, deneyi yapan tarafından rastgele bir numara kullanılarak tedaviye veya kontrole atanmamaktadırlar. Gözleme dayalı çalışmalarda rastlantısal olarak atanmayan tedavi ve kontrol gruplarının gerek demografik özelliklerinde gerekse gözlenen ortak değişkenlerinde büyük farklar olabilir. ‘‘Eşleştirilmiş örnekleme’’ iki grubun önemi az yollardan ayrıştırılmasını sağlamakta, ancak bunun dışındaki ayrılabilmeyi garanti edememektedir. Gruplar tedaviye girmeden önce ayrılamıyorsa ve önemli yollardan farklar varsa, bu durumda hayatta kalmadaki farklar bu başlangıç farklarının yansımasından başka bir şey olmayacaktır.

İlerlemiş kanser tedavisi için C vitamini tedavi olarak kullanımı isimli gözleme dayalı çalışma da, çalışmaya başlarken kontrol grubundaki hastaların kayıtları ölülerden, tedaviye giren hastaların kayıtları ise yaşayanlardan alınmıştır. Bu çalışmada tedavide olan hastalar, kısa bir süre sonra ilerlemiş kanserden dolayı ölmüş ve tedavi alan şimdi ölü olan hastaların kayıtları C vitamini tedavisini almamış hastaların hayatta kalma durumlarında kullanılmıştır. Yapılan deneyde ise böyle bir durum söz konusu değildir tedavi ve kontrol grubundaki bütün

hastaların kayıtları deneyin başında yaşarlarken alınmıştır.

C vitamini çalışmalarından sonra elde edilen ilk sonuçlar ise; gözleme dayalı çalışmalar ve deneyler çok farklı sonuçları sağlayabileceği şeklinde olmuştur. Bu gerçekleştiğinde ise deneyler daha güvenilirdir. İkinci olarak gözleme dayalı çalışmada eşleştirme ve benzer düzeltmeler yapılarak gruplar ayırt edilebilmekte, ancak tedavi ve kontrol grubunun ayrılmasını garanti edememektedir. Üçüncü olarak kontrollü bir deneyde rastlantısallık kullanılır ancak gözleme dayalı çalışmada rastlantısallık bulunmamaktadır.

Deney yapmak mümkün olmadığında gözleme dayalı çalışmalar yol göstermektedir. Gözleme dayalı çalışmaların ve deneylerin doğrudan mukayesesinde ortak yön az bulunmaktadır. Kanser ve C vitamini çalışması bu nedenle birer istisnadır.

Gözleme dayalı çalışmaya bir diğer örnek de sigara içimi ve kalp hastalığı üzerine yapılan çalışmadır.

Doll ve Hill (1966) kalp hastalığı olan ve çeşitli sigara içimi davranışına sahip olan İngiliz doktorların ölüm oranları üzerinde çalışmışlardır.[*] Bu çalışma sonucunda, akciğer kanseri ile sigara içimi arasında ciddi bir ilişki bulunurken, kalp hastalığı ile sigara kullanımı arasında düşük bir ilişki bulunmuştur. Araştırmacılar kalp hastalığı riski arttığında ölüme sebep olmasına rağmen, kalp hastalığı bilinen genel ölüm nedenlerinden uzak olduğu sonucuna varmışlardır.

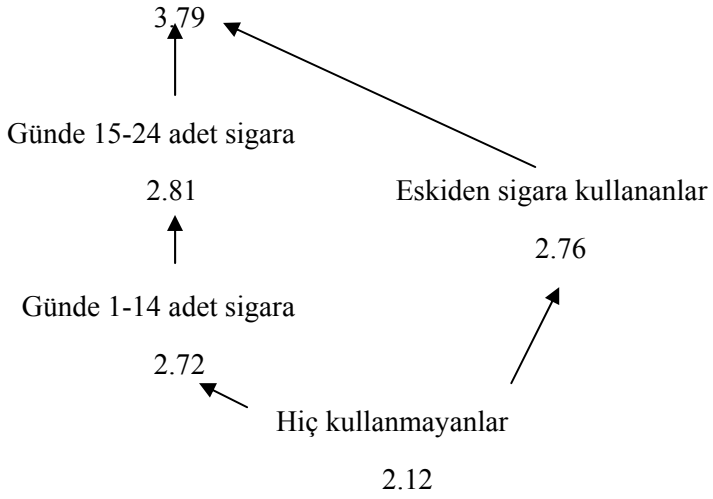
Doll ve Hill'in çalışmalarının başlangıcında yaptıkları ''yaş ayarlaması'' (adjust for age) olmuştur. Çalışmalarında yaşlılar kalp hastalığında gençlere göre daha fazla risk altındadır. Çok sayıda sigara kullanan genç insan ve çok sayıda sigara kullanmayan yaşlı insan olmasına rağmen, grup olarak sigara kullananlar, kullanmayanlara göre nedense daha yaşlılardan oluşmuştur. Sigara kullanan ve kullanmayanları doğrudan ayırt etmek için, yaşı görmezden gelerek oldukça yaşlı grubu oldukça genç gruba ayırt edilmiş, böylelikle sigaranın kullanım etkisi olmasa bile kalbe ait ölüm oranlarında bir fark görüleceği umulmuştur. Burada ''yaş ayarlaması'' aynı yaştaki sigara kullanan ve kullanmayanları ayırt etmek içindir. Genellikle faklı yaşlardaki sonuçlar ayarlanmış-yaş ölüm oranı denilen tek bir sayıda birleşirler. Doll ve

---

[*] Doll, R. ve Hill, A., (1966), '' Mortality of British Doctors In Relation to Smoking: Observations on Coronory Thrombosis'', Epidemiological Approaches to the Study of Cancer and Other Cronic Diseases, 205-268.

Hill'in çalışmasında ayarlanmış-yaş ölüm oranındaki farklar, yaşlardaki farklarla yorumlanamamaktadır çünkü aynı yaştaki sigara kullanan ve kullanmayanların ayrılması sonucu meydana gelmişlerdir. Yaş ve diğer değişkenler için bu tür düzenlemelerin yapılması gözleme dayalı veri analizinin ana konusunu oluşturmaktadır.

Doll ve Hill ikinci olarak, sigara içimi kalp rahatsızlığına neden olduğunda bunun sonuçlarını araştırmışlardır. Sigara içenler de ölümlerin artması kesinlikle beklenmektedir.



Şekil 2.1 Günde 25 veya 25 den fazla sigara kullanımı ile kalp hastalığı ölümü arasındaki ilişki

Şekil 2.1 de Doll ve Hill tarafından başka önemli hastalıkla ilişkili olmadan, kalp hastalığından ölenler için 6 ayrı ayarlanmış-yaş ölüm oranları verilmiştir. Bu altı grup sigara kullanmayanlar, sigarayı bırakanlar ve halen günde 1–14, 15–24 adet sigara kullananlar ve ≥ 25 sigara kullananlardan oluşmaktadır. Doll ve Hill ilginç ve daha detaylı tahminlere olanak vermesine rağmen, burada sigarayı bırakanları eskiden içtikleri sigara miktarlarına göre ayırmamışlardır. Yine de, yaşlardaki fark, bu ölüm oranlarını etkilememiştir. Oranlar her yıl için 1000 ölüdür, 3.79 değerinin anlamı her 1000 doktordan her yıl yaklaşık dört ölüm olduğudur.

İçilen sigara adedi ile ölüm oranlarının artığı görülmüştür (Şekil 2.1). Alternatif bir iddia da sigara kullanımının kalp hastalığından ölüme neden olmadığı ancak kendi sonuçları ile karşılaştırmaya yaradığıdır.

2.3 Gözleme Dayalı Çalışmalarda Yapılan Hatalar ve Düzeltme Yolları

2.3.1 Gözleme Dayalı Çalışmalarda Bilinen Sistematik Hata

Tedavi ve kontrol grupları çalışmadaki çıktıları etkileyecek şekilde tedaviye başlamadan önce önemli yollardan farklılık gösteriyorsa gözleme dayalı çalışma için bu durumda sistematik hatalı (yanlı) denilir. Sistematik hata; $\hat{\theta}$'nın matematik ümidi $E(\hat{\theta})$ ile θ parametresi arasındaki pozitif veya negatif fark olarak tanımlanmaktadır. Bilinen (açık) sistematik hata mevcut verilerden anlaşılabilir. Örneğin; tedaviye başlamadan önce, tedaviye girecek olan bireyler kontrollerden daha az girdiye sahip olduklarını varsayacak olursak, burada bir bilinen (açık) sistematik hata (overt bias) olduğunu söylenebilir. Saklı (hidden) sistematik hata ise buna benzerlikle beraber gerekli bilgi gözlenemediğinden veya kayıt edilemediğinden, bu tarz sistematik hataya saklı sistematik hata denilmektedir. Açık sistematik hatalar, eşleştirme ve tabakalara ayırma gibi düzeltmeler yapılarak kontrol edilebilirler. Diğer bir deyişle, tedavi ve kontrol birimleri gözlenen ortak değişkenler bakımından farklılık gösterebilmekte, ancak bu görülebilen farklar tedavi ve kontrol birimlerini gözlenen ortak değişkenlerin aynı değerleri ile ayırt ederek ortadan kaldırılabilmektedir.

Bilinen (açık) sistematik hata ve bunun düzeltilerek ortadan kaldırılmasını ilk olarak Cochran 1968'deki çalışmasında ortaya koymuştur.[*] Cochran bu çalışmada Best ve Walker'ın yaptığı araştırma verilerini kullanarak 3 gruba ayrılmış olan erkeklerin ölüm oranlarını incelenmiştir. Bu gruplar, sigara kullanmayan, sigara kullanan, puro ve/veya pipo kullananlardan oluşmaktadır. Sigara kullanmayanların ölüm oranı her yıl için 1000 kişide 20,2 ölüdür, sigara kullananların ölüm oranı ise 20,5 ve puro ve/veya pipo kullananlar için ise bu oran 35,5 ölüdür. Bu durumda sigaranın zararsız olduğu bunun yanında hem puronun hem de piponun çok tehlikeli olduğu söylenebilir. Cochran her grup için yaş ortalamalarını vermiştir. Sigara kullanmayanların yaş ortalaması 54,9, sigara kullananların 50,5 ve puro ve pipo kullananlar için yaş ortalaması 65,9 dur. Ortalamalardan açıkça görüldüğü gibi puro ve/veya pipo kullananlar yaşlılardan oluşmaktadır, bu durumda ölüm oranlarında ki farklılık şaşırtıcı olmamalıdır çünkü bu puro ve piponun etkisini yansıtmayabilir. Diğer taraftan, sigara kullananlar en genç gruptur ve sigara kullanmayanlar daha yaşlı olmasına rağmen ölüm oranı sigara kullananlardan daha yüksek çıkmış ve sigaranın zararsız olmadığı saptanmıştır.

---

[*] Cochran, W. G., (1968), ''The Effectiveness of Adjustment By Subclassification In Removing Bias In

Cochran daha sonra yaş için ölüm oranlarını düzenleyerek ortak değişkendeki dengesizliği düzeltmiş, böylece çıktıdaki açık sistematik hatayı ortadan kaldırmıştır. Yaş, erkekleri 3 küme veya tabakaya ayırt etmek için kullanmış, bu şekilde aynı kümedeki erkekler benzer yaşa sahip olmuşlardır. Aynı yaştaki sigara kullanmayanların, sigara kullananların ve puro ve pipo kullananların her biri kümelere ayrılmış ve sonuçlar doğrudan düzeltme kullanılarak tek bir oranı oluşturmuştur. Düzeltilmiş ölüm oranları sigara kullanmayanlar için her yıl için 1000 kişide 20,3 ölüdür, sigara kullananlar için 28,3 ölü ve puro ve pipo kullananlar için 21,2 ölüdür. Bunun sonucunda da sigaranın tehlikeli olduğu iyice ortaya çıkmıştır.

Çalışmada sigara kullanımının etkilerini tahmin etmek için düzeltilmemiş oranlar yanlış sonuçlar vermektedir. Düzeltilmiş oranlar ise sigara kullanımının etkilerini doğru tahmin edebilmişlerdir. Yaşa etki yapabilecek kayıt edilmemiş başka ortak değişkenler olması da mümkündür, böyle bir durumda saklı sistematik hatadan söz edilebilecektir.

Cochran çalışmasında 3 yaş tabakası kullanmıştır. Bu 3 tabakanın yeterli olup olmadığı Cochran'ın çalışmasının ana sorununu oluşturmaktadır. Üç tabaka yerine, 12 tabaka kullanıldığında düzeltilmiş oranlar sigara kullanmayanlar için 20,2, sigara kullananlar için 29,5 ve puro ve/veya pipo kullananlar için 19,8 şeklinde bulunmuştur ve üç tabaka ile on iki tabakanın benzer düzeltilmiş oranlar ürettikleri saptanmıştır. Cochran bireylerin %20 sini içeren her beş tabakanın, yaş gibi sürekli bir ortak değişkendeki sistematik hatanın %90 ını ortadan kaldırdığı şeklinde kuramsal bir tartışmayı bu çalışmasında başlatmıştır.

2.3.2 Bilinen Sistematik Hata İçin Düzeltmelerin Planlanması

Bilinen sistematik hatanın kontrolü çalışma şekillenmeden önce başlamaktadır. Çalışma tasarımında gözlenen ortak değişkenlerdeki bilgilerin birleşmesi ile örneğin eşleştirilmiş örnekleme yaparak veya tedavi etkisinin tahmininde tabakalara ayırma veya kovaryans düzeltmesi yapılması suretiyle bu farklardan kaçınılabilmektedir.

Planlamadaki ilk adım gözleme dayalı çalışmada hangi tedavilerin yapılacağının belirlenmesi ve bu süreçte ortak değişkenlerden çıktıların ayırt edilmesidir. Ortak değişkenler etkilenmediği varsayımı altında çıktılar tedavi tarafından etkilenebilecek nicelikleri ölçmektedir.

Rosenbaum'un, plasebo (içinde etken madde bulunmayan ilaç) ve ilacı ayırarak kan

basıncının düşürülmesi için yaptığı çalışmasını örnek verilebilir [*]; Burada çıktı nabzın oranıdır. Burada gruplar tedaviye başladıktan 6 ay sonra kan basıncı seviyeleri için düzeltme yapıldıktan sonra ayırt ediliyorsa, nabzın düzeltilmiş oranları ilaç ve plasebo grupları için birbirine benzer olacaktır çünkü ilacı kullanan grupta, ilaç kan basıncını düşürmüş ve nabzı azaltmıştır. İlacın kan basıncı üzerindeki etkisi ortadan kalkarsa, nabız üzerindeki etkisi de onunla birlikte kalkmış olacaktır.

Bu örnekten de görülebileceği gibi, bir çıktı için yapılan düzeltmeler tedavi etkisinin bir kısmını ortadan kaldırmaktadır ancak tedavinin belli bir çıktıda önemsiz etkilerinin olması da şüphe uyandırabilmektedir, çünkü bu çıktıda ölçülmemiş başka önemli bir ortak değişken ile güçlü şekilde ilişkili olabilmektedir.

