



**T.C.
RECEP TAYYİP ERDOĞAN ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI**

**LOJİSTİK REGRESYON ANALİZİ: SAĞLIK
VERİLERİ ÜZERİNE BİR UYGULAMA**

(Yüksek Lisans Tezi)

Hilal POLAT

**Prof. Dr. Ali Sait ALBAYRAK
Danışman**

**RİZE
2017**

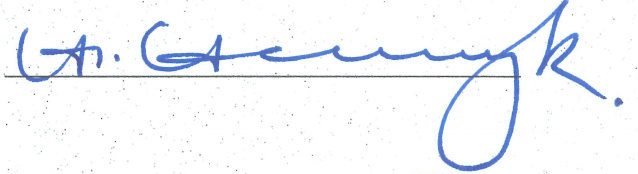
KABUL VE ONAY

Recep Tayyip Erdoğan Üniversitesi, Sosyal Bilimler Enstitüsü, İşletme Ana Bilim Dalında, Hilal POLAT tarafından hazırlanan “*Lojistik Regresyon Analizi: Sağlık Verileri Üzerine Bir Uygulama*” başlıklı bu çalışma, 06.07.2017 tarihinde yapılan savunma sınavı sonucunda oy birliği ile başarılı bulunarak jürimiz tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

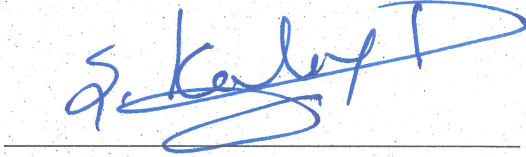
Jüri Üyesini Unvanı Adı-Soyadı

İmzası

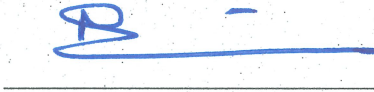
Prof. Dr. Ali Sait ALBAYRAK (Başkan)



Prof. Dr. Şeref KALAYCI (Üye)



Yrd. Doç. Dr. Mustafa SAVCI (Üye)



Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylıyorum.

...../...../2017

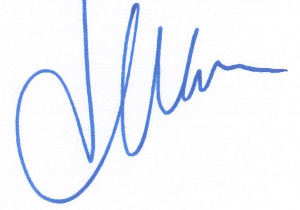


Yrd. Doç. Dr. Musa GÜN
Müdür V.

ETİK BEYAN

Bu tezdeki bütün bilgileri etik davranış ve akademik kurallar çerçevesinde elde ettiğimi ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yaptığımı bildiririm. İfade ettiklerimin aksi ortaya çıktığında ise her türlü yasal sonucu kabul ettiğimi beyan ederim. 06.07.2017

Hilal POLAT



ÖZET

Üniversite : Recep Tayyip Erdoğan Üniversitesi
Enstitü : Sosyal Bilimler Enstitüsü
Tezin Adı : Lojistik Regresyon Analizi: Sağlık Verileri Üzerine Bir Uygulama
Anabilim Dalı : İşletme
Tez Türü : Yüksek Lisans Tezi
Danışman : Prof. Dr. Ali Sait ALBAYRAK
Hazırlayan : Hilal POLAT
Yıl : 2017

Günümüz koşullarında yapılan birçok çalışma eldeki mevcut veriler kullanılarak ileriye dönük sonuçlar elde edebilmek üzerine kurulmuştur. Gerek sosyal bilimler gerekse sağlık alanındaki birçok çalışma bunu amaçlamıştır. Elde edilmek istenen sonuç, kullanılan istatistiksel yöntemin önemini arttırmıştır. Bu tezde, öngörü amaçlı istatistiksel analizler kullanılarak yapılan çalışmalarda çoğunlukla tercih edilen yöntem olarak karşımıza çıkan lojistik regresyon analizi detaylı olarak ele alınmıştır. Çalışmanın ilk bölümde lojistik regresyon analizi ile ilgili bilgiler verildikten sonra lojistik regresyon analizi yöntemlerinden iki kategorili lojistik regresyon analizi ve lojistik regresyon analizinin tercih edilme nedenleri üzerinde durulmuştur. Tek değişkenli ve çok değişkenli lojistik regresyon modellerinde parametre tahmin yöntemleri, katsayıların önem testleri, model katsayılarının yorumlanması, çoklu regresyon analizi, model seçim yöntemleri detaylı bir şekilde anlatılmış, lojistik regresyon analizi kullanılarak yapılan çalışmaların yer aldığı bir literatür özetine yer verilmiştir. Rahim kanseri tanısı almış ve retrospektif olarak taranarak hastane veri tabanından alınan, operasyon geçirmiş hastalar ile operasyon geçirmiş ancak kanser tanısı konmamış kontrol grubu arasında ileriye doğru değişken ekleme ve geriye doğru değişken eleme yöntemleri kullanılarak lojistik regresyon analizi uygulanmış ve analiz sonuçları karşılaştırılmıştır. Elde edilen veriler ile hastalık teşhisi için oldukça önemli sonuçlara ulaşılmıştır.

Anahtar Kelimeler: Lojistik Regresyon, Geriye Doğru Lojistik Regresyon, Eleme, İleriye Doğru Lojistik Regresyon, Adımsal Lojistik Regresyon Modeli.

ABSTRACT

University : Recep Tayyip Erdogan University
Institute : Institute of Social Sciences
Thesis Name : Logistic Regression Analysis: An Application on Health Data
Field of Study : Department of Business Administration
Thesis Type : Master's Thesis
Supervisor : Prof. Dr. Ali Sait ALBAYRAK
Author : Hilal POLAT
Year : 2017

Many studies conducted in today's conditions are based on the ability to obtain forward results using available data. This is the aim of many studies in the field of social and health sciences. The result to be obtained has increased the importance of the statistical method used. In this thesis, logistic regression analysis, which is mostly used as a preferred method in the studies carried out using statistical analysis for predictive purposes, was discussed in detail. In the first part of the study, some information on logistic regression analysis was given and then binary logistic regression analysis and the reasons why this method has been preferred were emphasized. Parameter estimation methods in univariate and multivariable logistic regression models, significance tests of coefficients, interpretation of model coefficients, multiple regression analysis, and model selection methods were explained in detail. In the following sections, a literature review of the studies using logistic regression analysis was given. Then, by using full and forward variable addition methods, logistic regression analysis was performed between the operated patients diagnosed with uterine cancer and selected from the hospital database by retrospectively screening and the patients in the control group operated but not diagnosed with cancer. Afterwards, the values found using the backward variable elimination method on the same data set were compared. Significant results for diagnosis of the disease were achieved with the obtained data.

Keywords: Logistic Regression, Backward Logistic Regression, Elimination, Forward Logistic Regression, Stepwise Logistic Regression Model.

ÖN SÖZ

Tez çalışmam sırasında kıymetli bilgi, birikim ve tecrübeleri ile bana yol gösterici ve destek olan, bu konuda tez çalışması yapma olanağı sunan değerli danışman hocam sayın Prof. Dr. Ali Sait ALBAYRAK'a sonsuz teşekkür ve saygılarımı sunarım.

Anaokulundan üniversiteye tüm eğitim hayatım boyunca destek olan tüm öğretmenlerime ve akademisyen hocalarıma çok teşekkür eder, saygılarımı sunarım.

Çalışmalarım boyunca maddi manevi destekleriyle beni hiçbir zaman yalnız bırakmayan annem, babam, eşim ve kızıma da sonsuz teşekkürler ederim.

Hilal POLAT
Rize 2017

İÇİNDEKİLER

	<u>Sayfa</u>
KABUL VE ONAY	2
ETİK BEYAN	3
ÖZET	4
ABSTRACT.....	5
ÖN SÖZ	6
İÇİNDEKİLER	7
TABLolar LİSTESİ	10
ŞEKİLLER LİSTESİ	12
KISALTMALAR LİSTESİ.....	13
GİRİŞ	14

BİRİNCİ BÖLÜM

1.1. LOJİSTİK REGRESYON ANALİZİ	16
1.1.1. Lojistik Regresyon Analizinin Tanımı ve Tarihçesi.....	16
1.1.2. Lojistik Regresyon Analizinin Tercih Nedenleri.....	18
1.1.3. İki Kategorili Lojistik Regresyon Analizi.....	19
1.1.3.1. Lojistik Regresyon Modeli	19
1.1.3.2. Katsayı Tahmin Yöntemleri ve En Çok Olabilirlik Yöntemi	22
1.1.3.3. Katsayıların Önemliliğinin Test Edilmesi	23
1.1.3.4. Lojistik Sınıflandırma ve Katsayıların Yorumlanması.....	25
1.1.3.5. Değişkenlerin Lojistik Regresyon Analizine Dâhil Edilmesi.....	26
1.1.3.6. Çoklu Doğrusal Regresyon Analizi	27
1.1.3.7. Model Geliştirme Yöntemleri	29
1.1.3.7.1. Tam Regresyon Yöntemi	30
1.1.3.7.2. İleri Doğru Değişken Seçim Yöntemi.....	30
1.1.3.7.3. Geriye Doğru Değişken Seçim Yöntemi	31
1.1.3.7.4. Adımsal Değişken Seçim Yöntemi.....	31
1.1.3.8. Lojistik Regresyon Analizinde İlişki Ölçüsü.....	31
1.1.3.9. Uyum İyiliği.....	32

1.1.3.10. Lojistik Regresyon Model Uygunluğunun Değerlendirilmesi.....	33
1.1.3.10.1. Standart Olmayan Hatalar.....	33
1.1.3.10.2. Standart Hatalar	34
1.1.3.10.3. Sapma (Deviance) Değerleri.....	34
1.1.3.10.4. Gözlem Uzaklığı (Leverage Values)	34
1.1.3.10.5. Cook Uzaklığı.....	35

İKİNCİ BÖLÜM

2.1. LOJİSTİK REGRESYON ANALİZİ KULLANILARAK YAPILAN ÇALIŞMALAR:LİTERATÜR İNCELEMESİ	36
---	-----------

ÜÇÜNCÜ BÖLÜM

3.1. LOJİSTİK REGRESYON ANALİZİ: SAĞLIK VERİLERİ ÜZERİNE BİR UYGULAMA	42
--	-----------

3.1.1. Lojistik Regresyon Analizi Öncesi Yapılan Çalışmalar ve Kullanılan Değişkenler	42
3.1.2. Araştırmada Kullanılan Önemli Değişkenlerin Tanımsal İstatistikleri ...	44
3.1.3. SPSS ile Lojistik Regresyon Analizinin Elde Edilmesi.....	49
3.1.3.1. Model Seçim Yöntemleri ve Uygunluğunun Değerlendirilmesi	54