Planlamadaki diğer bir adım ise ölçülecek ortak değişkenlerin listesini oluşturmaktır. Bu aşamada sistematik hata, hem bilinen (açık) , hem de saklı sistematik hatadan oluşmaktadır. Tam anlamıyla saklı bir sistematik hata olmasa bile, bu listede yapılan küçük bir değişiklik çalışmanın inandırıcı olup olmadığını belirleyebilmektedir. Planlama aşamasında kolaylıkla düzeltilebilecek küçük bir dikkatsizlik, daha sonraki aşamalarda düzeltilemeyecek problemler yaratabilmektedir. Rassal klinik deneylerin tasarımında, standart uygulama ile toplanacak verileri ve uygulanacak analizleri açıklayan yazılı bir protokol hazırlanmakta ve deney başlamadan önce eleştirici yorumlar için bu protokol araştırmacılar arasında dolaştırılmaktadır. Gözleme dayalı çalışmalarda da bu yazılı protokollerden ve eleştirici yorumlardan faydalanılmaktadır.

Bilinen sistematik hata için yapılan düzeltmeler veri analizinden çok verinin toplanmasında başlamaktadır. Genellikle tedaviye giren bireyler, kontrol bireyleri ile çift veya eşleşmiş kümeler oluşturmak için eşleştirilirler, böylece gözlenen ortak değişkenlerde ayrılabilecek olanlar gözlemlenir. Eşleştirme çıktılar ölçülmeden önce yapılabilmektedir. Burada üzerinde durulması gereken üç nokta vardır. Birincisi, analitik düzeltmelerden farklı olarak, çalışma tasarımında yapılan eşleştirme tabakalara ayırma gibi yapılan düzeltmeler değiştirilemezdir. Plasebo ve ilacın nabız atışını düşürülmesi için yapılan deney örneğinde olduğu gibi tedaviden sonra kan basıncı için düzeltme yapmak yanlış olacaktır. Bu hata analitik bir metot kullanılarak yapılmış olsaydı, başka bir analiz uygulanarak düzeltilebilirdi ancak hata

---

[*] Rosenbaum, P. R., (1995), Observational Studies, Springer-Verlag , New York.

eşleştirilmiş örnekleme kullanılarak yapıldığında bunu düzeltmek zorlaşacaktır.

İkinci olarak, tasarımda eşleştirme yapılması analitik düzeltmelerden daha kolay belli ortak değişkenlerde kontrol sağlanmaktadır. Bunlar tipik olarak birimleri birçok küçük kategoride sınıflayan ortak değişkenlerdir. Eşleştirme yapmak tedavi ve kontrol birimlerinin aynı kategorilerde bulunmasını temin eder ancak eşleştirme çalışmanın tasarımında kullanılmazsa, bazı kategoriler kontrol birimleri olmadan tedavi birimlerini ya da tedavi birimleri olmadan kontrol birimlerini içerebilmektedir.

Üçüncü olarak, maliyet eşleştirme yapıp yapmama konusuna karar vermede diğer önemli bir etkendir. Bazı ortak değişken bilgileri kolayca elde edilebilir, fakat diğer verileri elde etmek pahalı ve zor ise bu durumda eşleştirme cazipliğini yitirmektedir. Birçok çalışmada bazı kontrol birimleri tedavi birimlerinden çok farklı olabilir ki bu karşılaştırma için kullanışlı bir durum değildir. Eşleştirme yapılarak daha sonra az kullanılışa sahip olacak kontrollerdeki bu verilerden kaçınılabilmektedir.

Eşleştirilmiş çiftlerin seçiminde her tedavi bireyini potansiyel birden çok kontrol ile eşleştirilebilir. Her tedavi bireyinin birkaç kontrol bireyi ile eşleştirilmesinde, her bir tedavi bireyi için dörtten fazla kontrol bireyi ile eşleşmesinin çalışmaya kazanç sağlamadığı sonucuna varılmıştır. [*]

Ortak değişkenden veri toplamak bazı sorunları da beraberinde getirmektedir. Tüm gözlenen ortak değişkenler için düzeltme yapılıp yapılamayacağı sorusu bu sorunlardan biridir. Esas olarak, tedaviye girmeden önce bireyleri tanımlayan bir değişken olan gerçek bir ortak değişken için düzeltmeden kaçmak için az veya çok bir neden bulunmamaktadır. Rastlantısallık deneyde tedavi ve kontrol gruplarının ilgili ve ilgili olmayan tüm ortak değişkenlerin ayrılabilir olmasını sağlamaktadır. Uygulamada durum genelde ortak değişkenlerin (düzeltmelerde kullanılanlar) sayısının artması ile maliyet karmaşıklığı artacak ve bu durumda en önemli ortak değişkenler için düzeltmelerin yapılması zorlaşacaktır.

Diğer bir sorun da, verinin kalitesi ve bütünlüğüdür. Daha çok ortak değişken toplandığında

---

[*] Ury, H., (1975), ‘’Efficiency of Case-Control Studies With Multiple Controls Per Case: Continuous or Dichotomous Data’’, Biometrics, 31: 643–649.

ve analiz edildiğinde, tüm ortak değişkenlerin doğruluk ve bütünlüğünün yüksek standartlara ulaşmasını garanti etmek zorlaşacak ve her bir ortak değişkenin model kurmada ve eşleştirmede gerekli dikkati almasını garanti etmesi zorlaşacaktır. Her birinde çok sayıda kayıp veri olan ortak değişken olduğunda, tüm ortak değişkenlerde tam veriye sahip birey az olacak ve bu analiz yapmayı oldukça güçleştirecektir.

Ortak değişkenleri seçmek için kullanılan en yaygın yöntemler oldukça tartışılmıştır. t testi kullanılarak ve sadece farkların anlamlı çıktığı ortak değişkenler için düzeltme yapılarak uzun ortak değişken listesinden tedavi ve kontrol grupları ayrılmaktadır. Burada üç problem vardır. Birincisi, test ortak değişken ile çıktı arasındaki ilişki dikkate almamaktadır. İkinci olarak, istatistiksel anlamlılığın bulunmaması ortak değişkendeki dengesizliğin görmezlikten gelinecek kadar küçük olması anlamına gelmesine inanmak için hiçbir neden söz konusu değildir. Üçüncü olarak, düzeltmeler aynı zamanda tüm ortak değişkenleri kontrol ederken, test her seferinde bir ortak değişkeni dikkate almaktadır. Cochran (1965) kolay bir regresyon modeli altında tüm niceliklerin normal dağıldığı ve tek bir ortak değişkenin sistematik hatanın kaynağı olduğu durumda bu tekniği uygulamıştır. ‘’Tek bir ortak değişken 1,5 üzerinde bir t değeri gösteriyorsa, çıktıların değerleri bilindiğinde bu sonuçlar için tekrar ortak değişkenlere bakılması tavsiye edilir.’’ sonuca varmıştır.[*]

Aşağıda anlatılacak olan yaklaşım genellikle akla yatkın ve pratiktir. Başlangıçta düzeltmeler için geçici bir ortak değişken listesi seçilir. Tedavi ve kontrol gruplarında ortak değişkenlerin araştırmaya yönelik karşılaştırılmaları ile geçici ortak değişkenlerden elde edilen bilgi kullanılarak düzeltmeler yapılır. Geçici liste yardımıyla düzeltme için geçici yöntemler belirlenir. Bunlar, eşleşmiş çiftlerin veya kümelerin seçimi, tabakayı tanımlamak ve hangi teknik kullanılacaksa onu belirtmektir. Geçici listeye dâhil edilmeyen ortak değişkenlere de bu düzeltme yöntemi uygulanarak, düzeltmeden sonra büyük dengesizlik sergileyen ortak değişkenler belirlenir. Bu analizin ışığında ortak değişkenlerin geçici listesinin tekrar gözden geçirilmesi gerekmektedir.

---

[*] Cochran W.G., (1965),’’ The Planning of Observational Studies of Human Populations’’, Journal of the Royal Statistical Society, Series A, 128:134–155.

2.3.3 Tabakalara Kesin Ayırma ve Eşleştirme

Araştırmanın başlangıcında, çeşitli değerler alabilen, gözlenen X ortak değişkeninin sadece tek değer alabilen M birimleri bulunmaktadır. Söz konusu X ortak değişkenleri genelde, X için eşleştirme veya tabakalandırma amacıyla verilerin analiz öncesi yeniden düzenlenmesinde kullanılmaktadır. X 'lerde eşleştirme veya tabakalara ayırma yapmak tekrar düzenleme yapmaya örnek olarak verilebilir. M birim $j = 1,.....,m$ ise $x_{[j]}$, j. birim için ortak değişkendir ve bu birim için tedaviye atama $Z_{[j]}$ dır. Köşeli parantez deki $[j]$ ifadesi birimlerin düzenlemeden önceki birimlerin numarasını ifade etmektedir. Düzenleme yapıldıktan sonra birim, köşeli parantez olmayan bir ifadeye sahip olacaktır.

Gözleme dayalı bir çalışma için j'inci birimin $\pi_{[j]} = prob\left(Z_{[j]} = 1\right)$ olasılıkla tedaviye ve $\left(1 - \pi_{[j]}\right)$ ile de kontrole atandığını varsayalım. Kuşkusuz $\pi_{[j]}$, $0 < \pi_{[j]} < 1$ arasında yer almaktadır. Modelde, tedaviye atamaların hatalı para atışı ile yapıldığını düşünülürse, muhtemelen her birim için farklı sistematik hataya sahip para atılmış olacak ve paraların sistematik hataları veya π leri bilinmediği durumda, model aşağıdaki gibi olacaktır.

$$prob(Z_{[1]} = z_1,...,Z_{[M]} = z_M) = \prod_{j=1}^{M} \pi_{[j]}^{z_j} \left\{1 - \pi_{[j]}\right\}^{1-z_j} \tag{2.2}$$

Gözleme dayalı çalışmada, $\pi_{[j]}$ ler bilinmediğinden, $Z_{[1]},...,Z_{[M]}$ lerin dağılımı da bilinmemekte ve rastlantısallığın yarattığı bilinen tedavi atamalarının dağılımını ortaya koymak burada mümkün olmamaktadır.

Gözleme dayalı çalışmada modelimizde bilinen sistematik hatanın olduğunu, saklı sistematik hatanın olmadığı varsayılırsa, π lerin bilinmediği fakat gözlenen ortak değişkenler $x_{[j]}$ e bağlı olarak bilindiği durumda gözleme dayalı çalışmada saklı sistematik hata yoktur denilmektedir. Böylece x'in aynı değerine sahip iki birim, aynı π tedaviye atama şansına sahip olacaktır. $\pi_{[j]} = \lambda\left(x_{[j]}\right)$ $j = 1,.....,m$ de olduğu gibi şekli bilinmeyen bir $\lambda(.)$ fonksiyonu varsa, bu durumda çalışma da saklı sistematik hata yoktur denilir. Eğer çalışma da saklı sistematik hata yoksa bu durumda (2.2) yerine,

$$prob(Z_{[1]} = z_1,...,Z_{[M]} = z_M) = \prod_{j=1}^{M} \lambda\left(x_{[j]}\right)^{z_j} \left\{1 - \lambda\left(x_{[j]}\right)\right\}^{1-z_j} \tag{2.3}$$

yazılabilecektir.

Kısaca, (2.3) denklemi sağlandığında gözleme dayalı çalışmada saklı sistematik hata yoktur denilir.

Çalışmada saklı sistematik hata olmadığında $\lambda(x)$ fonksiyonuna propensity skor denilir. Propensity skor saklı sistematik hatanın olduğu durumlarda da kullanışlı bir yöntemdir. Rubin (1977) propensity skoru ortak değişkenler temelinde rastgeleleştirme olarak tanımlamıştır.*

Saklı sistematik hatanın bulunmadığı, fakat açık sistematik hatanın varlığı halinde $\lambda(x)$ bilinmediğinden durumu düzeltmede kullanılan en basit yaklaşım tabakalara ayırma yöntemidir.

Burada anlatılan düzeltmeler bilinen sistematik hata olduğu, saklı sistematik hata olmadığı durumdaki modeller içindir.

Tabakalara Ayırma:

Genellikle, birimler x ortak değişken temel alınarak tabakalara gruplanırlar. M birimden, $n \le M$ birim seçilir ve onları s tabakada n_s birim ile S birbiri ile örtüşmeyen tabakalara gruplanırlar. N birimin seçim ve onları tabakalara atanması için sadece x ler ve rastgele sayılar tablosunu kullanılmaktadır. Birimler tekrar numaralandırılmakta böylece s tabakasındaki i. birim Z_{si} tedavi atamasına ve x_{si} ortak değişkenine sahip olmaktadır. N-sıra için Z, $\left(Z_{11},...,Z_{S,n_s}\right)^T$ olarak yazılır. s tabakadaki tedavi birimlerinin sayısı için m_s yazılır. Burada $m_s = \sum_i Z_{si}$ ve $m = \left(m_1,...,m_s\right)^T$ dir.

X'i tabakalara kesin ayırmak, x de homojen olan tabakalara kesin ayırmanın sonucunda tabakalar homojen birimlerden oluşmaktadır, böylece iki birim sadece aynı x değerine sahip ise, tüm s, i ve j için $x_{si} = x_{ji}$ ise, aynı tabakadandır denilir. Tabakalara kesin ayırma, sadece x az boyutlu ve koordinatları süreksiz ise kullanışlıdır, diğer durumda aynı x değerine sahip çok sayıda birimi kullanmak zorlaşacaktır.

---

* Rubin, D.B., (1977), '' Assignment to The Treatment Group on The Basis of A Covariate'', Journal of Educational Statistics, 2:1–26

X'i tabakalara kesin ayırmada, çalışmada saklı sistematik bulunmadığı ve (2.3) denkleminin geçerliliği halinde, bu durumda aynı tabakadaki tüm birimler aynı tedaviye atanma şansına sahip olacaklardır. Bu durumda $\lambda(x_{si})$ yerine λ_s yazılabilir ve (2.3) denklemi şöyle olur:

$$prop(Z=z)=\prod_{s=1}^{S}\prod_{i=1}^{n_s}\lambda_s^{z_{si}}\left(1-\lambda_s\right)^{1-z_{si}} \tag{2.4}$$

(2.4) de, tedaviye atamanın dağılımı $prop(Z=z)$, λ_s bilinmediğinden ve $m_s=\sum_i Z_{si}$ bir rastgele değişken olduğundan ötürü bilinememektedir. m verildiğinde Z'nin koşullu dağılımını düşünecek olursak; bu dağılım, sadece $s=1,...,S$ için $m_s=\sum_i Z_{si}$ olduğu durumda, $z\in\Omega$ de elemanları 0'ın ve 1'in N-sırası olan bir Ω kümesindeki dağılımdır. Böylece Ω nin $K=\prod_{s=1}^{S}\binom{n_s}{m_s}$ elemanı vardır ve her tedaviye atama $z\in\Omega$, (2.3) deki aynı koşulsuz olasılığa,

$$prop(Z=z)=\prod_{s=1}^{S}\lambda_s^{m_s}\left(1-\lambda_s\right)^{n_s-m_s} \tag{2.5}$$

sahip olacaktır.

Verilen koşullu olasılık m sabit olduğunda $prop(Z=z\,|\,m)=1/K$ dır. Bu da tekdüze bir rassal deneyde Z'nin dağılımını vermektedir.

Kısaca gözleme dayalı bir çalışmada, saklı sistematik hata bulunmaması ve x'ler tam olarak tabakalara ayrılabildiği durumlarda, her tabakada, tedavi birimlerinin sayısı m verildiğinde tedaviye atamanın koşullu dağılımı Z, prop(Z=z|m), tekbiçim rassal deneydeki tedaviye atanmanın dağılımı ile aynı olacaktır. Tedaviye atanma olasılıkları $\lambda(x)$ bilinmediğinde dahi bu durum geçerli olacaktır. Çalışmada saklı sistematik hata yoksa ve x tabakalara ayrılırsa, çalışma bir tekbiçim rassal deney için kullanılan yöntemler ile analiz edilebilmektedir.

Özetle, gözleme bağlı bir çalışma saklı sistematik hatasız olup x tam olarak tabakalara ayrıldığı takdirde, her tabakadaki işlem gören (tedavi) m birimin veri olduğu işlemine Z'nin koşullu dağılımı $prop(Z=z\,|\,m)=1/K$ tekdüze bir rastgele deneydeki işlem dağılımının aynıdır.

Eşleştirme:

Tabakalarda tedavi birimlerinin m_s sayısı üzerinde ve kontrol birimlerinin $n_s - m_s$ sayısı üzerinde kısıtlamalar olması eşleştirme ile tabakalara ayırma arasındaki başlıca farklılıktır. Örneğin, çift eşleştirmede her s için $n_s = 2$ ve $m_s = 1$ gerek duyulmakta iken çoklu kontrol ile eşleştirme için ise $n_s \geq 2$ ve $m_s = 1$ söz konusudur.

X'de eşleştirme yapmak;

1) S, m ve n'e kısıt uygulayarak
2) Tabakalarda x'in seyrine dayanan kısıtları karşılayabilen bir tabakalalaştırma

elde edilen örnekleri oluşturabilmesine bağlıdır.

Örneğin; s=1,…,100 için $n_s = 2$ ve $m_s = 1$ ile tüm mümkün tabakalara ayırma düşünülerek ve tüm mümkün tabakalardan , tabakalardaki x'lere ve rastgele sayılara dayanarak birinin seçilmesi ile , S=100 çift ile bir çift eşleşmiş örnek şekillenmiş olacaktır. Bir tam eşleştirme her bir eşleştirilmiş kümede tüm n_s birim için x'in aynı olduğu eşleştirmedir, bu da her s için $i, j = 1,...,n_s \quad x_{si} = x_{sj}$ dir. Tam tabakalara ayırma ile tam eşleştirme mümkün olabilmekte ancak bu durum x in az boyutlu ve süreksiz olduğu durumda olanaklıdır.

(2.3) denklemindeki gibi, gözleme dayalı çalışmada sistematik hata bulunmuyor ve biri tam olarak x de eşleşiyorsa bu durumda verilen m ile tedaviye atamanın koşullu dağılımı prop(Z=z$\mid$ m)=1/K , Z tekbiçim rassal deneylerinkinin aynısı olacaktır. Çalışmada saklı sistematik hatanın da bulunmaması halinde bir eşleştirilmiş rassal deneymiş gibi analiz edilebilmektedir.

Çalışma tasarımda gözlenen ortak değişkenlerdeki bilgilerin birleşmesi ile örneğin eşleştirilmiş örnekleme yaparak veya tedavi etkisinin tahmininde tabakalara ayırma veya kovaryans düzeltmesi yaparak bu farklardan kaçınılabilir. Geleneksel düzeltme yöntemleri (eşleştirme, tabakalara ayırma, kovaryans düzeltmesi) genellikle sınırlıdır, çünkü düzeltme için sadece sınırlı sayıda ortak değişken kullanabilmektedir.

3. PROPENSİTY SKOR

Ortak değişken sayısı p sayıda arttıkça, aynı veya benzer x değeri ile eşleştirilmiş çiftler bulmak zorlaşacaktır. Her bir ortak değişken bir ikili değişken olursa, x in 2^p olası değeri olacak, bu durumda p=20 ortak değişken ile x'in bir milyondan fazla değeri olacaktır. Yüzlerce ve binlerce birim ve p=20 ortak değişken olduğunda, çok sayıda birim, x'in eşsiz büyük değerini alacaktır. Bundan ötürü, çok sayıda ortak değişken olduğunda, eşleştirme yapmak olanaklı değildir.

Eşleştirme ve tabakalara ayırmanın X'de homojen olan eşleştirilmiş kümeleri oluşturmasının yanında iki hedefi vardır. Birincisi, saklı sistematik hata olmadığında x için düzeltme yapmak yeterlidir, böylece düzeltme için tabakalar veya eşleştirilmiş kümeler oluşturmak için alışıla gelmiş rassal deneylerde kullanılan metotlar kullanılmaktadır. İkincisi, saklı sistematik hata olsun ya da olmasın, kontrol ve tedaviye giren grupların eşleştirilmiş bireyler x den farklı değerlere sahip olsalar bile x in benzer dağılımları ile ayrılmak istenmektedir. Bu ikinci amaca ise ortak değişken dengesi denilir.

Gerçek propensity skor değeri bilindiğinde daha önce sözü edilen iki amaca propensity skor da eşleştirme veya tabakalara ayırma yapılarak ulaşılmaktadır. Bu, tabakalar veya eşleştirilmiş kümeler x de heterojen olsalar bile propensity skorda homojen şekilde şekillenirler.

3.1 Propensity Skorun Tanımı

Tedavinin x gözlenen ortak değişkenlere göre koşullu olasılığı olarak, tedaviye atamalar için $\pi_{si} = prob(Z_{si} = 1)$ ve $0 < \pi_{si} < 1$ ile (2.3) denkleminden, s tabakasındaki tüm n_s bireyler için propensity skor;

$$\lambda(x_s) = \frac{\sum_{i=1}^{n_s} \pi_{si}}{n_s} \tag{3.1}$$

şeklinde tanımlanmıştır.

π_{si}, $0 < \pi_{si} < 1$ sağladığından, her s için $0 < \lambda(x_s) < 1$ olur.

Propensity skorun $\lambda(x_s)$, işleyişi şöyle yorumlanabilir: s tabakasından rastgele bir birim seçilir, her birimin seçilme olasılığı $1/n_s$ dir. Bu rastgele birim $\lambda(x_s)$ olasılık ile tedaviyi alınır. Bu $1/n_s$ olasılık ile (s,i) den birimin seçilmesi anlamındadır. Burada (s,i) π_{si} olasılık

ile tedaviye alınır, böylece bu rastgele birimin tedaviyi almasının marjinal olasılığı $\sum_{i=1}^{n_s} \pi_{si} / n_s = \lambda(x_s)$ şeklinde olacaktır.

Propensity skorun iki özelliği bulunmaktadır. Birincisi, saklı sistematik hata olmadığında yani her s tabaka ve i birim için $\pi_{si} = \lambda(x_s)$ olması halinde söz konusudur. Saklı sistematik hata bulunmadığında x_s de homojen olan tabaka ve eşleştirilmiş kümelerin oluşturulmasına gerek yoktur; $\lambda(x_s)$ de homojen olan tabaka ve eşleştirilmiş kümelerden elde etmek yeterlidir. Saklı sistematik hata yoksa ve $\lambda(x_s)$ de tabakalar homojen ise, tedaviye atananların koşullu dağılımları tek biçimdir ve rassal deneyler için kullanılan istatistik yöntemleri kullanılabilir. x_s çok boyutlu olabilir ancak $\lambda(x_s)$ sadece bir sayıdır, genellikle $\lambda(x_s)$ in benzer değerleri ile birimleri bulmak, x_s in benzer değerleri ile bulmaktan daha kolaydır. Saklı sistematik hata olmadığında sadece x_s e bağlı olarak bilinen sistematik hata olduğunda propensity skor $\lambda(x_s)$ için düzenleme yapmak yeterlidir.

İkinci özellik ise, saklı sistematik hata olsun veya olmasın, $\pi_{si} \neq \lambda(x_s)$ olsa bile propensity skorun geçerliliğidir. $\lambda(x_s)$ homojen olan tabakalar veya eşleştirilmiş kümeler x_s'i dengeleme eğiliminde olup, bir anlamda aynı tabaka veya aynı eşleştirilmiş kümedeki tedavi ve kontrol birimleri x_s'in aynı dağılımına sahip olacaklardır. Bir deneyde, rastlantısallık tüm gözlenen veya gözlenmeyen ortak değişkenlerin kontrolünü sağlamaktadır, yani tedavi ve kontrol grupları ortak değişken değerlerinin aynı dağılımına sahip olma eğilimindedir. Gözleme dayalı çalışmalarda ise, propensity skor $\lambda(x_s)$ de homojen olan tabakalar veya eşleştirilmiş kümeler gözlenen ortak değişken x_s leri dengeleme eğilimindedir. Propensity skor bir dengeleme ölçüsüdür çünkü propensity skorlar veri olduğunda tedavi öncesi özelliklerin koşullu dağılımları tedavi ve kontrol grupları için aynıdır.[*] Açıkça ifade etmek gerekirse, propensity skor bir kişinin sadece ortak değişken skoru kullanılarak tedavi edilmesinin bir olasılık ölçüsüdür.

---

[*] Rosenbaum, P. R. ve Rubin, D. B., (1983), ''The Central Role of The Propensity Score In Observational Studies for Causal Effects'', Biometrika, 70: 41–55.

3.2 Propensity Skor Yaklaşımı

Propensity skorda ortak değişkenlerin sayısı bakımından herhangi bir sınırlama yoktur. Ortak değişken sayısı arttığında, aynı veya benzer eşleştirilmiş çiftler bulmak zorlaşacaktır. Propensity skor, çok sayıda ortak değişken içerdiğinde, eşleştirilmiş setlerin ve tabakaların oluşturulması için bir düzenektir ve ortak değişken bilgilerinin özet ölçüsünü içermektedir.

Genellikle x de tabakalara kesin ayırma veya eşleştirme zordur veya mümkün değildir. x çok boyutluysa veya sürekli ölçüler içeriyorsa N birimin her biri farklı x değerine sahip olabilmekte böylece hiçbir tabaka aynı x ile tedavi veya kontrol birimlerini içermemektedir.

Gözleme dayalı çalışmada saklı sistematik hatanın olmadığını ve (2.3) denkleminin geçerli olduğu varsayılırsa, X'de tabakalara kesin ayırma veya eşleştirme yerine, aynı tabakada tedaviyi alma şansı aynı olan $\lambda(x)$ birimler ile tabakalar veya eşleştirilmiş kümeler oluşturmayı düşünecek olursak; bu durumda bir tabaka veya bir eşleştirilmiş küme arasında birimler x'lerin farklı değerlerine sahip olabilir ancak aynı propensity skor'a $\lambda(x)$'a, sahip olacaklardır yani $x_{si} \neq x_{sj}$ olabilir ancak her zaman $\lambda(x_{si}) = \lambda(x_{sj})$ olacaktır. Buna propensity skorda tam eşleştirme veya tabakalara ayırma denilmektedir. Bu durumda (2.3) ve (2.4) değişikliğe uğramadan geçerli olacaktır. Bu denklemlerde, tabaklardaki eşit x ler sadece $\lambda(x)$leri sağlamak için kullanılmaktadır. Kısaca, saklı sistematik hata olmayan gözleme dayalı çalışmada, propensity skor da tam eşleştirme veya tabakalara ayırma, verilen m'in tedaviye atamanın koşullu dağılımı Z yi sağlamaktadır. Bu, tekbiçim rassal deneydekiyle prop(Z=z|m)=1/K ile aynıdır. Rassal deney ile ulaşılan sonuçlar ile eğer tabakalar veya eşleştirilmiş kümeler $\lambda(x)$'e ve $\lambda(x)$ da homojen olan tabakalar ve eşleştirilmiş kümeleri içeren x'in bölümlerine dayanarak şekilleniyorsa aynı sonuçlara varılmaktadır.

Uygulamada, $\lambda(x)$ bilinmediğinden, $\lambda(x)$ de eşleştirme veya tabakalara ayırma gerçekleşemez. Bu nedenle tahmini propensity skorun kullanılması gerekmektedir.

3.3 Propensity Skorun Dengeleme Özellikleri

$\lambda(\mathrm{x}_s) = \Lambda$ olduğu en az bir s'in seçildiği varsayılsın. Propensity skorun $\lambda(\mathrm{x}_s) = \Lambda$ değerine sahip birkaç s olduğundan, böylece bu tabakalardaki birim toplam sayısı için;

$$n_\Lambda = \sum_{s:\lambda(x_s)=\Lambda} n_s = \sum {}^{*} n_s. \qquad (3.2)$$

yazılabilir.

Burada $\sum^*$ ile $\lambda(\mathrm{x}_s)=\Lambda$ gibi tüm s'lerin toplamını ifade etmektedir. Bu tabakalardan $\{(s,i):\lambda(x_s)=\Lambda\}$ rastgele bir birim (s,i) seçilir. Burada birimin seçilme olasılığı $1/n_\Lambda$ her birim için aynıdır. Bu rastgele birim için tedaviye atama ve gözlenen ortak değişken için sırasıyla Z ve X yazılır; eğer seçilen birey (s,i) ise $Z=Z_{si}$ ve $X=x_s$ şeklinde yazılır.

$Z=1$ in olasılığı vurgulamak için $\lambda(\mathrm{x}_s)=\Lambda$ ile n_Λ bireyden rasgele seçilen bir bireyden söz edilir ve olasılığı $prob\{Z=1|\lambda(X)=\Lambda\}$ olarak yazılır. $\lambda(\mathrm{x}_s)=\Lambda$ olduğunda, rastgele birey n_s/n_Λ olasılık ile s tabakasından çekilecektir. Böylece,

$$prob\{Z=1|\lambda(X)=\Lambda\}=\sum{}^*\frac{n_s\lambda(x_s)}{n_\Lambda}=\sum{}^*\frac{n_s\Lambda}{n_\Lambda}=\Lambda \tag{3.3}$$

olmaktadır.

Burada yine $\sum^*$ ile $\lambda(\mathrm{x}_s)=\Lambda$ gibi tüm s lerin toplamını ifade edilmektedir.

Bu $prob\{Z=1|\lambda(X)=\Lambda\}$ sonucu farklı bir yolla da ifade edilebilir. $\lambda(\mathrm{x}_s)=\Lambda$ olduğu durumda, (s,i) bireyi $1/n_\Lambda$ olasılık ile seçilir ve tedaviyi π_{si} olasılık ile alır; böylelikle

$$prob\{Z=1|\lambda(X)=\Lambda\}=\frac{1}{n_\Lambda}\sum{}^*\sum_{i=1}^{n_s}\pi_{si}=\sum{}^*\frac{n_s\lambda(x_s)}{n_\Lambda}=\Lambda \tag{3.4}$$

olmaktadır.