BULGULAR.....	57
----------------------	-----------

SONUÇ VE ÖNERİLER.....	66
-------------------------------	-----------

KAYNAKLAR	68
------------------------	-----------

EKLER	75
--------------------	-----------

Ek-1 Hastalığın Derecesi	75
Ek-2 Ortalama Tümör Boyutu	75
Ek-3 Miyometrial İnvazyon.....	75
Ek-4 Doku Türü (Histoloji)	75
Ek-5 Servikal Stromal Tutulum	75
Ek-6 Adneksiyal Tutulum.....	75
Ek-7 Lenfovasküler İnvazyon.....	76
Ek-8 Pelvik Lenf Tutulumu	76
Ek-9 Paraaortik Lenf Nodu Tutulumu	76
Ek-10 Kanser Hastalığının Evresi.....	76
Ek-11 Peritoneal Sistoloji	76

Ek-12 Lökosit Düzeyi	76
Ek-13 Danışman Hekim Tarafından Risk Değerine Göre Yeniden Kodlanan Operasyon Öncesi Lökosit Değerleri (WBC5).....	77
Ek-14 Tanımlayıcı Grafikler.....	77
ÖZ GEÇMİŞ	83



TABLolar LİSTESİ

	<u>Sayfa</u>
Tablo 1.1 Üç Kategorili Bir Değişkenin Kukla Değişken Olarak Kodlanması	29
Tablo 3.1 Araştırmada Kullanılan Değişkenlerin Simgeleri ve Açıklamaları	43
Tablo 3.2 Hasta ve Kontrol Grubunun Yaş Ortalaması ve Standart Sapma Değerleri	44
Tablo 3.3 Hasta ve Kontrol Grubu Operasyon Öncesi ve Sonrası Lökosit ve Trombosit ile Yaş Değişkenlerinin Tanımsal İstatistikleri	45
Tablo 3.4 Yaş, Lökosit ve Trombosit Değerleri İçin Bağımsız İki-Örnek T-Testi	46
Tablo 3.5 Hasta Grubu Lökositöz ve Trombositöz Düzeyleri Frekans Tablosu ...	46
Tablo 3.6 Hasta Grubu Operasyon Öncesi Lökositöz Değişkeninin Aritmetik Ortalaması, Ortancası, Tepe Değeri ve Standart Sapması	47
Tablo 3.7 Lökositöz ve Evre, Paraaortik Lenf, pelvik Lenf Nodu Tutulumu, Andeksiyel Tutulum, Histoloji Arasında Ki-Kare Analizi Sonuçları	48
Tablo 3.8 Analizde Kullanılan Kategorik Değişkenlerin Kategorik Kodlamaları ve Frekansları	49
Tablo 3.9 Lojistik Regresyon Analizi İlk Adım Sınıflandırma Tablosu	50
Tablo 3.10 Lojistik Regresyon Analizi İleriye Doğru Değişken Ekleme Metodu Sınıflandırma Tablosu	51
Tablo 3.11 İleriye Doğru Değişken Ekleme Metodu Lojistik Regresyon Analizi Adımlar ve Güven Aralığı Değerleri	52
Tablo 3.12 Geriye Doğru Değişken Eleme Metodu Lojistik Regresyon Analizi Sınıflandırma Tablosu	53
Tablo 3.13 Geriye Doğru Değişken Eleme Metodu Lojistik Regresyon Analizi Sınıflandırma Tablosu	53
Tablo 3.14 Geriye Doğru Değişken Eleme Metodu Lojistik Regresyon Analizi Adımlar ve Güven Aralığı Değerleri	54
Tablo 3.15 İleri Doğru Değişken Ekleme Model Özeti	55

Tablo 3.16 Geriye Doğru Değişken Eleme Metodu Model Özeti	56
Tablo 4.1 Lökositöz Modele Dâhil Edildiğinde Oluşan Sonuçlar	58
Tablo 4.2 Lökositöz ve Grup Değişkeni Arasında Lojistik Regresyon	58
Tablo 4.3 Lökositöz Olan ve Olmayan Hasta ve Kontrol Grubu Çapraz Tablosu	58
Tablo 4.4 Lökositöz, WBC1 ve WBC3 Değişkenleri ile Lojistik Regresyon Analizi Sonucu Oluşan Durumlar (1.Adım)	59
Tablo 4.5 Lökositöz, WBC1 ve WBC3 Değişkenleri ile Lojistik Regresyon Analizi Test Uyum İyiliği (Hosmer-Lemeshow Testi).....	59
Tablo 4.6 Sınıflandırma Tahmin Edilebilirliği	60
Tablo 4.7 Grup Değişkeni ile Lökosit, WBC1, WBC3 Lojistik regresyon Analizi	60
Tablo 4.8 Lökositöz ile WBC1, WBC2, TRM1, TRM2, WBC3 ve WBC5 Değişkenleri ile Sınıflandırma Sonuçları.....	60
Tablo 4.9 WBC3 Değişkeni ile WBC5 Arasında Lojistik Regresyon Analizi	61
Tablo 4.10 Lökositöz Sınıflandırma Tahmin Tablosu	62
Tablo 4.11 Lökositöz ve Evre Değişkeni Arasında Lojistik Regresyon Analizi Sonuç.....	62
Tablo 4.12 Miyometrial İnvazyon ve Evre Arasında Lojistik Regresyon Analizi	62
Tablo 4.13 Geriye Doğru Eleme Yöntemi İle Lojistik Regresyon Analizi Sonuçları	63
Tablo 4.14 Tüm Değişkenler Dâhil İken Geriye Doğru Değişken Eleme Metodu Sınıflandırma Sonuçları	64
Tablo 4.15 Geriye Doğru Değişken Eleme Yöntemi Lojistik Regresyon Analizi	64

ŞEKİLLER LİSTESİ

	<u>Sayfa</u>
Şekil 1.1 Lojistik Regresyon Grafiği	20
Şekil 3.1 Hasta Olgularda Ortalama Tümör Boyutları	47
Şekil 3.2 Çalışmaya Alınan Rahim Kanserli Olguların Cerrahi Evreleri	48



KISALTMALAR LİSTESİ

Anl.	: Anlamlılık Düzeyi
AO	: Aritmetik Ortalama
ASD	: Güven Aralığı Alt Sınır Değeri
BDÖ	: Beck Depresyon Ölçeği
Bk.	: Bakınız
CE	: Sisplatin-Etoposid
ENB	: En Büyük
ENK	: En Küçük
GA	: Güven Aralığı
İMKB	: İstanbul Menkul Kıymetler Borsası
KHDAK	: Küçük Hücreli Dışı Akciğer Karsinomu
KOR	: Koroner Arter Hastalığı
MIC	: Mitomisin-İfosfamid-Sisplatin
OAF	: İki Ortalama Arasındaki Fark
SBTÖ	: Stresle Başa Çıkma Tarzları Ölçeği
Sd	: Serbestlik Derecesi
SH	: Standart Hata
SS	: Standart Sapma
ÜSD	: Güven Aralığı Üst Sınır Değeri

GİRİŞ

Birçok alanda yapılan bilimsel çalışmalarda, özellikle sağlık ve sosyal bilimlerde incelenen olayların karmaşık ve olayı açıklamada kullanılabilir değişkenlerin sayısı fazla olabilmektedir. Araştırmacılar, birden fazla değişkeni inceleyip, değişkenler arasındaki ilişkilerin yorumlanması, üzerinde çalıştıkları konuda etken olması durumunda tek tek bağımlı değişken üzerine etkisi yanında, bunların birlikte etkisini de öğrenmek ya da incelemek istemektedir. Bu istek değişkeni inceleyip, değişkenler arasındaki anlamlı ya da anlamsız ilişkilerin yorumlanmasında kolaylıklar sağlayan çok değişkenli istatistiksel analizlerin kullanılmasının gerekliliğini ortaya çıkarmıştır.

Çok değişkenli istatistiksel analizlerle ilgili çalışmalar 1900'lü yıllarda başlamıştır (Koçak,1997). Yapılan çalışmalar gün geçtikçe artmış, ihtiyaçlar farklılaşmış ve analiz ile elde edilmek istenen sonuçlar değişiklik göstermiştir. Bu çalışmada birbirleriyle ilişkili çok sayıda değişkenin oluşturduğu sistemin yapısı belirlenmeye çalışılıp, daha basit bir yapıya indirgendikten sonra elde edilen bilgilerin değerlendirilmesi, yorumlanması ve sonuçta bir karara ulaşılması hedeflenmektedir.

Veri setinde bulunan değişkenler ve bu değişkenlerin özellikleri kullanılacak yöntemin seçiminde büyük öneme sahiptir. Bir bağımlı değişken ile bir ya da daha fazla bağımsız değişken arasındaki ilişki incelenmek istenirse regresyon analizi yöntemleri tercih edilebilmektedir. Regresyon analizinin iki alt başlığı olan basit ve çoklu doğrusal regresyon analizi yöntemlerinde bağımlı değişken metrik (aralık veya oran) ölçekli bir değişken tipindedir. Buna karşılık, çalışmalarda bağımlı değişkenin kategorik veri tipinde olma durumu ile de sıklıkla karşılaşmaktadır. Bu durumda, doğrusal regresyonda parametre kestirimlerini hesaplamak için kullanılan en küçük kareler yönteminden yararlanmak bu yöntemle ilgili varsayımlar sağlanmadığı için uygun olmamaktadır. Bu nedenle bağımlı değişken iki ya da ikiden çok kategorili nitel veri tipinde olduğunda lojistik regresyon yöntemi ile çözümleme

gerçekleştirilebilmektedir. Lojistik regresyon analizi yöntemi çeşitli varsayımlardan (normallik, sabit varyansa vs.) sapmalar olması durumunda diskriminant analizi ve çapraz tablo analizine alternatif olarak kullanılmaktadır (Alpar, 2013).

Değişkenlerin tipi ve dağılımı ile ilgili varsayımların az olması ve sonuçların daha kolay yorumlanabilmesi gibi nedenlerle lojistik regresyon son yıllarda sıklıkla kullanılan bir yöntem haline gelmiştir.