Aşağıda gösterilecek olan 3.1 önermesi ise, propensity skorun dengeleme özelliklerini açıklamaktadır.

Önerme 3.1. $\lambda(\mathrm{x}_s)=\Lambda$ ise,

$$prob\{X=x_s|\lambda(X)=\Lambda,Z=1\}=prob\{X=x_s|\lambda(X)=\Lambda,Z=0\}. \tag{3.5}$$

İSPAT. $\lambda(\mathrm{x}_s)=\Lambda$ ise Bayes Teoreminden,

$$prob\{X=x_s|\lambda(X)=\Lambda,Z=1\}=\frac{prob\{Z=1|\lambda(X)=\Lambda,X=x_s\}\,prob\{X=x_s|\lambda(X)=\Lambda\}}{prop\{Z=1|\lambda(X)=\Lambda\}}.$$

$prob\{Z=1 \mid \lambda(X)=\Lambda, X=x_s\} = prob\{Z=1 \mid X=x_s\} = \lambda(x_s) = \Lambda$ yazılır ve

$prob\{Z=1 \mid \lambda(X)=\Lambda\} = \Lambda$, olduğundan

$prob\{X=x_s \mid \lambda(X)=\Lambda, Z=1\} = prob\{X=x_s \mid \lambda(X)=\Lambda\}$, □

ile teorem ispatlanır.

Aynı propensity skor değerine sahip tedavi ve kontrol birimlerin gözlenen ortak değişken X ile aynı dağılıma sahip olduğunu ifade edilmiştir. Bunun anlamı, propensity skorda homojen olan bir tabaka veya eşleştirilmiş kümede, tedavi ve kontrol birimlerinin farklı X değerlerine sahip olabileceği fakat bu farklar sistematik farklardan çok şans farkları olacağıdır. Dengeleme burada gözlenen ortak değişkenler ile ilgili olup ancak rassal deneyden farklı olarak gözlenemeyen ortak değişkenleri dengelememektedir. Başka bir deyişle, önerme 3.1'e göre; propensity skorun Λ bir değeri seçilip, bu propensity skor değeri ile n_Λ birimden rastgele bir birim seçilirse, sonra da bu birim için, propensity skorun $\lambda(x_s)=\Lambda$ değeri verildiğinde tedaviye atama Z, ortak değişken X değerinden bağımsızdır.

3.4 Propensity Skorun Tahmini

Ortak değişkenlerde kayıp veri durumu söz konusu değilse, propensity skor diskiriminant analizi veya lojistik regresyon yardımıyla tahmin edilir. Ortak değişkenlerin çok değişkenli normal dağılıma sahip olduğu varsayıldığında ise diskriminant analizi kullanılır, lojistik regresyon için ise bu varsayıma gerek yoktur.[*] Propensity skor genellikle lojistik regresyon kullanılarak aşağıda ifade edilen yöntem ile tahmin edilmiştir.

$$\ln\left(\frac{PS}{1-PS}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p \tag{3.6}$$

$$PS = \frac{\exp\left(\beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p\right)}{1+\exp\left(\beta_0 + \beta_1 X_1 + \ldots + \beta_p X_p\right)} \tag{3.7}$$

Odds oranı = exp (constant + (coeff. x var1) + … + (coeff. x varN)) (3.8)

---

[*] D'Agostino RB, Jr., (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: 2265–2281.

PPSkor = Oddsoranı / (1+Oddsoranı) (3.9)

Propensity skoru hesaplamanın amacı, tedavi alacak birimin olasılığının yardımıyla (propensity skor) tedavi etkisinin tahmininin düzeltilerek hemen hemen rassal bir deneyin etkisini yaratmaktır. Aynı propensity skor ile biri tedavi grubunda biri de kontrol grubunda iki birim bulunursa, bu durumda iki birimin her gruba rassal olarak atandığını, bir anlamda eşit benzerlikte tedavi veya kontrolde olduğu düşünülebilir.

3.5 Propensity Skorun Amacı ve Yararları

Propensity skor gözleme dayalı çalışmalarda sistematik hatanın azaltılmasında, kesinliğin artırılmasında ve belirli ortak değişkenlerin etkilerini anlamak için kullanılmaktadır.

Propensity skor bir kez tahmin edildikten sonra eşleştirme, tabakalara ayırma (sınıflara ayırma), regresyon düzeltmesi (kovaryans düzeltmesi) veya bu üçünün bazı birleşiminin kullanılması yöntemleriyle sistematik hata azaltılabilir.[*] Bu yöntemler, tedavinin gözlenen ortak değişkenlere göre koşullu olasılığı olan $\lambda(x) = pr(Z = 1 \mid X)$, propensity skorun tanımı nedeniyle oldukça kullanışlıdır ve bu $\lambda(x)$ göre Z ve X koşullu olarak bağımsızdır anlamına gelmektedir.

Böylelikle, tedavi ve kontrol gruplarındaki eşit (veya eşite yakın) propensity skora sahip birimler, kendi geçmiş ortak değişkenlerine bağlı olarak aynı (veya hemen hemen aynı) dağılıma sahip olma eğiliminde olacaktır. Propensity skor ile kesin düzeltmelerin yapılması, ortak değişkenlerdeki tüm sistematik hatayı ortadan kaldıracaktır. Bu nedenle sistematik hatayı ortadan kaldırmak için yapılan düzeltmeler, tek tek tüm ortak değişkenleri kullanmak yerine propensity skor kullanılarak yapılabilir.[**]

---

[*] Rosenbaum, P. R. ve Rubin, D. B., (1983), ‘’The Central Role of The Propensity Score In Observational Studies for Causal Effects’’, Biometrika, 70: 41–55.

[**] D’Agostino RB, Jr., (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: 2265–2281.

3.6 Propensity Skor ile Sistematik Hatayı Azaltma Yöntemleri

3.6.1 Eşleştirme

Bu tip araştırmalarda, genellikle sınırlı sayıda tedavi hastası buna karşılık fazla sayıda kontrol birimi ile karşılaşılmaktadır. Mümkün tüm birimlerden çıktı ölçülerini sağlamak finansal açıdan oldukça belirsizdir. Bu durumda bazı örnekleme biçimlerinin uygulanması gerekmektedir. Eşleştirme ise, araştırmacılar tarafından sıklıkla araştırmada kontrol edilmesine gerek duyulan ortak değişkenlerde, tedavi bireyleri ile eşleşmiş kontrol birimleri seçmek için sıkça kullanılmaktadır. Propensity skor eşleştirmesi ile araştırmacı tek bir skaler değişken yardımıyla eşleştirme yaparak çok sayıda ortak değişken üzerinde kontrol sağlanmaktadır.

Uygulamada, propensity skorun bir eşleştirme değişkeni olarak kullanılmasından önce bazı sorunlar üzerinde durulmalıdır. İlk olarak, $\lambda(x)$ 'in fonksiyonel biçimi ender olarak bilinmekte ve $\lambda(x)$ 'in kullanılan veriden tahmini gerekmektedir. Ayrıca, tam eşleştirme genellikle mümkün olmadığından $\lambda(x)$ deki yakınlık sorunları üzerinde durulmalıdır. Üçüncü olarak ta, $\lambda(x)$ için yapılan düzeltmelerin x'i sadece dengelemesi beklenmektedir.

Eşit-yüzdeli sistematik hata azaltma eşleştirme yöntemleri (EPBR): Ortalama sistematik hata veya x deki eşleştirmeden önceki beklenen fark $E(x \mid z=1) - E(x \mid z=0)$ dır, x de eşleştirme yaptıktan sonra ortalama sistematik hata ise $E(x \mid z=1) - \mu_{0M}$ dır. Burada μ_{0M} eşleşmiş kontrol grubundaki x in beklenen değeridir. Genel olarak μ_{0M} , kullanılan eşleştirme yöntemlerine bağlı iken $E(x \mid z=1)$ ve $E(x \mid z=0)$ sadece kütle özelliklerine bağlıdır. Sistematik hatadaki azalma x in her koordinatı için aynıysa eşleştirme yöntemi bir eşit-yüzdeli sistematik hata azaltma yöntemidir*, bu bazı skaler $0 \le \gamma \le 1$ için şöyledir;

$$E(x \mid z=1) - \mu_{0M} = \gamma \left\{ E(x \mid z=1) - E\left(x \mid z=0\right) \right\} \tag{3.10}$$

Bir eşleştirme yöntemi eşit yüzdeli sistematik hata azaltma yöntemi değilse, eşleştirme x in bazı doğrusal fonksiyonları için sistematik hatayı arttıracaktır. X ve yanıt değişkenleri arasındaki eşleştirme yapıldıktan sonra toplanacak ilişki hakkında az şey bilindiğinde, eşit

---

* Rubin, D. B., (1976), ''Matching Methods That Are Equal Percent Bias Reducing: Some Examples'', Biometrics, 32: 109–120.

yüzdeli sistematik hata azaltma yöntemi oldukça cazip bir yöntemdir.

Propensity skor kullanarak eşleşmiş örnek oluşturmak için Rosenbaum ve Rubin üç teknik önermişlerdir[*].

Tahmin Edilen Propensity Skora Bağlı Olarak Mümkün Olan En Yakın Eşleştirme: Yöntemde tedavi ve kontrol birimleri rassal şekilde dizilir. İlk tedavi birimi ile propensity skoru mümkün olan en yakın kontrol birimi ile eşleştirilir sonra da her iki birim de eşleştirme yapılacak örneklemden çıkartılır. Bir sonraki tedavi birimi seçilir, bu süreç tüm tedavi birimleri için eşler bulunana kadar tekrar edilir. Eşleştirme yapmak için tahmini propensity skorun logitlerinin kullanılması tavsiye edilmektedir. $\hat{q}(X) = \log[1 - \hat{\lambda}(X) / \hat{\lambda}(X)]$ kullanılır çünkü $\hat{q}$ (X) in dağılımı genellikle yaklaşık olarak normaldir.

İki-örneklem t testi tedavi ve kontrol gruplarındaki ortak değişkenlerin dağılımlarını karşılaştırabilmek için kullanılan uygun bir testtir. Eşlemik (paired) t testi ise eşleştirilmiş çiftlerin farklarına bağlı olarak çıktı değişkenlerinin analizinde, ortak değişkenlerdeki sistematik hatanın etkisini değerlendirmek için uygun bir testtir.

Propensity Skoru İçeren Mahalanobis Metrik Eşleştirmesi: Ortak değişkenler çok değişkenli normal dağılıyorsa ve tedavi ve kontrol gruplarının ortak kovaryans matrisleri varsa bu durumda Mahalanobis metrik eşleştirmesi eşit yüzdeli sistematik hata azaltıcı bir tekniktir. Burada sistematik hata tedaviyi alanların ortalaması eksi kontrollerin ortalamasıdır. Diğer bir deyişle, tüm ortak değişkenlerdeki sistematik hatanın azalma yüzdesi eşittir ve eşleştirme ile sistematik hatası yükselecek hiçbir ortak değişken yoktur.[**] Maholanobis metrik eşleştirmesinde, tedavi ve kontroller rassal şekilde dizilirler. Tedavi birimi i ile kontrol birimi j arasındaki Mahalanobis farkı;

$$d(i, j) = (u - v)^T C_{0R}^{-1} (u - v), \tag{3.11}$$

İlk tedavi birimi Maholanobis farkın en yakın olduğu kontrol birimi ile eşleşir. Burada u ve

---

[*] Rosenbaum, P. R. ve Rubin, D. B., (1985), ‘’Constructing A Control Group Using Multivariate Matched Sampling Methods That Incorporate The Propensity Score’’, American Statistician, 39: 33–38.

[**] Rubin, D. B., (1976), ‘’Matching Methods That Are Equal Percent Bias Reducing: Some Examples’’, Biometrics, 32: 109–120.

v, $\left\{x^T, \hat{q}(x)\right\}^T$ nin tedavi birimi i ve kontrol birimi j için eşleşen değişkenlerin değerleridir ve C_{0R} kontrol grubundaki $\left\{x^T, \hat{q}(x)\right\}$ in örnek kovaryans matrisidir. Eşleşen bu iki birim örneklemden çıkarılır ve bu süreç tüm tedavi birimleri için eşler bulunana kadar tekrar edilmektedir.

Propensity Skorlar Tarafından Tanımlanan Aralıklar Arasında Mümkün Olan En Yakın Mahalanobis Metrik Eşleştirmesi: Bu yöntem yukarıda anlatılan iki yöntemi birleştirmektedir. Tüm tedavi birimleri rassal şekilde sıralanır ve ilk tedavi birimi seçilir. Tedavi birimlerinin tahmini propensity skorlarının değer aralığındaki tüm kontrol birimleri seçilir ve az sayıda ortak değişkene bağlı olarak bu birimler ile tedavi birimi arasındaki Mahalanobis mesafesi hesap edilir. En yakın kontrol birimi ve tedavi birimi örneklemden çıkarılır, geriye kalan tüm kontrol birimleri bir tedavi birimi ile bir sonraki eşleşme için hazır bulunurlar ve bu süreç bu şekilde tekrarlanmaktadır.

Bu üç yöntem de standart hatanın azaltılması için kullanışlı tekniklerdir. Tahmin edilen propensity skora bağlı olarak mümkün olan en yakın eşleştirme hesaplaması en kolay eşleştirme tekniğidir.[*] İkinci yöntem, propensity skoru içeren Mahalanobis metrik eşleştirmesi yalnız değişkenler için daha küçük standart biçime dönüştürülmüş farklar üretirken, propensity skorda önemli farklar kalmaktadır. Üçüncü yöntem, propensity skorlar tarafından tanımlanan aralıklar arasında mümkün olan en yakın Mahalonobis metrik eşleştirmesi bu üç yöntem arasındaki en iyi tekniktir. Tedavi ve kontrol gruplarındaki ortak değişkenler arasındaki en iyi dengeyi üretmektedir. Bu üçüncü yöntem propensity skora bağlı değer aralığı oluşturularak, araştırmacı rassal bir deney oluşturmaya çalışmaktadır şeklinde düşünülebilir.

Standart biçime dönüştürülmüş farklarda, burada eşleştirme yapıldıktan sonra ortak değişken dengesini ölçmek için kullanılmaktadır. Standart biçime dönüştürülmüş farklar hem sürekli hem de nitel değişkenler için ortak değişken dengesinin yorumu için uygun bir yöntemdir.

Standart biçime dönüştürülmüş farklar;

---

[*] Rosenbaum, P. R. ve Rubin, D. B., (1985), ‘’Constructing A Control Group Using Multivariate Matched Sampling Methods That Incorporate The Propensity Score’’, American Statistician, 39: 33–38.

$$d = \frac{100\left(\overline{x}_{tedavi} - \overline{x}_{kontrol}\right)}{\sqrt{\frac{s^2_{tedavi} + s^2_{kontrol}}{2}}} \quad \text{sürekli değişkenler için,} \tag{3.12}$$

$$d = \frac{100\left(p_{tedavi} - p_{kontrol}\right)}{\sqrt{\frac{p_t\left(1-p_t\right) + p_k\left(1-p_k\right)}{2}}} \quad \text{ikili değişkenler için,} \tag{3.13}$$

şeklinde hesaplanır. $\left|\text{standart biçime dönüştürülmüş farklar}\right| > \%10$ olduğu durumda ortak değişkende ciddi bir dengesizliğe işaret etmektedir.

Propensity skor kullanılarak yapılan eşleştirme sonucu ortak değişkende ciddi bir dengesizlik ortaya çıkarsa, bu durumda;

- Eşleştirmeden sonra dengesizliğe sahip olan ortak değişken için regresyon düzeltmesi yapılması düşünülebilir.
- Ortak değişken tarafından belirlenen kavrama uygun ilave veya alternatif bir ölçümün propensity skor modeline eklenmesi ek bir çözümdür.
- Tedavi birimlerinin farklı bir rassal sıraya sokarak tekrar eşleştirme yapılabilir veya farklı bir standart kullanılarak tekrar eşleştirmenin yapılması düşünülebilir. Propensity skorlar tarafından tanımlanan aralıklar arasında da Mahalonobis metrik eşleştirmesi yöntemi kullanılabilir.

3.6.2 Tabakalara Ayırma

Tabakalara ayırma yöntemi de gözleme dayalı çalışmalarda tedavi ve kontrol grupları arasındaki sistematik farkları kontrol etmek için kullanılmaktadır. Bu teknik gözlenen geçmiş özellikleri açısından birimleri tabakalara gruplar. Açıkça, bu işlem sadece gözlenen ortak değişkenlerdeki dengesizliklere bağlı olarak ortaya çıkan sistematik hatayı kontrol edebilir. Tabakalar bir kez belirlendikten sonra aynı tabakadaki tedavi ve kontrol birimleri doğrudan ayrışırlar. Eşleştirmede olduğu gibi tabakalara ayırmada da karşılaşılan problem, ortak değişken sayısı arttığından dolayı ortaya çıkmaktadır çünkü tabakaların sayısı üstel şekilde artmaktadır. Örneğin; tüm değişkenler nitel değişken ise, bu durumda k ortak değişken için 2^k tabaka olacaktır. Eğer k büyük ise bazı tabakalar sadece tedavi grubundan birimleri içerecek böylelikle tabakadaki tedavi etkisini tahmin etmek olanaksız hale gelecektir. Bu yöntemde yine propensity skor çok kullanışlıdır çünkü propensity skor gözlenen ortak

değişkenlerin skaler bir özetidir. Buna bağlı olarak tabakalara ayırmak, tedavi ve kontrol gruplarında tabaka sayısı üstel olarak artmaksızın ortak değişkenlerin dağılımlarının dengelenmesini sağlamaktadır. Propensity skora bağlı olarak tabakalara kesin ayırma yapıldığında, tabakalar arasında ortalama tedavi etkisi, gerçek tedavi etkisinin yansız (sistematik hatasız) tahmincisini üretir. Propensity skora bağlı olarak tabakalara ayırma, propensity skoru tahmin etmek için kullanılan k ortak değişkeni dengeler ve genellikle propensity skora bağlı beş tabaka, bu ortak değişkenlerin her birindeki sistematik hatanın %90 nını ortadan kaldırmaktadır[*].

Önce propensity skor lojistik regresyon veya diskriminant analizi ile tahmin edilir. Araştırmacı, her iki grubun birlikte veya tedavi grubunda ve kontrol grubunda ayrı olarak propensity skor değerlerine bağlı olarak tabaka sınırlarına karar vermelidir. Çalışmalarda, genellikle tedavi ve kontrol gruplarını birlikte farklı tabakalara ayırmak için ayırma değerlerine tahmini propensity skoru eşit beş tabakaya ayırarak karar verilir. İlk dengesizlik gruplar ayrılarak F testi ölçülür. Sonra da propensity skorda tabakalara ayırma yapıldıktan sonra dengesizliği hesaplamak için 2 yönlü varyans analizi kullanılır.

Tabakalara ayırmada propensity skor kullanımı birçok araştırmada kullanılmıştır. Bunlardan biri Stone ve diğerleri tarafından yapılmıştır. Bu çalışmada, araştırmacılar hastanede yatan 265 veya hastalığı ayakta geçiren 482 birim olmak üzere toplam 747 zatürree hastasının çıktılarını karşılaştırmak istemişlerdir. Ağaç sınıflaması tekniği ile propensity skorlar tahmin edilmiş, daha sonra hastalar propensity skora bağlı olarak yedi tabakaya atanmışlardır. 44 değişkenin 29 unda iki grup arası dengesizlik olduğu saptanmış ve propensity skorda tabakalara ayırma yapıldıktan sonra bunlardan sadece 13 ü p=0.05 de anlamlı kalmıştır. Araştırmacılar böylece doğrudan standardizasyon yöntemini kullanılarak tabakanın-spesifik ortalaması ile tedavi etkisini tahmin etmişlerdir[**].

Tahmini propensity skor ile eşleştirme ve tabakaları ayırma yöntemine başka bir örnek daha verilebilir:

---

[*] Rosenbaum, P. R. ve Rubin, D. B., (1984), ‘’Reducing Bias In Observational Studies Using Subclassification On The Propensity Score’’, Journal of the American Statistical Association, 79: 516–524.

[**] Stone, R. A., Obrosky, S., Singer, D. E., Kapoor, W. N. ve Fine, M. J., (1995), ‘’Propensity Score Adjustment For Pretreatment Differences Between Hospitalized And Ambulatory Patients With Community-Acquired Pneumonia’’, Medical Care, 33:AS56-AS66.

Sorun, kalp atardamarı hastalığının tedavisi için bir kalp bypass ameliyatı ile ilaç veya tıbbi tedavinin karşılaştırılmasıdır (Rosenbaum ve Rubin, 1984). n=1515 birey, 590 nı ameliyat hastası, 925 tıbbi hastadan oluşmaktadır. x vektörü 74 gözlenen ortak değişkeni içermekte ve bu değişkenler hasta geçmişi ile ilgili bilgilerden oluşmaktadır. 74 ortak değişkenin her biri tedavi ve kontrol gruplarında istatistiksel olarak anlamlı bir dengesizlik göstermektedir.

Bu 74 ortak değişken tahmini propensity skordan şekillenen 5 tabaka kullanılarak kontrol edilmiştir. Bu örnekte propensity skor doğrusal bir logit model kullanılarak tahmin edilmiştir;

$$\log[\lambda(x)/1-\lambda(x)] = \alpha + \beta^T f(x) \tag{3.14}$$

Bu gözlenen x ortak değişkenlerden tahmini tedaviye atama Z dir. 1515 hasta, tahmini propensity skora dayanarak, her biri 303 hastayı içeren 5 tabakaya ayrılmış, en yüksek ameliyat olasılıklarına sahip tabaka 69 tıbbi hasta ile 234 ameliyat hastası içermektedir. Diğer taraftan en düşük ameliyat olasılıklarına sahip tabaka ise 277 tıbbı hastası ile 26 ameliyat hastasını içermektedir. Önerme (3.1) de tahmin yerine gerçek propensity skor kullanıldığında aynı tabakadaki tıbbi ve ameliyat hastalarının 74 ortak değişkeninin benzer dağılımına sahip olacaklarını önerilmiştir.

Çizelge 3.1, propensity skor öncesi ve sonrası tabakalara ayrılan 74 değişkenden beşinin değişimini göstermektedir. ''Önce'' diye belirtilen sütun, F oranını vermektedir, bu tıbbi ve ameliyat hastalarından her bir değişkeninin ortalamalarını karşılaştırır. Bu kolondaki değerlerin büyük olması tıbbi ve ameliyat hastalarının her bir değişkende istatistik anlamda farklılıklar olduğunu göstermektedir. Son iki sütun her bir değişken için iki yönlü 2×5 varyans analizinden F-oranlarını vermekte, buradaki iki faktörü ameliyatın ilaca (tıbbi) karşı olması ve beş ise propensity skor tabakalarıdır.

Çizelge 3.1 Tahmini bir propensity skordan önce ve sonra ayrılmış tabakalardaki ortak değişken dengesizlikleri: seçilmiş ortak değişkenler için F oranları (Kaynak: Rosenbaum, P. R., (1995), Observational Studies, Springer-Verlag , New York.)

		Sonra:	Sonra:
Değişken	Önce	Ana Etki	Etkileşim
Anormal sol kalp karıncığı küçülmesi	51,8	0,4	0,9
Göğüs ağrısında ilerleme	43,6	0,1	1,4
Sol ana kanalın daralması	22,1	0,3	0,2
Kardiomegali	25,0	0,2	0,0

Şimdiye kadar uzun süreli nitrat kullanımı ile tedavi olanlar	31,4	0,1	2,2

‘’Sonra: Ana Etki’’ yazılmış sütun, ilaca karşı ameliyatın ana etkisi için F-oranlarını göstermektedir. ‘’Sonra: Etkileşim’’ yazılmış sütun ise iki yönlü etkileşim için F-oranlarını vermektedir. Çizelge 3.1 deki hiçbir F oranı 0.05 düzeyinde anlamlı değildir ve F oranlarının birçoğu 1 den küçüktür. Böylece, bu ortak değişkenlerde sistematik dengesizlik işareti bulunmamaktadır. Tüm 74 ortak değişken için, $2\times 74=148$ F oranından sadece biri 0.05 düzeyinde anlamlı bulunmuştur. Rassal deneyde $0.05\times 148=7,4$ F oranının sadece şans tarafından anlamlı çıkması beklenir. Bu 74 ortak değişken 5 tabaka ile rassal bir deneyden beklenildiğinden daha fazla tabakalar arasında dengelenmişlerdir. Rasgeleleştirme hem gözlenen hem de gözlenmeyen ortak değişkenlerde denge sağlamaktadır, diğer yandan propensity skorda tabakalara ayırma ile gözlenmeyen ortak değişkenlerin dengelenmesi beklenemez.

Bu durumda tahmini propensity skorda tabakalara ayırma, 74 gözlenen ortak değişkeni dengelemiştir. Beş tabaka kullanılarak, tıbbi ve ameliyat hastaları çıktıları bakımından birbirinden ayrılmışlardır.

3.6.3 Regresyon Düzeltmesi (Kovaryans Düzeltmesi)

Regresyon düzeltmesinde, tedavi etkisi τ, t tedaviyi c kontrolü simgelemek üzere

$$\hat{\tau}=(\bar{Y}_t-\bar{Y}_c)-\beta\left(\bar{X}_t-\bar{X}_c\right) \tag{3.15}$$

olarak tahmin edilmektedir.

Geçmiş ortak değişkenlerin etkisi, bu denklemin sağ tarafındaki ikinci terim çıkartılarak düzeltilir. Burada β geçmiş ortak değişkenlere bağlı olarak tedavi ve kontrol grupları için bir regresyon tahminidir. Bu nedenden ötürü, tedavi ve kontrol gruplarında propensity skora bağlı yanıtların regresyonunu bulmak ve bunu kullanarak tedavi etkisinin en son tahminini düzeltmek için propensity skor değişkeninin kullanımı oldukça kullanışlıdır. Roseman çalışmasında tedavi ve kontrol gruplarında yanıtların yüzeyleri paralel ve ya doğrusal ya da doğrusal değilse, propensity skor kullanılarak regresyon düzeltmesi tedavi etkisinin

tahmininde sistematik hatayı azalttığını bulmuştur.* Buna ek olarak, tabakalara ayırma yapılabilir ve sonra da bu tabakalar arası regresyon düzeltmesinde kullanılırsa tahmini tedavi etkisinin bu tahmincisi, yalnız eşleştirmeye dayanan tahminciden daha etkilidir.

Regresyon düzeltmesi yönteminde başka bir yaklaşım ise, büyük bir ortak değişken kümesinin propensity skoru tahmin etmek için kullanılması, sonrada ortak değişkenlerin bir alt kümesini ve propensity skoru alarak bunları regresyon düzeltmesinde kullanılması şeklindedir.

* D'Agostino RB, Jr., (1998),'' Tutorial In Biostatistics: Propensity Score Methods for Bias Reduction In The Comparison of A Treatment To A Non-Randomized Control Group'', Stat Med, 17:2265-2281.

4. UYGULAMA

Uygulamada, Marmara Üniversitesi Hastanesi göğüs cerrahisi bölümünde 1996–2003 yılları arasında aynı doktor tarafından göğüs cerrahisi ameliyatı geçirmiş n=478 hasta kullanılmıştır. Ameliyat sonrası delirium tanısı alan 18 hasta ile ve 460 delirium tanısı almayan hastaya ait ameliyat öncesi 10 risk faktörü ve ameliyat sonrası 14 risk faktörüne istatistik analiz uygulanmıştır.

Çizelge 4.1 Ameliyat öncesi risk faktörlerinin tek değişkenli analizleri

	Delirium tanısı almayan hasta (n=460)	Delirium tanısı alan hasta (n=18)	p değeri †
Ameliyat Öncesi			
Yaş (YASKOD) (18–59=0) (59>=1)	% 33	% 66	0,005**
Cinsiyet (CINS) (erkek=1,bayan=0)	% 67	% 61	0,38
Kronik Hastalık [a] (K2) (var=1,yok=0)	% 3	% 0	0,55
Alkol Kullanımı (K3) (var=1,yok)	% 19	% 44	0,017**
Psikiyatri sorunu (K4) (var=1,yok=0)	% 12	% 22	0,171
Serebrovasküler hastalık (K5) (var=1,yok=0)	% 0,6	% 0	0,891
Diyabet (DM) (var=1,yok=0)	% 7	% 22	0,046**
Kemoterapi(KEMO) (var=1,yok=0)	% 3	% 0	0,536
Tümör türü (TTÜR) (var=1, yok=0)	% 53	% 55	0,514
Operasyon türü acil (OTÜR) (1–4)	2.76 ± 1,15	2,88 ± 1,13	0,667 [d]

† Ki-kare testi p değerleri, ** p<0,005 ; [a] Herhangi bir akciğer, kalp, böbrek ve karaciğer hastalığının mevcudiyeti [d] . t istatistiği p değeri (sürekli veri, ortalama ± standart sapma şeklinde gösterilmiştir.)