Epidemiyolojik araştırmalarda incelenen hastalığa neden olan faktör veya faktörler ile açıklanan değişken arasındaki ilişki bir hastalığın tespit edilmesi ve tedavi sürecine erken geçilmesi açısından oldukça önem arz etmektedir. Bir örnek verecek olursak; rahim kanseri tanısı almış hastalar ve diğer nedenlerden hastaneye başvurmuş ancak kanser tanısı almamış kontrol grubunu ele aldığımızda kanser tanısı konmuş hastalar 1 değerini, kontrol grubu olguları da 0 değerini alarak bağımlı değişkeni oluşturursun. Hastalıkla ilgili yaş, kilo, sigara içme durumu, trombosit (kanın pıhtılaşmasını sağlayan hücresi) ve lökosit (kandaki beyaz küre, akyuvar) değerleri gibi diğer değişkenlerde bağımsız değişken olarak kullanılan bir modelde lojistik regresyon analizi kullanılarak aynı anda tüm değişkenlerin bağımlı değişken üzerindeki etkisi incelenebilir. Aynı anda tüm bağımsız değişkenlerin bağımlı değişken üzerindeki pozitif ya da negatif etkisinin görülebilmesi zaman kaybının azalması ve maliyetin düşürülmesi açısından oldukça önemlidir.

Mevcut veriler göz önünde bulundurularak bu tezde lojistik regresyon analizi detaylı olarak ele alınmış, teorik olarak ne olduğu, nasıl yapıldığı ve neden tercih edildiği anlatılmıştır. Bir literatür taraması yapılmış ve çalışma özetlerine yer verilmiştir. 19 Mayıs Üniversitesi Tıp Fakültesi Kadın Hastalıkları ve Doğum Kliniğine 2005-2013 yılları arasında gelen hastalardan elde edilmiş sağlık verileri kullanılarak lojistik regresyon analizi uygulanmış ve en uygun model bulunmaya çalışılmıştır. İleriye doğru değişken ekleme ve geriye doğru değişken eleme metotları kullanılarak lojistik regresyon analizi gerçekleştirilmiş lökosit, trombosit gibi kan hücrelerinin ve yaş değişkeninin hasta grubu tespit etmede önemli bir değişken olarak kullanılabileceği sonucuna ulaşılmıştır.

BİRİNCİ BÖLÜM

1.1. LOJİSTİK REGRESYON ANALİZİ

Belirli sayıda özelliği bilinen değişkenleri, bu özelliklerine göre önceden bilinen veya sonradan belirlenen gruplara atamak amacıyla kullanılan bir kaç yöntem bulunmaktadır. Gözlemlerin gruplara ayrılmasında en sık kullanılan geleneksel yöntemler arasında kümeleme analizi, diskriminant analizi ve lojistik regresyon analizi yer almaktadır. Kümeleme analizinde gözlemlerin atanacağı grup sayısı önceden bilinmemektedir. Bu analiz tekniğinde gözlemler ya da değişkenler uzaklık-benzerlik ölçütlerine göre gruplandırılırlar. Buradaki amaç yalnızca gözlemlerin oluşturduğu grupların yapısını anlamaktır. Diskriminant analizi ile lojistik regresyon analizinde ise grup sayısı önceden belirlidir ve mevcut veriler kullanılarak geliştirilen en uygun modelle birimler yeniden sınıflandırılarak bir sınıflandırma fonksiyonu elde edilir. Elde edilen model yardımıyla, yeni elde edilen ya da edilecek olan gözlemlerin hangi gruba atanacağı belirlenebilmektedir (Başarır, 1990; Alpar,2013).

1.1.1. Lojistik Regresyon Analizinin Tanımı ve Tarihçesi

Lojistik regresyon; açıklayıcı değişkenlere göre bağımlı değişkenin önceden tanımlanan gruplara atama olasılıklarını hesaplayabilen çok değişkenli bir sınıflandırma tekniğidir (Powers ve Yu,2000).

Lojistik regresyon analizi, temelde regresyon analizi olmakla birlikte bir sınıflandırma analizi olma özelliğini de taşımaktadır. Bu sebeple lojistik regresyon analizini klasik regresyon analizinden ayıran üç önemli farklılığı vardır. Bu farklılıklar şunlardır:

1. Regresyon analizinde bağımlı değişken metrik ölçekli iken, lojistik regresyon analizinde nominal ölçekli bir değişkendir.
2. Regresyon analizinde bağımlı değişkenin birim değeri, lojistik regresyon analizinde ise bağımlı değişkenin alabileceği değerlerden birinin gerçekleşme olasılığı tahmin edilir.

3. Regresyon analizinde bağımsız değişkenlerin çoklu normal dağılım göstermesi şartı aranırken, lojistik regresyon analizinin uygulanabilmesi için bağımsız değişkenlerin bir kuramsal dağılıma uygunluk varsayımı yapılmaz(Atakurt, 1999).

Lojistik modelinin kullanılması 1845’li yıllara kadar dayanmaktadır. Daha çok sosyoekonomik ağırlıklı konuların incelenmesinde kullanılan lojistik analiz, ilk önceleri toplumdaki nüfus artışının, matematiksel bir ifadeyle açıklanmasına yönelik çalışmalarda kullanılmıştır (Gürcan, 1998).Lojistik modelin biyolojik deneylerin analizi için kullanımı ilk olarak Berkson (1944) tarafından önerilmiştir. Cox (1970) biyolojik deneylerin analizi modelini gözden geçirerek çeşitli uygulamalar yapmıştır. Truett vd. (1967) ile Halperin vd. (1971) lojistik regresyonun normal dağılım varyansının sağlanamadığı durumda diskriminant analizine alternatif olarak önermişlerdir. Lojistik regresyon modellerinin yaygın bir şekilde kullanılabilir hale gelmesi, katsayı tahmin yöntemlerinin geliştirilmesi ve lojistik regresyon modellerinin daha ayrıntılı incelenmesine sebep olmuştur. Cornfield (1962), lojistik regresyondaki katsayı tahmin işlemlerinde diskriminant fonksiyonu yaklaşımını ilk kez kullanarak popüler hale getirmiştir. Lee (1984) basit dönüşümlü (cross-over) deneme planları için doğrusal lojistik modeller üzerinde durmuştur. Bonney (1987) lojistik regresyon modelinin kullanımı ve geliştirilmesi üzerinde çalışmıştır. Lojistik regresyon analizinin özellikle son 20 yıldır askeri konularda, meteorolojide, iç göç hareketlerinde ve eğitim alanında kullanımının arttığı görülmektedir. Bu artışın en önemli nedenlerinden biri, istatistik paket programlarının kullanımının yaygınlaşmasıdır. Ancak yine de en yaygın kullanıldığı alanlardan birinin tıp olduğu gözlenmektedir (Seven, 1997). Örneğin bir tümörün kanser olup olmadığını ayırt etmeyi sağlayan bir model oluşturmak, bir hastalığın tedaviye cevap verip vermediğini belirlemek gibi farklı amaçlarla sıkça kullanılmaktadır (Field, 2005).

Bu ve buna benzer birçok örneğe 1800’lü yıllardan günümüze kadar yapılan çalışmalarda, lojistik regresyon analizinin model seçiminde kullanıldığına rastlanmaktadır.

1.1.2. Lojistik Regresyon Analizinin Tercih Nedenleri

Kategorik verilerin çok deęişkenli istatistiksel analizi, hemen hemen tüm alanları ilgilendirmektedir. Lojistik regresyon analizi, regresyon tipi mantığıyla ve dięer analizlere göre daha kullanışlı olmasının sağladığı avantaj sebebiyle, kategorik veri analizinde önemli bir yere sahiptir.

Basit ve çoklu doğrusal regresyon analizleri, bağımlı deęişken ile bağımsız deęişken ya da deęişkenler arasındaki ilişkiyi analiz etmede kullanılmaktadır. Bu yöntemlerin uygulanabileceği veri setlerinde bağımlı deęişkenin normal dağılım göstermesi, bağımsız deęişkenlerin normal dağılım gösteren deęişken ya da deęişkenlerden oluşması ve hata terimlerinin varyansının normal dağılım göstermesi gerekir. Bu ve benzeri koşulların sağlanamaması durumunda ise, basit ya da çoklu doğrusal regresyon analizi kullanılamaz(Kılıç, 2000). Bilindiği gibi regresyon analizinin bağımlı deęişkeni kategorik bir deęişken olduğunda normal dağılımı ve hata terimlerinin sabit varyanslı olduğu varsayımları ihmal edilmiş olur. Lojistik regresyon analizinde de doğrusal regresyon analizinde olduğu gibi bazı deęişken değerlerine dayanarak tahmin yapılmaya çalışılır. Doğrusal regresyon analizinde bağımsız deęişkenin çoklu normal dağılım göstermesi şartı aranırken, lojistik regresyon analizinde böyle bir şart yoktur. Bir dięer husus olan örneklem büyüklüğü için, özellikle bağımlı deęişkenin ikiden fazla kategorisinin olduğu durumlarda, geçerli bir hipotez testi için her bağımsız deęişkende en az 50 kişilik bir grup büyüklüğüne ihtiyaç vardır. Bazı kaynaklarda bu sayının her bağımsız deęişken için minimum 20, toplamda minimum 60 olması gerektiği vurgulanmaktadır(Akkuş ve Çelik, 2004; Tatlıdil, 1996; Tabachnick ve Fidell, 1996). Dolayısıyla lojistik regresyonun dięer iki teknikten daha esnek olduğu ifade edilmektedir.

Örneklem büyüklüklerinin aynı olması durumunda, bağımlı deęişkenin her bir kategorisinde bağımsız deęişkenlerin çok deęişkenli normalliğe sahip olması, her bir kategori için varyans ve kovaryansların eşitliği kabullerinin karşılanması durumunda, daha önce de değinildiği gibi diskriminant analizi, lojistik regresyon analizine tercih edilmektedir. Bununla birlikte, lojistik regresyon analizi ile yapılan

çözümlemeden elde edilen ilişkinin yorumlanmasının daha kolay olduğunu belirtmekte yarar vardır(Akkuş, 2004). Belirtilen bu gibi durumlarda çoğu zaman lojistik regresyon analizi kullanılmaktadır.

1.1.3. İki Kategorili Lojistik Regresyon Analizi

Bağımlı değişkenin iki kategorili olduğu lojistik regresyon yöntemidir. Kodlama yapılırken, sıklıkla riskin olmadığı durum için 0 ve riskin olduğu durum için 1 kodu kullanılır. Bağımsız değişkenlerin tipi için herhangi bir kısıtlama yapılmaz sürekli, kesikli ya da sıralı veri tipinde olabilir.