Çizelge 4.2 Ameliyat sonrası risk faktörlerinin tek değişkenli analizleri

	Delirium tanısı almayan hasta (n=460)	Delirium tanısı alan hasta (n=18)	p değeri †
Ameliyat Sonrası			
Solunum yetmezliği [b] (RD) (var=1,yok=)	% 18	%27	0,237
Serumdaki kimyasal değerlerin anormalliği [c] (ELEKTR) (var=1,yok=0)	% 5	% 27	0,004**
Operasyon süresi (AMLSÜRE) (var=1,yok=0)	% 38	% 72	0,004**
Yoğun bakım süresi (YBSTAY)	2,56 ± 7,89	3,05 ± 1,92	0,79 [d]

Uyku bozukluğu (UYKU) (var=1,yok=0)	% 0	% 11	0,001**
Hipertansiyon (K10) (var=1,yok=0)	% 12	% 22	0,272
Enfeksiyon (K11) (var=1,yok=0)	% 19	% 22	0,760
Kan transfüzyonu (K12) (var=1,yok=0)	% 29	% 22	0,606
Aminofilin kullanımı (AFİLİN) (var=1,yok=0)	% 3	% 5	0,464
Antiaritmi kullanımı (ARİTMİ) (var=1,yok=0)	% 7	% 11	0,371
Antibiyotik kullanımı (K16) (var=1,yok=0)	% 24	% 22	0,564
Streoid kullanımı (STREOİD) (var=1,yok=0)	% 5	% 11	0,285
Antihipertansif kullanımı (K17) (var=1,yok=0)	% 16	% 16	0,576
Hareket edememe (İMMOB) (var=1,yok=0)	% 10	% 16	0,288

† Ki-kare testi p değerleri; ** p<0,005 ; [b] Oda havasındaki Po_2<55 mm Hg kandaki oksijen miktarının yetersizliği ve Po_2 >44 mm Hg kandaki anormal karbondioksit seviyesi; [c] sodyum <130 veya >150 meq/L, potasyum <3.0 veya >6 meq/L, glukoz <60 veya >300 mg/dL; [d] .t istatistiği p değeri (sürekli veri, ortalama ± standart sapma şeklinde gösterilmiştir.)

Çizelge 4.1 ve Çizelge 4.2 her iki grup için ayrı ayrı 24 risk faktörünün tanımlayıcı istatistiklerini içermektedir. Örneğin delirium tanısı almayan 460 hastada solunum yetmezliği çekenlerin sayısı % 18 iken, delirium tanısı konan 18 hastanın solunum yetmezliği çekenlerin sayısı %27 dir. Tedavi ve kontrol grupları arasında ortak değişkenler bakımından fark olup olmadığını görmek için sürekli değişkenler için iki örnek t testi, süreksiz değişkenler için Ki-kare testi ve Fisher's kesin test uygulanmıştır. Burada sıfır hipotezimiz tedavi ve kontrol grupları arasında ortak değişkenler bakımından ortalamalarda fark olmadığı şeklindedir. Uygulama SPSS paket programında yapılmıştır. Nitel değişkenler için Ki-kare testi uygulanırken, sürekli veriye sahip değişkenler için t testi uygulanmıştır. Sıfır hipotezimiz burada $H_0 : \mu_1 = \mu_2$, alternatif hipotezimiz ise $H_1 : \mu_1 \neq \mu_2$ şeklinde kurulmuştur. Buna göre ameliyat öncesi yaş, alkol kullanımı, diyabet risk faktörleri ile ameliyat sonrası serumdaki kimyasal değerlerin anormalliği, operasyon süresi, uyku bozukluğu risk faktörlerinin p değerleri 0.05 altında çıkmış, yani iki grup arasında ortak değişkenler bakımından fark olduğu sonucuna varılmıştır.

Propensity skorları tahmin etmek için lojistik regresyon uygulanmış ve delirium tanısı için önemli risk faktörleri belirlenmiştir.

Çizelge 4.3 Verilerin özeti

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	478	100,0
	Missing Cases	0	,0
	Total	478	100,0
Unselected Cases		0	,0
Total		478	100,0

a. If weight is in effect, see classification table for the total number of cases.

Yukarıdaki Çizelge 4.3 de, çalışmadaki verilerin hiçbirinde kayıp veri olmadığını göstermektedir.

Çizelge 4.4 Bağımlı değişkenin kodlanması

Dependent Variable Encoding

Original Value	Internal Value
kontrol	0
hasta	1

Çizelge 4.4, bağımlı değişkende delirium tanısı alanların hasta ve 1 kodu verilerek kodlandığını, delirium tanısı almayanların ise kontrol ve 0 kodu alarak kodlandığını göstermektedir. Lojistik analiz hasta olanların olasılığına dayanarak hesaplanmıştır. Nitel değişkenlerin kodlaması ise Ek 1 de verilmektedir.

Çizelge 4.5 Denklemdeki değişkenler

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-3,241	,240	181,937	1	,000	,039

Bu SPSS çıktısı model için yapılan ilk testtir. Burada sıfır hipotezi $H_o : \beta_1 = ... = \beta_k = 0$ katsayılar tüm bağımsız değişkenler için sıfırdır şeklindedir. Alternatif hipotez ise $H_1 : \beta_1 \neq ... \neq \beta_k \neq 0$ şeklinde kurulmuştur. Test sonucunda sıfır hipotezinin reddedilmesi gerektiğine hükmedilmiştir. Bu da modelde en az bir beta katsayısının sıfırdan farklı olduğu anlamına gelmektedir.

Çizelge 4.6 Modelin açıklama gücü

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	78,443	,145	,529

Modelimizin Nagelkerke R^2 değeri 0,529 tür. Buna göre modelde bağımsız değişkenlerin beraberce bağımlı değişkenin % 53 açıklamaktadır anlamına gelmektedir.

Çizelge 4.7 Hosmer ve Lemeshow testi sonucu

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	1,081	8	,998

Hosmer ve Lemeshow modele uygunluk testi, tahmini olasılıklara dayanarak bireyleri desillere bölmekte, sonra gözlenen ve beklenen frekanslardan ki-kare istatistiğini hesaplamaktadır. Burada p= 0,998 dır. Sıfır hipotezi reddedilmemekte, dolayısıyla da tahminlerin modele iyi uyduğu söylenebilmektedir.

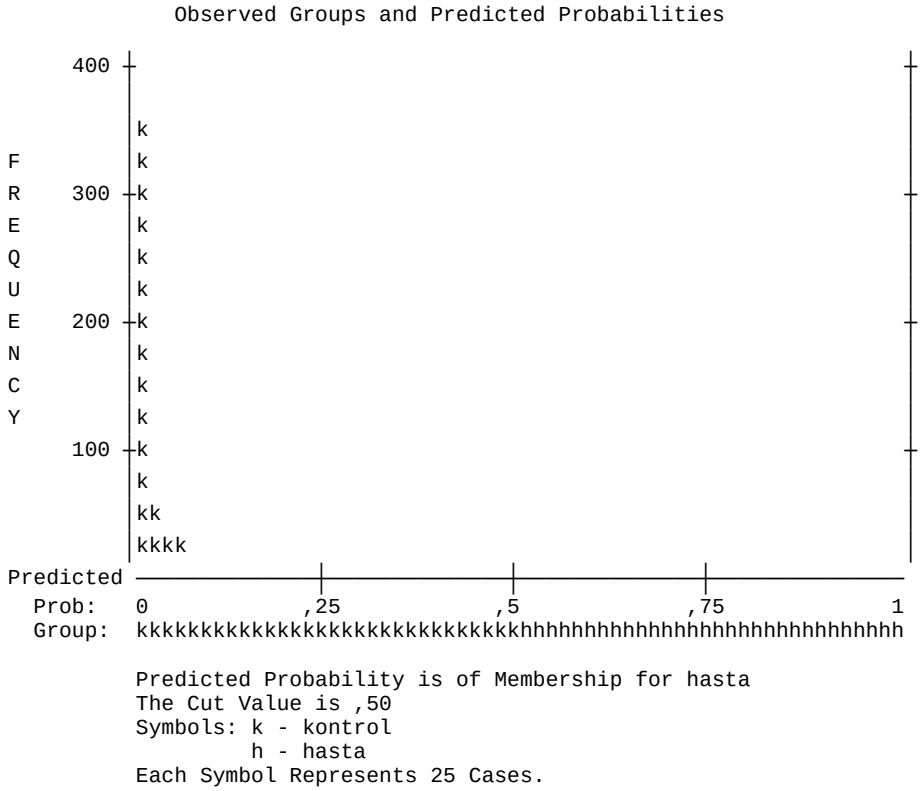
Çizelge 4.8 Lojistik regresyon çıktısı

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95,0% C.I.for EXP(B) Lower	Upper
Step 1[a]	TTÜR(1)	,515	,966	,284	1	,594	1,673	,252	11,112
	YASKOD(1)	-1,907	,884	4,654	1	,031	,149	,026	,840
	K2(1)	20,568	7836,056	,000	1	,998	8,6E+08	,000	.
	K3(1)	-,787	,920	,733	1	,392	,455	,075	2,760
	K4(1)	-2,095	,926	5,118	1	,024	,123	,020	,756
	K5(1)	-1,360	22029,444	,000	1	1,000	,257	,000	.
	AMLSÜRE(1)	-2,732	,968	7,960	1	,005	,065	,010	,434
	IMMOB(1)	,183	1,269	,021	1	,886	1,200	,100	14,422
	RD(1)	-,300	,775	,149	1	,699	,741	,162	3,386
	UYKU(1)	-43,443	28943,294	,000	1	,999	,000	,000	.
	K10(1)	-,032	1,058	,001	1	,976	,968	,122	7,708
	K11(1)	-,309	,964	,102	1	,749	,735	,111	4,862
	K12(1)	1,256	,810	2,407	1	,121	3,512	,718	17,176
	DM(1)	-1,198	,955	1,574	1	,210	,302	,046	1,961
	AFILIN(1)	-1,987	1,508	1,736	1	,188	,137	,007	2,634
	ARITMI(1)	18,993	5811,123	,000	1	,997	1,8E+08	,000	.
	K16(1)	,533	,977	,298	1	,585	1,704	,251	11,568
	K17(1)	2,078	1,437	2,090	1	,148	7,988	,477	133,623
	KEMO(1)	16,539	8692,541	,000	1	,998	1,5E+07	,000	.
	STEROID(1)	-1,688	1,366	1,527	1	,217	,185	,013	2,688
	ELEKTR(1)	-1,944	,939	4,290	1	,038	,143	,023	,901
	CINS(1)	1,149	,797	2,078	1	,149	3,156	,661	15,058
	OTÜR			1,432	4	,839			
	OTÜR(1)	-42,552	25344,458	,000	1	,999	,000	,000	.
	OTÜR(2)	-24,325	25075,102	,000	1	,999	,000	,000	.
	OTÜR(3)	-24,169	25075,102	,000	1	,999	,000	,000	.
	OTÜR(4)	-25,559	25075,102	,000	1	,999	,000	,000	.
	YBSTAY	,016	,055	,085	1	,771	1,016	,913	1,131
	Constant	17,271	45148,061	,000	1	1,000	3,2E+07		

a. Variable(s) entered on step 1: TTÜR, YASKOD, K2, K3, K4, K5, AMLSÜRE, IMMOB, RD, UYKU, K10, K11, K12, DM, AFILIN, ARITMI, K16, K17, KEMO, STEROID, ELEKTR, CINS, OTÜR, YBSTAY.

Çizelge 4.8 de lojistik regresyon sonucu yaş, psikiyatri sorunu, operasyon süresi, serumdaki kimyasal değerlerin anormalliği risk faktörlerinin p değerleri 0.05 altında çıkmış yani bu risk faktörlerinin parametre tahminleri anlamlı kabul edilmiştir. Exp(B) ise odds oranını göstermektedir.



Şekil 4.1 Sınıflama tablosu

Şekil 4.1 lojistik regresyon altında doğru veya yanlış tahmincileri tespit etmek için yararlanılan alternatif bir yöntemdir. X ekseni 0.0 dan 1.0 a kadar ‘’1’’ de sınıflanmış bağımlı değişkenin tahmini olasılıklarıdır. Y ekseni, sınıflanmış birimlerin sayısı; frekanslardır. Yukarıdaki şeklin içinde sıfırın ve birin gözlendiği sütunlardır. Burada ‘k’lar kontrol, ‘h’ ler ise hastadır ve her bir sembol 25 bireyi göstermektedir. Şekil modelin zor bireyleri iyi sınıfladığını göstermektedir. Şekil aslında hasta iken, kontrol olarak gruplanmış bireyleri görmemizi sağlamaktadır. Şekil aynı zamanda çalışmada aslında hasta iken, kontrol olarak gruplanmış hiçbir bireyin olmadığını göstermektedir.

Propensity skorların SPSS programı kullanılarak elde edilmesinin gösterimi Ek 2 de verilmektedir. Propensity skor genellikle lojistik regresyon kullanılarak aşağıda ifade edilen yöntem ile tahmin edilmiştir ve ilk 10 delirium tanısı konmayan ve delirium tanısı alan hasta

için propensity skorlar şöyledir:

Çizelge 4.9 İki gruptaki ilk 10 bireyin propensity skorları

	Delirium tanısı alan grup	**Delirium tanısı almayan grup**
1	1	0.01644
2	0,21501	0.01951
3	0,08691	0.02112
4	1	0.00635
5	0,13866	0.00768
6	0,07576	0.05915
7	0,02145	0.01570
8	0,22228	0.07505
9	0,00687	0.13616
10	0,12452	0.03201
…	…	…

$$\ln\left(\frac{PS}{1-PS}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p \qquad PS = \frac{\exp\left(\beta_0 + \beta_1 X_1 + ... + \beta_p X_p\right)}{1 + \exp\left(\beta_0 + \beta_1 X_1 + ... + \beta_p X_p\right)}$$

Odds ratio = exp(constant + (coeff. x var1) + … + (coeff. x varN))

PPSkor = Oddsratio / (1+Oddsratio)

Cochran (1968), ortak değişken veya değişkenin alt sınıflamasına bağlı olarak yapılan beş tabakaya ayırmanın sistematik hatayı %90 olarak azalttığını göstermiştir.[*] Bu nedenle tüm

[*] Cochran, W. G., (1968), ‘’The Effectiveness of Adjustment By Subclassification In Removing Bias In Observational Studies’’, Biometrics, 24, 205–*213*.

delirium tanısı alan ve almayan hastalar için propensity skorlar tahmin edildikten sonra her iki grubun birlikte bu değerlerine bağlı olarak eşit örneklem sayısına sahip 5 tabaka sınırına karar verilmiştir. Çizelge 4.10 bu çalışma için oluşturulan tabakaları ve sınırlarını göstermektedir.

Çizelge 4.10 5'li tabakalara ayrıldığında propensity skor düzeltmesi sonrası kontrol ve tedavi gruplarına düşen örneklem sayısı

Tabakalar	Sınırlar	Delirium tanısı alan	Delirium tanısı almayan
1	0.5–0.115	7	28
2	0.115–0.05	2	33
3	0.05–0.022	1	34
4	0.022–0.0121	1	34
5	0.0121–0.006	1	34

Tabakalara ayırdıktan sonra çalışmamızda delirium tanısı alan 12 hasta ile almayan 163 hasta kalmış ve böylece birbiri ile benzer dağılıma sahip tedavi ve kontrol bireyleri seçilmiştir. Tabakalara ayırmadan önceki ve sonraki F istatistik sonuçları Çizelge 4.11 ve 4.12 de verilmiştir. İlk dengesizlik için gruplar ayrılarak F istatistiği ölçülür. Daha sonra propensity skorda tabakalara ayırma yapıldıktan sonra dengesizliği hesaplamak için 2 yönlü varyans analizi kullanılır.