Bağımsız değişkene ilişkin veri matrisi $X=x$ olduğunda yani X 'in değeri bilindiğinde olayın olma olasılığı ya da $Y=1$ olması olasılığı π ile gösterilsin. π 'yi tanımlamak için standart doğrusal model kullanılarak Eşitlik 1.1'deki model oluşturulabilir.

$$\pi(x) = P(Y=1 / X=x) = \beta_0 + \beta_1 x + e \quad (1.1)$$

Bu doğrusal fonksiyonun sonucu 0 ile 1 arasında olmadığından olasılığı modellemek için kullanılamaz. Burada doğrusal regresyonda katsayı kestirimi için kullanılan en küçük kareler yönteminin kullanılamamasının nedeni varyansların eşitliği varsayımı sağlanamamasıdır. Bu sorunu çözebilmek için ağırlıklandırılmış en küçük kareler yaklaşımı kullanılabilir; ancak bu yaklaşımla ilgili de sorunlar vardır. π 'nin değeri bilinmez. Oysa ağırlıklandırılmış en küçük kareler yaklaşımında π 'nin tahmini başlangıç değeri ile başlanması ve tekrarlanması ve yinelemeli yöntemle sonuca gidilmesi gerekmektedir. Bu yöntemler yerine olasılığı modellemek için seçenek yöntemi tanımlanmıştır(Alpar,2013:674-715).

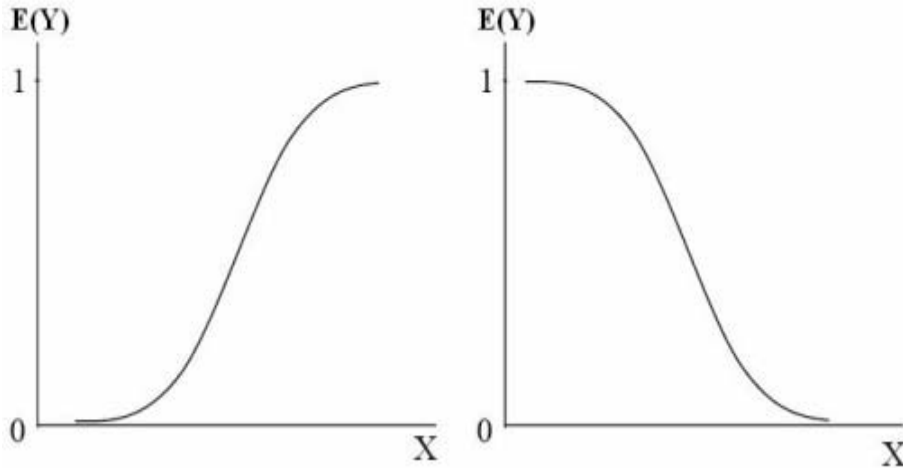
1.1.3.1. Lojistik Regresyon Modeli

Lojistik regresyon modeli, genel doğrusal modellerin binom dağılımlı bağımlı değişkenler için elde edilmiş olan özel bir biçimidir. Hem teorik hem de deneysel incelemeler bağımlı değişken iki sonuçlu iken cevap fonksiyonunun $\pi/(1-\pi)$ şeklinin S veya ters S şeklinde olacağını göstermiştir (Bk. Şekil 1.1). Başlangıçta X değerlerindeki artışla birlikte π değerleri yavaşça artar, sonra artış giderek

hızlanır ve durağanlaşır. Fakat hiçbir zaman 1'in üstüne çıkmaz. Bağımsız değişkenler ile olasılık arasındaki ilişki doğrusal değildir(Kaşko,2007).

Lojistik fonksiyonun 0 ile 1 arasında bir değişim aralığına sahip olması lojistik fonksiyonun tercih edilmesindeki en önemli nedendir. Lojistik model, ortaya çıkacak riski 0 ile 1 arasında herhangi bir değer olarak tahmin etmeye yarar. Başka bir deyişle 1'in üstünde veya 0'in altında bir risk olmaz. Lojistik regresyon modeli Eşitlik 1.2. ile belirtilmiştir. Bu bağıntıya göre $X = x$ olduğunda Y 'nin 1 olma olasılığı π olarak bulunur.

$$\pi(x) = P(Y = 1 / X = x) = \frac{e^{(\beta_0 + \beta_1 x)}}{1 + e^{(\beta_0 + \beta_1 x)}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \quad (1.2)$$



Şekil 1.1 Lojistik Regresyon Grafiği

$$\text{lojit}\pi(x) = g(x) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) \quad (1.3)$$

Eşitlik 1.3'ten bahis (odds) oranının doğal logaritması alınarak doğrusal modele dönüştürülerek lojistik regresyon modeli oluşturulur.

$$g(x) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \ln e^{(\beta_0 + \beta_1 x)} = \beta_0 + \beta_1 x \quad (1.4)$$

Eşitlik 1.4'teki modelin ayrıntılı açıklaması ilk olarak tek bağımsız değişken içeren model için yapılmaktadır. Tek değişkenli model Eşitlik 1.4 veya Eşitlik 1.5'deki gibi yazılır.

$$\ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x_i \quad (1.5)$$

Buradan π hastalık durumunu gösterirken $1-\pi$ hastalık bulunmama durumunu göstermektedir. Eşitlik 1.5'deki sonucun elde edilmesi ile lojit modelin aşağıdaki özellikleri yazılabilmektedir (Albayrak, 2006).

- ✓ Olasılıklar sıfırdan bire doğru büyürken lojit fonksiyonu $-\infty$ ile $+\infty$ arasında değerler almaktadır.
- ✓ Modelin bağımlı değişkeni ile bağımsız değişkenleri arasında doğrusal bir ilişki olduğu halde, olasılıklarla bağımlı değişkenler arasında ilişki doğrusal değildir.
- ✓ Bağımsız değişkenlerin verilen özelliklerine göre olasılık değerleri ise Eşitlik 1.6 yardımıyla hesaplanmaktadır.

Bir örnek ile model oluşturacak olursak, hastalık değişkeni olarak; 1 ve 0 olarak kodlanan histolojik tipi örnek alalım. Bir (1) hastalığın varlığını, sıfır (0) ise hastalığın olmadığı durumu gösterebilir. Bir grup hasta üzerinde çeşitli ölçümler yapılsın. Örneğin X_1 yaşı, X_2 sigara kullanma durumunu, X_3 sitolojisi olsun. Bu bireyler Cohort çalışmalarında olduğu gibi belirli bir zaman gözlemlensin. X_1, X_2, X_3 gibi bağımsız değişkenlere sahip hasta olmayan bireylerin belirli bir süre boyunca izlenmesiyle, hastalığın ortaya çıkma ihtimali açıklanmak istensin. Koşullu ihtimal, $\pi(Y=1/X_1, X_2, X_3)$ ile gösterilsin. Bu durumda lojistik model Eşitlik 1.6'daki gibi yazılır.

$$P(Y = 1 / X_1, X_2, X_3) = \pi = \frac{1}{1 + e^{-(\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3)}} \quad (1.6)$$

Burada Y ; hastalık değişkenini yani bağımlı değişkeni ifade etmektedir. β_i 'ler bilinmeyen parametreler olup, veriler yardımıyla tahmin edilmektedir.

1.1.3.2. Katsayı Tahmin Yöntemleri ve En Çok Olabilirlik Yöntemi

İki gruplu bir lojistik regresyon modelinin katsayılarının tahmin edilmesinde; en çok olabilirlik yöntemi, yeniden ağırlıklandırılmış en küçük kareler yöntemi, minimum lojit ki-kare yöntemi yaygın olarak kullanılan yöntemlerdendir. En çok olabilirlik yöntemi, gözlenen veri kümesini elde etmenin olasılığını maksimum yapan bilinmeyen parametrelerin değerlerini verir. Bu metodu uygulamak için önce en çok olabilirlik fonksiyonunun oluşturulması gerekmektedir. Bu parametrelerin en çok olabilirlik tahminleri, fonksiyonu maksimum yapacak şekilde seçilir. Bu fonksiyonu en büyük yapan değerler, bilinmeyen parametrelerin en çok olabilirlik kestiricileridir. Yani en çok olabilirlik yönteminde bir olayın olma olasılığı maksimum yapılmaya çalışılır. Böylece tahminler gözlenen değerlere en yakın olan değerler olarak seçilir(Elhan, 1997).

İki kategorili lojistik regresyon modelinde bağımlı değişken iki değer alır. Y değeri 0 ya da 1 olacak şekilde rastgele iki değer alabilir. Bu değerlerin hesaplanması olabilirlik fonksiyonu tarafından Eşitlik 1.7'deki şekilde ifade edilmektedir.

$$\pi(Y = 1) = \pi = \frac{e^{\beta x}}{1 + e^{\beta x}} \quad (1.7)$$

Bu eşitlikte β katsayıların vektörüdür ve X bağımsız değişkenlerin vektörü olarak ifade edilmektedir. Bağımsız değişkenlerin bir doğrusal fonksiyonu olarak Eşitlik 1.8'de yeniden ifade edilmektedir.

$$\ln\left(\frac{\pi}{1 - \pi}\right) = \beta x \quad (1.8)$$

Bağımlı değişkenin değerleri bilinmediği için bu fonksiyon ile doğrudan tahmin edilememektedir. Ancak olabilirlik fonksiyonu bu probleme bir çözüm sunmaktadır. Her gözlem öncelikle Bernoulli denemesi olarak kabul edilir. Yani çalışmaların toplam sayısı 1'e eşit olan sonuçların iki terimli olduğunu varsayarak n tane gözlemi Eşitlik 1.9'da, olabilirlik fonksiyonunu Eşitlik 1.10'da ve logaritmik olabilirlik fonksiyonu da Eşitlik 1.11'de verilmektedir.

$$P(Y = y_i) = p_i^{y_i} (1 - p_i)^{1-y_i} \quad (1.9)$$

$$L = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1-y_i} = \prod_{i=1}^n \left(\frac{e^{bX}}{1 + e^{bX}} \right)^{y_i} \left(\frac{1}{1 + bX} \right)^{1-y_i} \quad (1.10)$$

$$\ln(L) = l = \sum_{i=1}^n y_i \left(\frac{e^{\beta X}}{1 + e^{\beta X}} \right) + \sum_{i=1}^n (1 - y_i) \left(\frac{1}{1 + e^{\beta X}} \right) \quad (1.11)$$

Burada β parametreler vektörü Eşitlik 1.10'daki gibi maksimize edilmektedir. Olağan prosedür her denklemi sıfıra eşitlemek ve daha sonra ortaya çıkan denklemi çözmek, her bir bağımsız değişkene göre denklemin yeniden birinci türevini almaktır. Ancak ortaya çıkan denklemlerin analitik bir çözümü yoktur. Sonuç olarak $\hat{\beta}$, Newton-Raphson gibi yinelemeli yöntemler kullanılarak en üst düzeye çıkarılarak elde edilir (Sharma, 1996).

Örneklem büyüklüğü sonsuza yaklaştıkça, en çok olabilirlik yöntemi tutarlıdır, yeterlidir ve asimtotik olarak normal dağılır. Örneklem büyüklüğü küçük olduğunda da kötü sonuçlar vermediği belirtilmiş fakat değişken sayısının en az 10 katı kadar gözlem olmasının gerekliliğine değinilmiştir. Bu yöntemin yanında en küçük kareler yöntemi, yinelemeli olarak uygulanacak bir yöntem olup, her adımda ağırlıklar tahminlere bağlı olarak yeniden elde edilir (Tatlidil, 1996). Minimum lojit ki-kare yöntemi ise lojistik regresyon analizinde tekrarlı durumlar mevcutsa kullanılmaktadır. Lojit değeri olarak tanımlanan bağımlı değişkenin bağımsız değişkenler üzerinde ağırlık değeri ile ağırlıklandırılmış regresyondan en küçük kareler tahminlerini elde etmeye dayanır (Ünal, 1997).

1.1.3.3. Katsayıların Önemliliğinin Test Edilmesi

Modeldeki değişkenlerin önemliliği; Olabilirlik Oranı, Wald ya da Skor testlerinden biri kullanılarak gerçekleştirilir. Lojistik regresyonda katsayıların önem testi için ana prensip sorgulama altındaki değişkeni içeren ve içermeyen modellerden elde edilen tahmin değerlerinin sonuç değişkeninin gözlenen değerleriyle karşılaştırılmasıdır. Gözlenen ve tahmin edilen karşılaştırma işlemi kesiklilik nedeniyle log-olabilirlik (log-benzerlik) fonksiyonu ile yapılır. Olabilirlik fonksiyonları

kullanarak gözlenen ve tahmin edilen değerleri karşılaştırma işlemi Eşitlik 1.12'deki şekilde yapılmaktadır (Ünal, 1997).

$$D = -2Ln \left[\frac{\text{Şuandaki Modelin Olabilirliği}}{\text{Doymuş Modelin Olabilirliği}} \right] \quad (1.12)$$

Büyük parantezin içindeki ifadeye olabilirlik istatistiğine sapma (deviance) adı verilmektedir. Logaritma eksi iki (-2) katının alınmış olmasının matematiksel olarak bir anlamı olduğu kadar dağılımı bilinen bir değerde elde edilebilmektedir. Böyle bir teste, olabilirlik oran testi adı verilmektedir. Uyum iyiliğinde D ölçütü oldukça önemlidir. Bağımsız bir değişkenin önemine karar vermek için denklemde bağımsız değişkenin olduğu ve olmadığı durumlardaki D değerleri karşılaştırılarak karar verilir (Sanfort, 1985).

Bağımsız değişkeni içermesinden dolayı çıkan D 'deki değişim Eşitlik 1.13'teki gibidir:

$$G = D (\text{Değişkensiz Model}) - D (\text{Değişkenli Model}) \quad (1.13)$$

$$G = -2ln \left[\frac{\text{Değişkensiz Modelin Olabilirliği}}{\text{Değişkenli Modelin Olabilirliği}} \right] \quad (1.14)$$

Bütün değişkenleri içeren model ile tahmin edilen modele ilişkin olabilirlik oran değerlerinin farkına dayanan ölçütlerin ki-kare dağılımını göstereceği düşüncesinden hareketle durulan modelin geçerliliği sınanmaktadır. Bu yolla modele girecek bağımsız değişkenlere karar verilmektedir. $\beta_1=0$ hipotezi altında, G istatistiği 1 serbestlik dereceli ki-kare dağılımı gösterir (Ediz, 1997).

Olabilirlik oranı testinde modelleri karşılaştırabilmek için tüm modellerin aynı veri setinde karşılaştırılması gerekir. Bu nedenle veri setinde eksik ya da hatalı veri söz konusu olduğunda bu yöntem, mevcut veriler yerine tam ve doğru veri seti ile işlem yaparak devam edecektir. Önemlilik testinde kullanılan bir diğer yöntem ise Wald testidir. Wald testinde yine beta katsayılarının en çok olabilirlik kestiriminden yararlanılır. Bu istatistik, eğim parametresi β 'nin en çok olabilirlik kestiriminin istatistiği normal dağılıma uyar, β 'nin en çok olabilirlik kestiriminin karesi

standart hatasının karesine bölüldüğünde Wald istatistiği ki-kare dağılımına uy-
maktadır. Bu ifadeler Eşitlik 1.15 ve Eşitlik 1.16 şeklinde gösterilir.

$$t = \frac{\beta_j}{S(\beta_j)} \quad (1.15)$$

$$Wald = t^2 = \frac{\hat{\beta}_j^2}{S(\hat{\beta}_j)^2} \quad (1.16)$$

Kullanılacak bir diğer yöntem ise skor testidir (ST). Bu testin avantajı en çok olabilirlik kestirimin hesaplanmasına gerek olmamasıdır. Eşitlik 1.17'deki gibi hesaplanmaktadır. Bazı istatistik programlarında olabilirlik testi ve Wald testinin hesaplaması bulunurken skor testinin yapılmaması tercih edilmemesine neden olmaktadır.

$$ST = \frac{\sum_{i=1}^n x_i (y_i - \bar{y})}{\sqrt{y_i (1 - \bar{y}) \sum_{i=1}^n (x_i - \bar{x})^2}} \quad (1.17)$$

1.1.3.4. Lojistik Sınıflandırma ve Katsayıların Yorumlanması

Bir gözlemi birimlerin tahmin edilen olasılıkları doğrultusunda birkaç grup-
tan birine atamak sınıflama olarak tanımlanmaktadır. Genellikle $\pi(y_i = 1/x_i)$ de-
ğerini belirlemek amacıyla kullanılan lojistik regresyon modeli aynı zamanda bir
sınıflandırma modeli olarak da kullanılır. Seçilen n birimlik bir örneklem sonucu
elde edilen Eşitlik 1.18 şeklinde ifade edilir.

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + \dots + \hat{\beta}_k x_{ki} \quad (1.18)$$

Lojistik regresyon yardımıyla bulunan Eşitlik 1.19 ile olasılığın sonucu be-
lirlenmektedir. Buradaki kritik değer 0.50 olarak varsayılmaktadır. Birimlerin grup-
lara atanması belirlenen kritik değere göre yapılmaktadır. Olasılık değerinin
 $\pi_i \geq 0.5$ olması durumunda $y_i = 1$, $\pi_i < 0.5$ ise $y_i = 0$ biçiminde sınıflandırılır ve
gruplara atama işlemi gerçekleştirilir (Aktaş ve Yılmaz, 2001).

$$\pi_i = \frac{e^{y_i}}{1 + e^{y_i}} \quad (1.19)$$

Lojistik regresyon fonksiyonunda tahmin edilen regresyon katsayılarının yorumlanması ise doğrusal regresyon modelindeki kadar kolay değildir eksenindeki başlangıç noktasına göre hazırlanan lojistik regresyon modelinde x değişkenindeki bir birimlik artışın tesirini ölçmek zordur. b_j katsayısı yorumlanırken x 'de ki bir birimlik artış için $\pi_i / (1 - \pi_i)$ bahis oranı ile $EXP(b_j)$ çarpılarak elde edilen lojistik regresyon fonksiyonundan yararlanılır.

Lojistik modeldeki etkiler bahis (odds) oranına dayanır. x 'in bir değerinde kestirilen olasılık oranının, diğer değerinde kestirilen olasılık oranına oranı olarak ifade edilir. Bu istatistik $x=1$ olan bireylerin $x=0$ olan bireylere nazaran bağımlı değişkenin kaç kat daha fazla 1 olarak görüldüğü sonucunu verir (Bircan, 2004). Bu ifade ile bahis oranı 1'den büyük ise etkenin hastalık riskini arttırıcı olduğu, bahis oranı 1'den küçükse etkenin hastalık riskini azaltıcı olduğu söylenebilmektedir.

1.1.3.5. Değişkenlerin Lojistik Regresyon Analizine Dâhil Edilmesi

SPSS programıyla **Analyze>Regression>Binary Logistics** yoluyla iki-so-nuçlu (iki-gruplu veya iki-kategorili) lojistik regresyon analizi elde edilmektedir. Metrik ölçekli değişkenler modele olduğu gibi alınır. Ek herhangi bir tanımlama işlemine gerek yoktur. Nominal (sınıflayıcı) ölçekli değişkenlerin ise modele kategorik değişkenler olarak tanımlanması ve kategori değerlerinin belirtilmesi gerekmektedir. Değişkenler seçildikten sonra **Categorical** komutunun tıklanması ile açılan pencerede, kategorik değişkenin sağ tarafındaki kutuya aktarılması gerekir. Daha sonra **Change Contrast** komutu ile karşılaştırma (kontrast) tipinin belirlenmesi gerekmektedir. Burada 7 farklı karşılaştırma yöntemi bulunmaktadır (Albayrak, 2006).

1. **Deviation**; bağımsız değişkenin referans sınıfı hariç, diğer tüm sınıfları genel etkiyle karşılaştırmaktadır.
2. **Simple**; bağımsız değişkenin referans sınıfı hariç, diğer tüm sınıfları referans sınıfıyla karşılaştırmaktadır.