Çizelge 4.11 Tabakalara ayırmadan önce ve sonra iki grup arasındaki dengesizliği ölçmek amacıyla hesaplanan F-istatistiği sonuçları

	Tabakalara ayırmadan önce F istatistiği †	Tabakalara ayırdıktan sonra F istatistiği †
Ameliyat Öncesi		
Yaş (YASKOD) (18–59=0) (59>=1)	8,55 **	0,021
Cinsiyet (CINS) (erkek=1,bayan=0)	0,267	0,043
Kronik Hastalık (K2) (var=1,yok=0)	0,606	-
Alkol Kullanımı (K3) (var=1,yok)	6,578**	0,055
Psikiyatri sorunu (K4) (var=1,yok=0)	1,687	0,662
Serebrovasküler hastalık (K5) (var=1,yok=0)	0,118	-

Diyabet (DM) (var=1,yok=0)	5,207**	0,349
Kemoterapi(KEMO) (var=1,yok=0)	0,648	-
Tümör türü (TTÜR) (var=1, yok=0)	0,044	0,021
Operasyon türü acil (OTÜR) (1–5)	0,184 [a]	0,802 [a]

† F istatistiği: $\chi^2 = F_{(r,\infty)}.r$ (r= serbestlik derecesi); [a] F istatistiği= t^2 ; ** p<0,05

Çizelge 4.12 Tabakalara ayırmadan önce ve sonra iki grup arasındaki dengesizliği ölçmek amacıyla hesaplanan F-istatistiği sonuçları

	Tabakalara ayırmadan önce F istatistiği †	Tabakalar ayırdıktan sonra F istatistiği
Ameliyat Sonrası		
Solunum yetmezliği (RD) (var=1,yok=0)	0,980	1,284
Serumdaki kimyasal değerlerin anormalliği (ELEKTR) (var=1,yok=0)	13,982**	1,077
Operasyon süresi (AMLSÜRE) (var=1,yok=0)	8,357**	0,744
Yoğun bakım süresi (YBSTAY)	0,068 [a]	0,082
Uyku bozukluğu (UYKU) (var=1,yok=0)	51,32**	-
Hipertansiyon (K10) (var=1,yok=0)	1,418	0,197
Enfeksiyon (K11) (var=1,yok=0)	0,065	0,459
Kan transfüzyonu (K12) (var=1,yok=0)	0,451	0,058
Aminofilin kullanımı (AFİLİN) (var=1,yok=0)	0,282	0,936
antiaritmi kullanımı (ARİTMİ) (var=1,yok=0)	0,453	-
Antibiyotik kullanımı (K16) (var=1,yok=0)	0,027	0,705
Streoid kullanımı (STREOİD) (var=1,yok=0)	0,936	0,002
Antihipertansif kullanımı (K17) (var=1,yok=0)	0,004	0,010
Hareket edememe (İMMOB) (var=1,yok=0)	0,769	0,315

† F istatistiği: $\chi^2 = F_{(r,\infty)}.r$ (r= serbestlik derecesi); [a] F istatistiği= t^2 ; ** p<0,05

Yukarıda sunulan tablolardan görülebileceği gibi tabakalara ayırdıktan sonra birbirleri ile benzer ortalamalara sahip tedavi ve kontrol gruplarının elde edildiğine F testi sonucunda karar verilmiştir. Propensity skora bağlı olarak, tabakalara ayırmadan önce yaş, alkol kullanımı, diyabet, serumdaki kimyasal değerlerin anormalliği, operasyon süresi ve uyku bozukluğu değişkenlerinde delirium tanısı alanlar ile almayanlar arasında farklılık olduğu görülmüş fakat

tabakalara ayırma yapıldıktan sonra bu farkların düzeldiğine F istatistiği sonucunda karar verilmiştir. Kronik hastalık, serebrovasküler hastalık, uyku bozukluğu, antiaritmi kullanımı, kemoterapi değişkenlerinin var veya yok şeklindeki iki yanıtında, tabakalara ayırma yapıldıktan sonra yanıt değerlerinde sabit değer gözlenildiğinden bu değişkenlere istatistik analiz uygulanamamıştır. Elde ettiğimiz yeni örneklem ile lojistik regresyon yardımıyla risk faktörlerinin anlamlılıkları test edilmiş ve sonuçları aşağıda verilmiştir.

Çizelge 4.13 Verilerin özeti

Case Processing Summary

Unweighted Cases [f]		N	Percent
Selected Cases [a,b,c,d,e]	Included in Analysis	175	100,0
	Missing Cases	0	,0
	Total	175	100,0
Unselected Cases		0	,0
Total		175	100,0

a. The category variable K2 is constant for all selected cases. Since a constant was requested in the model, it will be removed from the analysis.

b. The category variable K5 is constant for all selected cases. Since a constant was requested in the model, it will be removed from the analysis.

c. The category variable UYKU is constant for all selected cases. Since a constant was requested in the model, it will be removed from the analysis.

d. The category variable ARITMI is constant for all selected cases. Since a constant was requested in the model, it will be removed from the analysis.

e. The category variable KEMO is constant for all selected cases. Since a constant was requested in the model, it will be removed from the analysis.

f. If weight is in effect, see classification table for the total number of cases.

Yukarıda değinildiği gibi kronik hastalık, serebrovasküler hastalık, uyku bozukluğu, antiaritmi kullanımı, kemoterapi değişkenlerinin var veya yok şeklindeki iki yanıtında, tabakalara ayırma yapıldıktan sonra yanıt değerlerinde sabit değer gözlendiğini ve hiç kayıp veri olmadığı görülmektedir.

Çizelge 4.14 Denklemdeki değişkenler

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	-2,609	,299	76,072	1	,000	,074

Test sonucunda 0,05 anlamlılık seviyesinde sıfır hipotezinin reddedilmesi gerektiğine karar verilmiştir. Bu da modelde en az bir beta katsayısının sıfırdan farklı olduğu anlamına gelmektedir.

Çizelge 4.15 Modelin açıklama gücü

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	76,229	,062	,158

Modelimizin Nagelkerke R^2 değeri 0,158 dur. Bu modelimizdeki bağımsız değişkenlerin beraberce, bağımlı değişkenin yaklaşık % 16 sını açıklamaktadır anlamına gelmektedir

Çizelge 4.16 Hosmer ve Lemeshow testi

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	9,969	8	,267

Burada p= 0,267 tür. Sıfır hipotezi reddedilmemekte, dolayısıyla da tahminlerin modele iyi uyduğu söylenebilmektedir.

Çizelge 4.17 Propensity skor düzeltmesi sonrası lojistik regresyon çıktısı

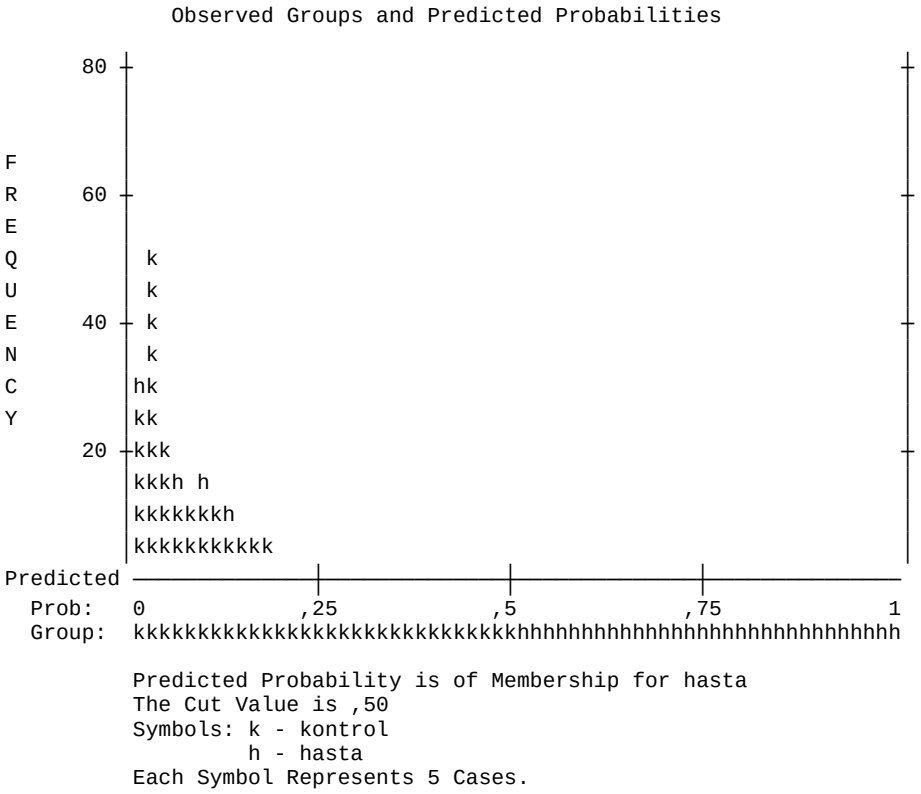
Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95,0% C.I.for EXP(B) Lower	Upper
Step 1[a]	TTÜR(1)	,649	1,006	,415	1	,519	1,913	,266	13,749
	YASKOD(1)	-1,363	,958	2,023	1	,155	,256	,039	1,674
	K3(1)	-,396	,997	,157	1	,692	,673	,095	4,752
	K4(1)	-1,743	,957	3,318	1	,069	,175	,027	1,142
	AMLSÜRE(1)	-2,109	1,064	3,927	1	,048	,121	,015	,977
	IMMOB(1)	,227	1,281	,031	1	,859	1,255	,102	15,450
	RD(1)	-,222	,791	,079	1	,779	,801	,170	3,771
	K10(1)	-,170	1,066	,025	1	,874	,844	,104	6,823
	K11(1)	-,171	,959	,032	1	,858	,843	,129	5,516
	K12(1)	,902	,826	1,192	1	,275	2,465	,488	12,450
	DM(1)	-1,050	,946	1,231	1	,267	,350	,055	2,236
	AFILIN(1)	-1,788	1,480	1,460	1	,227	,167	,009	3,043
	K16(1)	,227	,979	,054	1	,816	1,255	,184	8,543
	K17(1)	1,532	1,484	1,065	1	,302	4,625	,252	84,793
	STEROID(1)	-1,382	1,399	,975	1	,323	,251	,016	3,901
	ELEKTR(1)	-1,357	1,058	1,644	1	,200	,257	,032	2,049
	CINS(1)	,974	,795	1,499	1	,221	2,648	,557	12,581
	OTÜR			1,185	2	,553			
	OTÜR(1)	1,181	1,110	1,131	1	,288	3,256	,370	28,683
	OTÜR(2)	1,293	1,379	,879	1	,348	3,644	,244	54,397
	YBSTAY	,012	,054	,050	1	,823	1,012	,911	1,125
	Constant	1,930	3,465	,310	1	,577	6,892		

a. Variable(s) entered on step 1: TTÜR, YASKOD, K3, K4, AMLSÜRE, IMMOB, RD, K10, K11, K12, DM, AFILIN, K16, K17, STEROID, ELEKTR, CINS, OTÜR, YBSTAY.

Çizelge 4.17 deki çıktıya göre sadece operasyon süresi risk faktörünün parametre tahmini p değeri 0,05 altında çıkmış yani anlamlı kabul edilmiştir. Psikiyatri sorunu risk faktörünü ise p değeri 0,10 altında çıkmıştır yani 0,10 anlamlılık seviyesinde anlamlı kabul edilmiştir.

Sistematik hatanın ortadan kalktığı yeni örneklem sonucunda delirium tanısına etki eden en önemli risk faktörü operasyon süresi olarak bulunmuştur.



Şekil 4.2 Yeni örneklem için sınıflama tablosu

Şekil 4.2 de ‘k’lar kontrol, ‘h’ ler ise hastadır ve her bir sembol 5 bireyi göstermektedir. Bu şekli incelemek bize modelin zor bireyleri nasıl iyi sınıfladığını göstermektedir. Şekil aslında hasta iken, kontrol olarak gruplanmış bireyleri görmemizi sağlamaktadır. Şekil aynı zamanda çalışmada hasta olan 20 bireyin, kontrol olarak gruplanmış olduğunu göstermektedir.

5. SONUÇ VE TARTIŞMA

Özellikle, propensity skor uygulayarak rassal olmayan çalışmalarla ve geriye dönük çalışmalarla elde edilen sonuçlar, hemen hemen ileriye dönük rassal çalışmalar gibi çoğu kez doğru sonuçlar vermektedir. Propensity skor, istatistik analizlerde yoğun olarak özellikle uygulamalı tıp alanında kullanılmaktadır. Rassal klinik denemelerin maliyeti artıkça ve çoğu araştırmacının daha az pahalı olan araştırmalara yani gözleme dayalı çalışmalara yönelmesi sonucunda propensity skorun kullanımı artış göstermektedir. Çalışmanın tasarım aşamaları (eşleştirme veya tabakalara ayırma) göz önüne alındığında propensity skor yöntemi azami faydayı ürettiği görülmüştür. Bu fayda zaman ve maliyetten tasarruf ederek gerçek tedavi etkisinin daha kesin tahminlerini içermektedir. Bu tasarruf ile çalışma için uygun olmayacak bireylerin çalışmaya alınmasından kaçınılacaktır.[*] Çalışmanın en önemli tarafı ise propensity skor uygulanarak örneklemedeki sistematik hatanın azaltılmasıyla, (retrospective) rassal olmayan geriye dönük çalışmaların, ileriye dönük (prospective) rassal çalışmalar kadar yakın sonuçlar vermesidir. [**] [***]

Çalışmada propensity skor öncesi örneklem ile propensity skor sonrası örnekleme uygulanan analiz sonuçları risk faktörlerinin anlamlılığı bakımından oldukça farklıdır. Tabakalara ayırma yöntemi sonucunda birbiri ile ortak değişkenleri bakımından benzer dağılıma sahip tedavi ve kontrol bireyleri seçilmiş böylece yeni örneklemdeki tedavi ve kontrol grupları hemen hemen aynı karakteristiklere sahip olmuş ve böylelikle sistematik hata azalmıştır.

Delirium tanısı az rastlanır bir durum olduğundan çalışmamızda vaka sayısının oldukça düşük olması çalışma sonuçlarını olumsuz yönde etkilemiştir. Bu nedenle de propensity skoru tahmin ettikten sonra eşleştirme yöntemi yerine daha büyük örneklem ile çalışmak için tabakalara ayırma yöntemi kullanılmıştır. Vaka sayısı daha fazla olsaydı propensity skor

---

[*] D'Agostino RB, Jr., (1998),'' Tutorial In Biostatistics: Propensity Score Methods for Bias Reduction In The Comparison of A Treatment To A Non-Randomized Control Group'', Stat Med, 17:2265-2281.

[**] Boening A., Friedrich C., Hedderich J., Schoettler J., Fraund S. ve Cremer J. T., (2003), '' Early and Medium-Term After On-Pump and Off-Pump Coronary Artery Surgery: A Propensity Score Analysis'', Ann Thorac Surg, 76: 2000-2006

[***] Weitzen S. , Lapane K.L. , Toledano A. Y., Hume A. L. ve Mor V., (2004), ''Principles For Modeling Propensity Scores In Medical Research: A Systematic Literature Review'' Pharmacoepidemiology and Drug Safety, 13: 841–853

öncesi ve sonrası yapılan istatistik analizler farklı sonuçlar verebilirdi. Bu durum çalışmanın en önemli kısıtıdır.