3. **Difference**; bağımsız değişkenin ilk sınıfı hariç tüm sınıfları, önceki sınıfların ortalama etkisiyle karşılaştırmaktadır.
4. **Helmert**; bağımsız değişkenin son sınıfı hariç diğer tüm sınıfları, sonraki sınıfların ortalama etkisiyle karşılaştırmaktadır.
5. **Polynomial**; bu yöntemde sınıflar eşit aralıklı varsayılmaktadır. Sadece nicel değişkenle için kullanılabilir.
6. **Indicator**; karşılaştırma sınıf üyeliğinin var olup olmadığını gösterir. Karşılaştırma matrisinde referans sınıfı sıfırlardan oluşan bir satır ile gösterilir

Ordinal ölçekli değişkenlerde ise araştırmanın amacına bağlı olarak modele metrik ölçekli değişkenler olarak ya da uygun bir karşılaştırma yöntemi seçilerek nominal ölçekli değişkenler olarak tanımlanabilir. Bu iki yaklaşımla elde edilen modellerden hangisi daha uygun ise ilgili yaklaşım benimsenerek analizlere devam edilebilir.

1.1.3.6. Çoklu Doğrusal Regresyon Analizi

Pratikte genellikle herhangi bir olguyu tek değişkenle açıklamak uygun olmaz. Sözelimi satış elemanı performansını etkileyen faktörler üzerine bir araştırma yapıldığında içerisinde tutumsal performans, özel performans, görev dışı performans gibi birden fazla faktörün dikkate alınması gerekmektedir.

Birçok ekonomik değişken bir araya gelerek bir değişkeni etkileyebildikleri gibi, kendi aralarında da birbirlerini etkileyebilmektedir. Sağlık alanında yapılan çalışmalarda genellikle yine benzer bir durum söz konusudur. Bir hastalığın ortaya çıkmasında birden fazla faktör etki gösterebilmektedir. Örneğin kanser vakalarında alkol-sigara, genetik yatkınlık, kronik rahatsızlık gibi birden fazla bağımsız değişkenin bir bağımlı değişkeni etkileyebilmektedir (Birdal ve Bayar 2003). Diğer bir ifade ile tek bir metrik ölçekli bağımlı değişken ile iki veya daha çok sayıda bağımsız değişken arasındaki ortalama ilişkinin matematik bir fonksiyonla ifadesine çoklu doğrusal regresyon analizi adı verilmektedir.

Çok değişkenli lojistik regresyon analizi x_1, x_2, \dots, x_p gibi p sayıda bağımsız değişkenin olduğunu ve bu değişkenlerin her birisinin en az aralık ölçekli olduğunu kabul edelim. Bağımlı değişkenin var olduğu zaman ($Y=1$) koşullu olasılık Eşitlik 1.20'deki gibi ifade edilirken,

$$\pi(x) = P(Y = 1 / x) \quad (1.20)$$

Çok değişkenli lojistik modelin lojiti Eşitlik 1.21 ve Eşitlik 1.22'deki gibi yazılmaktadır (Taşkın, 1997).

$$g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (1.21)$$

$$\pi(x) = e^{g(x)} / 1 + e^{g(x)} \quad (1.22)$$

Eğer bazı bağımsız değişkenler kesikli, nominal ölçekli (ırk, cinsiyet vb.) ise, o zaman bu değişkenleri sürekli değişkenlermiş gibi denkleme dâhil etmek uygun olmaz. Çünkü bu tür değişkenlere verilen değerlerin büyüklüklerinin nicel olarak kesin bir anlamı yoktur. Bu durumda çeşitli dizayn değişkenlerinin ya da kukla değerlerinin (dummy değişkenlerinin) kategorik olan bu değişkenleri temsil etmesi için kullanılması gereklidir.

Bir örnek verecek olursak medeni durum; evli, bekâr ve boşanmış olarak kodlanmış olsun. Kategori sayısı 3 olduğundan iki tane kukla değişken kullanılmalıdır. SPSS'te kullanılan karşılaştırma yöntemlerinden basit (simler) karşılaştırma yöntemi benimsenmesi durumunda Tablo 1.1'de verilen kukla değişkenleri elde edilmektedir.

Eğer nominal bir değişkenin k kategorisi varsa, o zaman $k-1$ tane kukla değişkeni kullanılmalıdır. J .inci bağımsız değişken X_j 'nin k tane kategorisi varsa, $k-1$ kukla değişkeni D_{ju} ve $u=1,2,\dots, k-1$ için β_{ju} ile katsayılar gösterilsin, j .nci değişkeni kesikli, p adet değişkeni olan model için lojit Eşitlik 1.23'teki gibidir (Atakurt, 1999).

$$g(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \sum_{u=1}^{k-1} \beta_{ju} D_{ju} \dots + \beta_p x_p \quad (1.23)$$

Tablo 1.1 Üç Kategorili Bir Değişkenin Kukla Değişken Olarak Kodlanması

Medeni Durum	D1	D2
Bekâr	0	0
Evli	0	1
Boşanmış	1	0

Çoklu regresyon analizi modeli kurulduktan sonra modeldeki değişkenlerin önemliliği test edilir. Çoklu lojistik regresyon analizinde de diğer çok değişkenli çözümlene yöntemlerinde olduğu gibi en az değişken ile en iyi uyuma sahip model elde edilmek istenir. Bu amaçla, veriye en iyi uyan modeli belirlemek için ileriye doğru ekleme (forward) ya da geriye doğru eleme (backward) gibi adımsal yöntemlerden yararlanır. Aynı anda tüm değişkenlerin analize dâhil edilerek bulunan tam model (full model) yerine, tek değişkenli analizler yapılarak en doğru ve en güvenli sonucu veren değişkenlerin modele dâhil edilmesi ve ilgisiz olan değişkenlerin modelden çıkartılması önerilmektedir.

1.1.3.7. Model Geliştirme Yöntemleri

Model geliştirilme aşamasında modelde tutulması gereken ya da modelden çıkarılması gereken değişkenlerin belirlenmesine çalışılır. Model geliştirme stratejisi, çalışmanın amacına bağlı olarak değişkenlik göstermektedir. Bazı araştırmalarda eldeki tüm değişkenler kullanılarak hangi değişkenin bağımlı değişkeni nasıl etkilediğinin tespiti olabileceği gibi, bazı araştırmalarda araştırmacının amacı bilinen bir ya da birden fazla değişkenin bağımlı değişkeni nasıl etkilediğinin belirlenmesi olabilir. Model geliştirilmesinde dikkat edilmesi gereken nokta, sonucu etkileyecek bu değişkenlerin model oluşturulurken dikkate alınması gerektiğidir. Yani çoklu regresyon modelleri; tam, ileri doğru seçim yöntemi, geriye doğru değişken seçim yöntemi ile adımsal değişken seçim yöntemi olarak üç farklı model geliştirme stratejisinden söz edilebilir (Ergün, 1995). Diğer bir anlatımla çoklu regresyon modelleri tam (full) veya indirgenmiş regresyon modelleri olarak ikiye ayrılabilir. Tam

regresyon yönteminde bağımlı değişkeni açıklamada etkili olabilecek tüm olası değişkenler model dâhil edilmektedir. İndirgenmiş regresyon yöntemlerini ise üçe ayırmak mümkündür. Bu yöntemler şunlardır: İleriye doğru değişken seçim yöntemi, geriye doğru değişken seçim yöntemi ve adımsal değişken seçim yöntemi. En uygun bir regresyon modelinin araştırılmasında bu üç indirgenmiş modelle model tahmini yapıp aynı veya benzer modeller elde edilmesi stratejik açıdan elde edilen regresyon modelinin uygunluğunun bir kanıtı olarak değerlendirilmektedir.

1.1.3.7.1. Tam Regresyon Yöntemi

Değişken sayısı az olduğunda, olası tüm durumların bir blok halinde dâhil edildiği modele tam regresyon modeli adı verilmektedir. Diğer bir ifade ile bağımsız değişkenlerin hiçbirinin birbiri üzerinde herhangi bir üstünlüğünün bulunmadığı düşünülerek tüm bağımsız değişkenlerin aynı anda modele alındığı durum olarak ifade edilmektedir. Değişken sayısı fazla olduğunda tanımlayıcı modellerin sayısı da fazla olacaktır. Değişkenlerin anlamlılıklarının belirlenme süresi adım sayılarının artmasıyla beraber uzayacak ve zaman kaybına sebep olacaktır. Bu durumda diğer yöntemler daha cazip hale gelmektedir. Ancak bu yöntemlerinde bazı zayıf yönleri söz konusudur. Örneğin indirgenmiş regresyon yöntemlerinden ileriye ve geriye doğru değişken seçim yöntemleri ile model geliştirildiğinde farklı modeller elde edilebilmektedir. Bu nedenle indirgenmiş regresyon yöntemlerinden adımsal regresyon yönteminin daha çok tercih edildiği söylenebilir.

1.1.3.7.2. İleri Doğru Değişken Seçim Yöntemi

İleri doğru değişken ekleme yönteminde analize sadece sabit terimin olduğu bağımsız hiçbir değişkenin olmadığı model ile başlanmaktadır. Modele eklendiğinde log olabilirlik oranı üzerinde en fazla değişime neden olan bağımsız değişken ilk aşamada modele dâhil edilmektedir. Modele anlamlı katkısı olan değişken kalmayacağı aşamaya kadar değişken seçim sürecine devam edilmektedir. Bu yöntemle elde edilen modelin en uygun model olduğu konusunda bir garanti söz konusu değildir.

1.1.3.7.3. Geriye Doğru Değişken Seçim Yöntemi

Bu yöntemin esası, ilk olarak modelin tüm bağımsız değişkenlerle birlikte kurulması (tam model) ve ardından anlamsız bağımsız değişkenlerin sırasıyla elenmesi işlemidir. Her aşamada bir bağımsız değişken modelden çıkarıldığında sapma değerindeki minimum artışa neden olan değişken modelden atılarak işleme en son sapmada en çok değişikliğe neden olan değişkenin tespit edilmesine kadar devam edilir.

1.1.3.7.4. Adımsal Değişken Seçim Yöntemi

Adımsal lojistik regresyon yönteminin kullanılması birçok değişkenin hızlı ve etkin bir şekilde incelenmesini ve değişkenlerin regresyon modellerine uyumunu sağlar. Adımsal yöntem ileri doğru seçim ile geriye doğru seçim yöntemlerinin birleştirilmiş uygulamalarını içerir. Adımsal yöntemde mantık olarak ileri doğru değişken ekleme yöntemi gibi çalışır ancak aradaki fark bir adımda modele dâhil edilen değişken bir sonraki adımda geriye doğru eleme yönteminde olduğu gibi modele katkısı dolayısıyla model dışında bırakılabilmektedir.