Cochran (1968), ortak değişken veya değişkenin alt sınıflamasına bağlı olarak yapılan beş tabakaya ayırmanın sistematik hatayı %90 olarak azalttığını göstermiştir.[*] Bu nedenle çalışma için propensity skoru beş tabakaya ayırma yöntemi tercih edilmiştir. Tabakalara ayırma sonucunda tam eşleştirme sağlanamamış, denek sayısı 478 den 175 düşmüştür çünkü tedavi ve kontrol grupları propensity skorlarının birbirini tutmayan değişim aralıklarına sahiptir.[**] Çalışmamızda delirium tanısı alan hastaların minimum ve maksimum propensity skorları 0,00687 ve 1,00 , delirium tanısı almayan hastalar için ise minimum ve maksimum propensity skorlar 0,509 ve 0,00 dır. En yüksek propensity skora sahip delirium tanısı alanlar (>0,509) ve en düşük propensity skora sahip delirium tanısı almayanlar (<0,00687) çalışmadan çıkarılmıştır.

Propensity skorun tanımında sözü edilen Rosenbaum ve Rubin (1983) propensity skor hesaplanması ve sonrasında yapılan uygulamaları tamamlanmış veri kümesi için geliştirmiştir. Ancak günümüzde propensity skorun kayıp veriler için hesaplanması istatistik programlarda henüz yer almamakta, kayıp verilere sahip veri kümeleri için çalışmalar sürdürülmektedir. Bu konudaki en önemli çalışmalar olarak ta D'Agostino'nun yapıtlarına değinilmektedir.[***]

En nihayet yöntemin bu hali ile eşleştirmedeki kısıtlar yüzünden, büyük örneklerde iyi sonuç verdiği öne sürülebilir.

---

[*] Cochran, W. G., (1968), ''The Effectiveness of Adjustment By Subclassification In Removing Bias In Observational Studies'', Biometrics, 24, 205–213.

[**] Rosenbaum, P. R., (1995), Observational Studies, Springer-Verlag, New York.

[***] D'Agostino, R. B. Jr. ve Rubin, D. B., (2000), ''Estimating and Using Propensity Scores with Partially Missing Data'', Journal of the American Statistical Association, 95:749-759

KAYNAKLAR

Boening A., Friedrich C., Hedderich J., Schoettler J., Fraund S. ve Cremer J. T., (2003), ‘‘Early and Medium-Term After On-Pump and Off-Pump Coronary Artery Surgery: A Propensity Score Analysis’’, Ann Thorac Surg, 76: 2000-2006

Cameron, E. ve Pauling, L., (1976), ‘’Supplemental Ascorbate In The Supportive Treatment of Cancer: Prolongation of Survival Times In Terminal Human Cancer’’, Proceedings of the National Academy of Sciences (USA), 73: 3685–3689.

Cochran W. G., (1965),’’ The Planning of Observational Studies of Human Populations’’, Journal of the Royal Statistical Society, Series A, 128:134–155.

Cochran, W. G., (1968), ‘’The Effectiveness of Adjustment By Subclassification In Removing Bias In Observational Studies’’, Biometrics, 24, 205–213.

D’Agostino RB, Jr., (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: 2265–2281.

D’Agostino, R. B. Jr. ve Rubin, D. B., (2000), ‘‘Estimating and Using Propensity Scores with Partially Missing Data’’, Journal of the American Statistical Association, 95: 749-759

Doll, R. ve Hill, A., (1966), ‘’ Mortality of British Doctors In Relation to Smoking: Observations on Coronory Thrombosis’’, Epidemiological Approaches to the Study of Cancer and Other Cronic Diseases, 205–268.

Hays, W. L., (1991), Statistics, Fourth Edition, Harcourt Brace Collage Publishers, New York.

Rosenbaum, P. R. ve Rubin, D. B., (1983), ‘’The Central Role of The Propensity Score In Observational Studies for Causal Effects’’, Biometrika, 70: 41–55.

Rosenbaum, P. R. ve Rubin, D. B., (1984), ‘’Reducing Bias In Observational Studies Using Subclassification On The Propensity Score’’, Journal of the American Statistical Association, 79: 516–524.

Rosenbaum, P. R. ve Rubin, D. B., (1985), ‘’Constructing A Control Group Using Multivariate Matched Sampling Methods That Incorporate The Propensity Score’’, American Statistician, 39: 33–38.

Rosenbaum, P. R., (1995), Observational Studies, Springer-Verlag, New York.

Rubin, D. B., (1976), ‘’Matching Methods That Are Equal Percent Bias Reducing: Some Examples’’, Biometrics, 32: 109–120.

Rubin, D.B., (1977), ‘’ Assignment to The Treatment Group on The Basis of A Covariate’’, Journal of Educational Statistics, 2: 1–26

Stone, R. A., Obrosky, S., Singer, D. E., Kapoor, W. N. ve Fine, M. J., (1995), ‘‘Propensity Score Adjustment For Pretreatment Differences Between Hospitalized And Ambulatory Patients With Community-Acquired Pneumonia’’, Medical Care, 33:AS56-AS66.

Ury, H., (1975), ‘’Efficiency of Case-Control Studies With Multiple Controls Per Case: Continuous or Dichotomous Data’’, Biometrics, 31: 643–649

Weitzen S. , Lapane K.L. , Toledano A. Y., Hume A. L. ve Mor V., (2004), ‘’Principles For

Modeling Propensity Scores In Medical Research: A Systematic Literature Review'' Pharmacoepidemiology and Drug Safety, 13: 841–853

EKLER

Ek 1 Nitel değişkenlerin kodlama tablosu

Çizelge Ek 1.1 Nitel değişkenlerin Spss de kodlanması

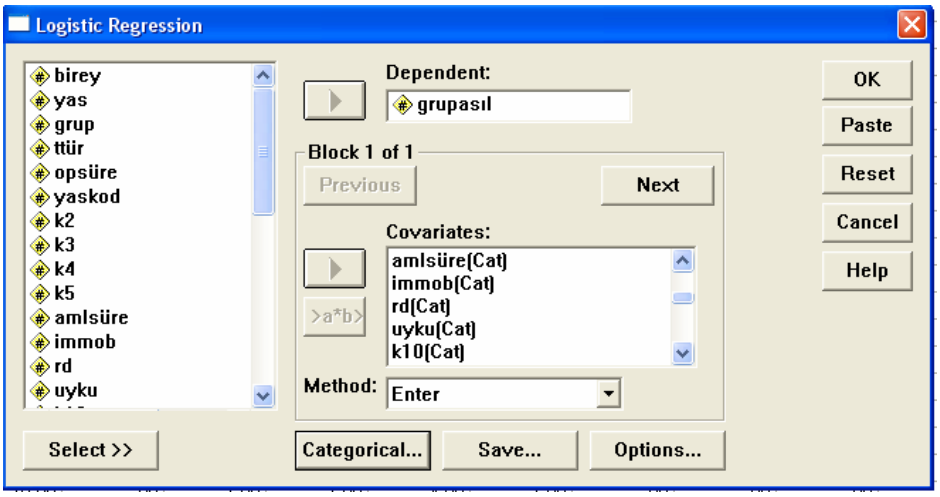
Categorical Variables Codings

		Frequency	Paramete (1)
K24	,00	456	1,000
	1,00	22	,000
YASKOD	,00	313	1,000
	1,00	165	,000
K2	,00	463	1,000
	1,00	15	,000
K3	,00	380	1,000
	1,00	98	,000
K4	,00	419	1,000
	1,00	59	,000
K5	,00	475	1,000
	1,00	3	,000
AMLSÜRE	,00	289	1,000
	1,00	189	,000
IMMOB	,00	428	1,000
	1,00	50	,000
RD	,00	388	1,000
	1,00	90	,000
UYKU	,00	476	1,000
	1,00	2	,000
K10	,00	416	1,000
	1,00	62	,000
K11	,00	383	1,000
	1,00	95	,000
K12	,00	338	1,000
	1,00	140	,000
CINS	1,00	321	1,000
	2,00	157	,000
ELEKTR	,00	447	1,000
	1,00	31	,000
STEROID	,00	450	1,000
	1,00	28	,000
KEMO	,00	462	1,000
	1,00	16	,000
K17	,00	401	1,000
	1,00	77	,000
K16	,00	364	1,000
	1,00	114	,000
DM	,00	440	1,000
	1,00	38	,000
AFILIN	,00	462	1,000
	1,00	16	,000
ARITMI	,00	444	1,000
	1,00	34	,000
TTÜR	,00	224	1,000
	1,00	254	,000

Spss programına nitel değişkenler Çizelge Ek 1.1 de görüldüğü gibi girilmiştir.

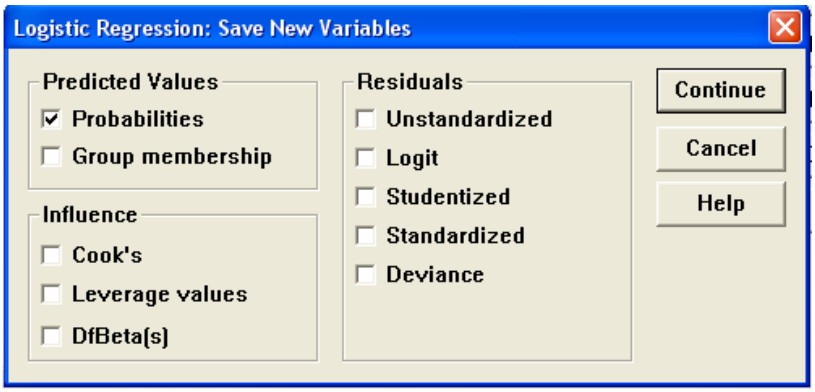
Ek 2 Propensity skorun Spss de hesaplanması

Spss programına veriler girildikten sonra, Analyze>Regression>Binary Logistic seçenekleri tıklanır. Şekil Ek 2.1 görüntülenir.



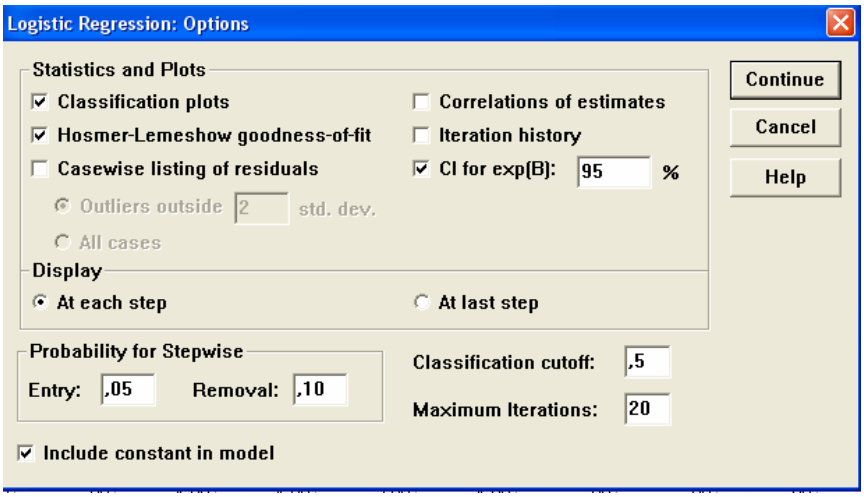
Şekil Ek 2.1 Lojistik regresyon işlem penceresi

'Dependent' kutucuğuna bağımlı değişken atanır. Ortak değişkenleri de 'Covariate(s)' kutucuğuna atadıktan sonra, ortak değişkenlerden nitel olanları 'categorical' butonuna basıldıktan sonra çıkan kutucuğa atanır. Lojistik regresyon hesaplanırken propensity skorun tahmin edilmesi için 'save' butonuna basılır, Şekil Ek 2.2 görüntülenir.



Şekil Ek 2.2 Propensity skor tahmin edilmesi için gerekli işlem penceresi

Görüntülenen (Şekil Ek 2.2) pencere de 'probabilities' seçeneği işaretlenir ve 'continue' butonuna basılarak propensity skor tahmini için gereken işlem yapılmış olur. Şekil Ek 2.1 deki ekrana geri dönülür. 'Options' butonuna basılarak lojistik regresyon analizinde sonuçları pekiştirecek seçenekler işaretlenir (Şekil Ek 2.3).



Şekil Ek 2.3 Lojistik regresyon 'options' penceresi

Continue> Ok butonları tıklanarak işlem sonlanır ardından çıktı ekranı ile veri sayfasında tahmin edilmiş propensity skorlar görüntülenir. Veri sayfasında propensity skorlar her birim için pre_1 isimli sütunda tahmin edilir (Şekil Ek 2.4).

	kemo	steroid	elektr	cins	otür	ybstay	pre_1
1	,00	,00	,00	1,00	2,00	,00	,01644
2	,00	,00	,00	1,00	4,00	,00	,01951
3	,00	1,00	,00	1,00	4,00	2,00	,02112
4	,00	,00	,00	1,00	3,00	,00	,00635
5	,00	,00	,00	1,00	2,00	3,00	,00768
6	,00	,00	,00	1,00	3,00	5,00	,05915
7	,00	,00	,00	1,00	3,00	8,00	,01570
8	,00	,00	,00	1,00	3,00	1,00	,07505
9	,00	,00	1,00	1,00	3,00	4,00	,13616
10	,00	,00	,00	,00	3,00	4,00	,03201
11	,00	,00	,00	1,00	4,00	,00	,04791
12	,00	,00	,00	1,00	2,00	3,00	,02646
13	,00	,00	,00	1,00	2,00	4,00	,03787
14	,00	,00	,00	1,00	2,00	,00	,01276
15	,00	,00	,00	1,00	2,00	2,00	,06215
16	,00	,00	,00	1,00	4,00	3,00	,00778
17	,00	,00	,00	,00	4,00	1,00	,00765
18	,00	,00	,00	,00	4,00	2,00	,00718
19	,00	,00	,00	,00	4,00	1,00	,01297
20	,00	,00	,00	,00	2,00	,00	,05889
21	,00	,00	,00	,00	4,00	,00	,13486
22	,00	,00	,00	,00	2,00	4,00	,05339
23	,00	,00	,00	,00	3,00	,00	,03009
24	,00	,00	,00	,00	2,00	2,00	,01042
25	,00	,00	,00	1,00	4,00	32,00	,06360
26	,00	,00	,00	,00	4,00	5,00	,01539
27	,00	,00	,00	,00	2,00	3,00	,08781
28	,00	,00	,00	1,00	2,00	,00	,01884
29	,00	,00	,00	,00	2,00	2,00	,05888
30	,00	,00	,00	1,00	2,00	,00	,00622

Şekil Ek 2.4 Tahmini propenstity skorun gösterildiği ekran

ÖZGEÇMİŞ

Doğum tarihi	27.03.1981	
Doğum yeri	İstanbul	
Lise	1992-1999	Özel Ata Lisesi
Lisans	1999-2003	Yıldız Üniversitesi Fen Edebiyat Fak. İstatistik Bölümü
Yüksek Lisans	2003-2006	Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik Yüksek Lisansı