Adımsal yöntemin sakıncası, modelde olmayan bütün değişkenlerin katsayılarının maksimum olabilirlik tahminlerinin her adımında tek tek hesaplanması gereğidir (Kleinbaum vd, 1998).

1.1.3.8. Lojistik Regresyon Analizinde İlişki Ölçüsü

Doğrusal regresyon analizinde belirlilik katsayısı (R^2), bağımlı değişkendeki değişimin ne kadarının regresyon modelindeki bağımsız değişkenler tarafından açıklandığını gösteren bir ölçüdür. Ancak lojistik regresyon analizinde bağımlı değişken nominal ölçekli bir değişken olduğundan doğrusal regresyon analizindeki R^2 istatistiği ile lojistik regresyon analizindeki R^2 değerini doğrudan karşılaştırmak çok doğru olmamaktadır. Bununla birlikte literatürde lojistik regresyon analizinde bir kaç R^2 istatistiğine yer verilmiştir. Bunlar Mc Faden, Cox-Snell ve Nagelkerke R^2 istatistikleri en sık kullanılan istatistiklerdir. Bu istatistikler genel olarak küçük çıkma eğiliminde olduğundan, model uyumunu değerlendirmekten çok en uygun modeli seçmek amacıyla kullanılması önerilmektedir (Albayrak, 2006:439-462).

Burada R -kare benzeri istatistiği daha büyük olan modelin daha uygun bir model olduğu kabul edilmektedir.

1.1.3.9. Uyum İyiliği

Lojistik varsayım çok değişkenli dağılımların geniş bir kesimi için sağlansa da kurulan lojistik modelin geçerliliği kontrol edilmelidir. Modelde bulunması gereken bütün değişkenler ele alındıktan sonra modelin sonuç değişkenini açıklama-daki etkinliğini araştırmaya "uyum iyiliği" adı verilmektedir (Ediz, 1997).

Uyum iyiliği testleri kategorik değişkenlerin kombinasyonu ile oluşan veri matrisinde, her bir hücre için beklenen frekansların değerlerini içerir. Eğer hücrelerin herhangi birinde beklenen frekans çok küçük ise (tipik olarak $f_b < 5$) analiz sonuçlarının güvenilirliği zayıf olur. Tüm kategorik değişken çiftleri, tüm hücrelerde beklenen frekansın 1'den büyük olduğundan ve beklenen frekansın 5'ten küçük olduğu gözenek sayısının %20'yi geçmediğinden emin olmak gerekmektedir. Bu koşullardan herhangi birinin karşılanmasında sorun olduğu takdirde, araştırmacı yapacağı analizin gücünün düşük düzeyde olacağını kabul etmeli veya ikiden fazla kategorisi bulunan değişkenlerde birleştirmelere gitmeli ya da toplam hücre sayısını azaltmak için bu kategorik değişkeni dışarıda bırakmalıdır (Tabachnick ve Fidell, 1996:483-554).

Uyum iyiliğinde kullanılan istatistik analizleri çeşitlidir. Bunlardan bazıları regresyondan ayrılışın test edilmesi gibi klasik anlayışı taşır. Diğerleri ise lojistik modelde normallik varsayımı sağlanmadığı için modelin uyum iyiliği testlerinde t ve F istatistiklerinin yerine Pearson ki-kare ve D (Deviance) uyum iyiliği testini kullanırlar (Ediz, 1997). Bu yöntemlerden Pearson ki-kare testi, Hosmer-Lemeshow testi ve sapma (deviance) istatistikleri en çok tercih edilenlerdir.

Pearson ki-kare ve sapma istatistikleri gözlenen ve beklenen değerler arasındaki farkı kullanarak modelin uyumunu değerlendirmektedir. Gözlenen ve kestirilen değerler arasındaki farklar, diğer bir deyişle artıklar kullanılarak bu iki istatistik elde edilmektedir. Pearson artıklarının toplamı Pearson ki-kare istatistiğini, sapma artıklarının toplamı sapma istatistiğini verir. Her iki istatistik değeri (kritik)

tablo değerinden büyük ve $p < 0,05$ ise bu durum modelin uyumlu olmadığını gösterir. Pearson ki-kare ve sapma istatistiği Eşitlik 1.24 ve Eşitlik 1.25’de verilmiştir.

$$\text{Pearson-}\chi^2 = \sum_{i=1}^j r_i^2 \quad (1.24)$$

$$D = \sum_{i=1}^j d_i^2 \quad (1.25)$$

Hosmer-Lemeshow ki-kare uyum iyiliği testinde, lojistik regresyon modelinin bir bütün olarak uyumunu değerlendirir (Bk. Eşitlik 1.26).

$$\chi_h^2 = \sum_{i=1}^c \sum_{j=1}^r \frac{(G_{ij} - B_{ij})^2}{B_{ij}} \quad (1.26)$$

Pearson ki-kare testi, özellikle bağımsız değişkenlerin sürekli değişkenler olduğu durumda ya da küçük örneklerle çalışıldığında, geleneksel ki-kare testinden çok daha güçlüdür. Formülde G , gözlenen değerleri, B ise beklenen değerleri ifade etmektedir. Bu anlamda da geleneksel ki-kare yöntemi ile hesaplanan Omnibus testinin daha güçlü bir alternatifidir. Bu teste ilişkin sonucun anlamlı olmaması ($p > 0,05$), model-veri uyumunun yeterli düzeyde olduğunu gösterir (Garson, 2012).

1.1.3.10. Lojistik Regresyon Model Uygunluğunun Değerlendirilmesi

Model uygunluğunun değerlendirilmesi genel olarak gerçek olasılıklar ile tahmin edilen olasılıklar arasındaki standart farklara bakılarak tespit edilmektedir. Lojistik regresyon prosedürüyle hesaplanabilen diğer önemli hatalar arasında standart olmayan hatalar, standart hatalar, sapma (deviance), kaldıraç değerleri (gözlem uzaklığı) ve Cook uzaklığı sayılabilir.

1.1.3.10.1. Standart Olmayan Hatalar

Standart olmayan hatalar (e_i) fiili olasılıklar ile tahmin edilen olasılıklar arasındaki farka eşittir. Modelin hataları logit ölçekle hesaplanmış ise bu hatalara logit hatalar adı verilmektedir. Logit hatalar Eşitlik 1.30’daki şekilde hesaplanmaktadır (Albayrak, 2006:439-462).

$$\text{LogitHata}_i = \frac{e_i}{\pi_i(1-\pi_i)} \quad (1.30)$$

1.1.3.10.2. Standart Hatalar

Standart hatalar standart olmayan hataların (e_i) kendi standart sapmalarına bölünmesi ile elde edilir. Standart hatalar Eşitlik 1.31'deki şekilde elde edilir.

$$Z_i = \frac{e_i}{\sqrt{\pi_i(1-\pi_i)}} \quad (1.31)$$

Her bir birimim standart hatası χ^2 uygunluk istatistiğinin bir bileşeni olarak görülebilir. Büyük örnekler için standart hatalar 0 ortalama ve 1 standart sapma ile normal dağılıma uymaktadır (Albayrak, 2006:439-462).

1.1.3.10.3. Sapma (Deviance) Değerleri

Doğrusal regresyonda her gözlem için artıkların karesi $(y - \hat{y})^2$, o gözlemin regresyon modelinin tümel uyum eksikliği (lack of fit) olan hata kareler toplamının katsayısını gösterir. Lojistik regresyonda sapma değerleri, her gözlemin sapma istatistiği D 'ye (modelin tümel uyum eksikliğini gösteren istatistik) olan katkısını belirtir. Sapma değerleri Eşitlik 1.32'deki gibidir (Alpar, 2013:675-700).

$$d_i = \text{sign}(y_i - \hat{\pi}_i) \sqrt{-2 \left[(-y_i \ln(\hat{\pi}_i)) - (1-y_i) \ln(1-\hat{\pi}_i) \right]} \quad (1.32)$$

Standartlaştırılmış sapma değerleri ise şapka matrisi ile hesaplanır. Standartlaştırılmış sapma artıkları d_{std} Eşitlik 1.33'deki gibidir.

$$d_{std} = \frac{d_i}{\sqrt{1-h_{ii}}} \quad (1.33)$$

1.1.3.10.4. Gözlem Uzaklığı (Leverage Values)

Doğrusal regresyondaki gibi hesaplanır. h_{ii} değerleri $(p+1)/N$ 'den büyük olan gözlemler uzak gözlemler olarak adlandırılır. Bazı kaynaklarda kesim noktası

olarak, toplam gözlem sayısı 30'un altında ise $3p/n$, gözlem sayısı yeterince büyük ise $2p/n$ değerlerinin kullanılması önerilmiştir.

1.1.3.10.5. Cook Uzaklığı

Etkili gözlemleri belirlemek için kullanılan istatistiklerden birisi de Cook uzaklığı D_j 'dir. Cook uzaklığı D_j Eşitlik 1.34'deki gibi hesaplanır.

$$D_j = \frac{z_j^2 h_j}{1 - h_j} \quad (1.34)$$

Cook uzaklığı D_j 'nin değeri gözlem uzaklığına ve onun standartlaştırılmış artışının büyüklüğüne bağlıdır. Cook uzaklığı bir gözlemin silinmesinin tüm gözlemler için değerleri ne kadar etkilediğini ölçer. Genellikle 1'in üzerindeki Cook değerleri etkili gözlem olarak değerlendirilmiştir (Alpar, 2013:675-690).

İKİNCİ BÖLÜM

2.1. LOJİSTİK REGRESYON ANALİZİ KULLANILARAK YAPILAN ÇALIŞMALAR:LİTERATÜR İNCELEMESİ

Literatür taraması için sağlık verileri, regresyon analizi, lojistik regresyon analizi anahtar kelimeleri kullanılarak tarama yapılmıştır. 1999-2016 yılları arasında yayınlanmış, lojistik regresyon analizi kullanılan 18çalışma incelenmiştir.

Bu çalışmalar yılı, örneklem hacmi ve özellikleri, yöntemi, bilgi kaynağı, kullanma nedeni algılanan etkinlik ve çalışmadan elde edilen önemli sonuçlar yönünden değerlendirilmiştir.

Atakurt (1999) “Lojistik regresyon analizi ve tıp alanında kullanımına ilişkin bir uygulama” isimli çalışmasında lojistik regresyon analizinin esasları ve ana basamaklarından bahsetmiştir. 500 bireyden elde edilen 34 değişken kullanılan bu çalışmada bağımlı değişken olarak belirlenen koroner arter hastalığı (*KOR*) değişkeni ile diğer bağımsız değişkenler arasındaki ilişkiyi incelenmiştir. Regresyon analizi yöntemi belirlenirken maksimum olabilirlik yöntemi kullanılmıştır.

Eskin, 2000 yılında yayınlamış olduğu “Lise öğrencisi ergenlerde depresyonun yaygınlığı ve ilişkili olduğu etmenler” isimli çalışmasında çocukluk depresyon envanteri değişkeninde belirlenen kesme noktası üzerinde değere sahip olan verileri etkileyen bağımsız değişkenler arasındaki durumu incelemiştir. Bu sonucun oluşmasında ne gibi etkilerinin olduğunu araştırmıştır. Depresyon olup olmama durumu bağımlı değişken olarak belirlenmiş ve diğer etkenlerin bağımlı değişken üzerindeki etkisi lojistik regresyon analizi ile incelenmiştir.

Maral vd. 2001 yılında yapmış olduğu “Depresyon yaygınlığı ve risk etkenleri: huzurevinde ve evde yaşayan yaşlılarda karşılaştırmalı bir çalışma” isimli çalışmasında huzurevi ve evde yaşama durumuna göre depresyon yaygınlığındaki artış incelemiştir. İki kategorili değişken evde ve huzurevinde yaşayanlar üzerinde etkili olan diğer bağımsız değişkenlerin tespiti için lojistik regresyon analizi kullanılmıştır.

Ersan vd. 2003 yılında yayınlanan “Mide kanserli hastalarda klinik patolojik prognostik faktörler” isimli çalışmalarında tek değişkenli iki kategorili değişkenler arasındaki ilişki ve çok değişkenli analiz için çok değişkenli lojistik regresyon analizi kullanılarak çalışma gerçekleştirilmiştir. Değerlendirme sonucunda evre, perinatal invazyon, metastatik nodul seviyesi, metastatik nodul sayısı ve splenektomi gibi değişkenlerin hastalık için anlamlı olduğu sonucuna lojistik regresyon analizi kullanılarak ulaşılmıştır.

Bircan, 2004 yılında yayınlamış olduğu “Lojistik regresyon analizi: tıp verileri üzerine bir uygulama” isimli çalışmasında ikili sonuç değişkeni ile hem sürekli hem de kesikli değişkenlerden oluşan bağımsız değişkenler kümesi arasındaki ilişkiyi tanımlayabilen lojistik regresyon analizinin incelenmeye çalışılmıştır. Lojistik regresyon analizine bir uygulama göstermek amacıyla çocuklarda doğum ağırlığını etkileyen önemli risk faktörlerini belirlemek için tıp üzerine bir çalışma yapılmıştır. Lojistik modele dâhil edilecek bağımsız değişkenler, tek değişkenli lojistik regresyon analiziyle belirlendikten sonra, çok değişkenli modele dâhil edilen her bir değişkenin önemliliği gösterilmiştir. Tek değişkenli modelde önemli bulunduğu halde çok değişkenli modelde önemsiz olan değişkenler model dışı bırakılmıştır. Böylece nihai model elde edilmiştir. Nihai modelin hem biyolojik olarak kabul edilebilir, hem de doğru sınıflama oranının yeteri kadar iyi olmasından dolayı, bu modelin risk faktörlerini belirleme de kullanılabileceği sonucuna varılmıştır.

Aktaş ve Giray 2005 yılında “Mali başarısızlığın çok değişkenli istatistiksel yöntemlerle belirlenmesi: tekstil sektörü örneği” isimli çalışmalarında mali başarısızlık riski olan işletmeleri tespit edebilecek bir model bulmaya çalışmışlardır. Tekstil sektöründe faaliyet gösteren İMKB’ye kayıtlı işletmelere ait veriler araştırma çerçevesini oluşturmaktadır. Çalışmada öncelikle bu işletmelerin 2001 yılına ait bilançoları yardımıyla mali oranları hesaplanmış, dönem sonu kar-zarar durumuna bakılarak, işletmeler o dönem için mali başarısız ya da başarılı olarak değerlendirilmiştir. Krizlere karşı dirençsiz olan tekstil sektörü işletme riski modelinin oluşturulması ve modelde anlamlı olan değişkenlerin belirlenmesi için öncelikle

elde edilen mali oranlara faktör analizi uygulanmış, elde edilen faktör skorları bağımsız değişken olarak alınarak uygulanan lojistik regresyon analizi sonuçlarına göre mali başarısızlığı etkileyen faktörler belirlenmiştir. Araştırma sonunda elde edilen bulgular yorumlanmıştır.

Çok vd. (2006) “İleri evre küçük hücreli dışı akciğer karsinomunda sisplatin-etoposid ile mitomisin-ifosfamid-sisplatin kombinasyonlarının karşılaştırılması” isimli çalışmalarında Sisplatin-etoposid (CE) ile mitomisin-ifosfamid-sisplatin (MIC) rejimleri küçük hücreli dışı akciğer karsinomunun (KHDAK) tedavisinde kullanılan konvansiyonel aktif kombinasyonlardır. Bu çalışmada önceden tedavi almamış evre IIIB ve IV KHDAK hastalarında CE ve MIC tedavilerinde gözlenen yanıt oranlarını, sağkalım, yanıt ve progresyon süreleri ve toksisite durumlarını karşılaştırmak amacıyla yapılan çalışmada yanıt oranı, yaşam ve yanıt süresi oranları açısından iki grup arasında anlamlı bir fark var mı yok mu lojistik regresyon analizi de kullanılarak test edilmiş ve fark olmadığı sonucuna varılmıştır.

Kaya vd. (2007) “Tıp fakültesi ve sağlık yüksekokulu öğrencilerinde depresif belirti yaygınlığı, stresle başa çıkma tarzları ve etkileyen faktörler” isimli çalışma kullanılan Beck Depresyon Ölçeği (BDÖ), Stresle Başa Çıkma Tarzları Ölçeği (SBTÖ) ve Sosyodemografik Bilgi Formu gibi ölçeklerden elde edilen bağımsız değişkenlerin bağımlı değişken olan depresif belirtiler ile ilişkili olup olmadığını araştırılmış ve ilişkili olduğu sonucuna ulaşılmıştır.

Pelitli vd. (2007) “Açıkgöz küresi yaralanmalarında sonuç görme keskinliği üzerine etkisi olan faktörler” isimli çalışmalarında; acil cerrahi onarım uygulanan 119 olgunun 119 gözü çalışmaya alınmıştır. Olgular yaş, cinsiyet, travmanın tipi, lokalizasyonu, ameliyat öncesi ve sonrası muayene bulguları ile başka cerrahi girişim yapılmış olması ve bunlar arasında ilişki lojistik regresyon analizi açısından araştırılmıştır.

Canbaz vd. (2007) “Tıp fakültesi son sınıf öğrencilerinin kaygı düzeyi ve etkileyen faktörler” isimli çalışmalarında tıp fakültesi son sınıf öğrencilerinin kaygı düzeylerini etkileyen faktörler uygulanan ölçeklerden elde edilen değişkenlerin bir-biri ile ilişkileri lojistik regresyon analizi ile incelenmiştir.

Düzenli vd. (2008) “Kronik solunum yetmezliği olan olgularda uzun süreli oksijen tedavisinin yaşam süresi üzerine etkisi” isimli çalışmalarında 15 saat kullanılan uzun süreli oksijen tedavisinin kronik solunum yetmezliği olan hastalarda yaşam süresini ve kalitesini arttığı bilgisinden yararlanarak lojistik regresyon analizi ile tedavi süresi arttırıldığında yaşam süresini arttığı yönünde bir sonuç elde edilmiştir. Parametreler arasındaki ilişkide, değişik faktörlerin bağımlılık derecesi için lojistik regresyon analizi kullanılmıştır.

Girginer ve Cankuş (2008) “Tramvay yolcu memnuniyetinin lojistik regresyon analiziyle ölçülmesi: ESTRAM örneği” isimli çalışmalarında toplu taşıma araçlarından biri olan tramvaya yönelik yolcu memnuniyeti, Eskişehir tramvay sistemi (ESTRAM) örneğinde, binominal lojistik regresyon analizi ile incelenmektedir. İki üniversiteye sahip olan Eskişehir’de öğrenci nüfusun fazla olması ve tramvay için önemli bir yolcu kitlesi olacakları düşüncesiyle çalışma, her iki üniversiteden basit tesadüfi örnekleme yoluyla seçilen 300 öğrenci üzerinde gerçekleştirilmiştir. Öğrencilerin memnuniyetleri ile ilgili binominal düzeyde, gizil bir değişken kullanılmıştır. Uygulanan binominal lojistik regresyon analizi sonucunda; öğrencilerin ESTRAM’dan memnuniyetleri üzerinde modele alınan tüm bağımsız değişkenlerin negatif etkileri olduğu belirlenmiştir

Özkan vd. (2009) “Bebek ölüm nedenleri üzerine bir araştırma” isimli çalışmalarında bebek ölümleri üzerine etkili olan faktörler araştırılmaya çalışılmışlardır. Annenin eğitim durumu, gebelik yaşı, bebeğin doğduğu yer ve doğumda bir sağlık personeli yardımı olması ya da olmaması, bebeğin sağlık kontrolünden geçmesi, bebeğin cinsiyeti, emzirilme durumu, annenin doğum sayısı ve bebeğin doğum zamanı, kilosu bağımsız değişkenler olarak alınıp, bebeğin ölümüne etkisi için iki kategorili bağımlı değişken üzerinde lojistik regresyon analizi uygulanmıştır.

Sulock (2009) “An Application of Binary Logistic Regression to College Admissions Data” konulu İkili Lojistik Regresyonun Üniversite Kabul Verilerine Uygulanması” tezinde üniversite kabul verilerini belirlemede bir genelleme yapılarak doğrusal regresyon analizi ile binary lojistik regresyon analizi incelenmiş normal dağılıma nasıl tepki verdiklerine bakılmıştır

Ege ve Bayrakdarođlu (2009) “İMKB Őirketlerinin hisse senedi getiri baŐarılarının lojistik regresyon tekniđi ile analizi” isimli alıŐmalarında İMKB 30 hisse senetlerinin getiri performansı lojistik regresyon tekniđiyle analizi yapılarak, İMKB’de faaliyet gsteren Őirketlerin baŐarı durumu belirlenmiŐtir. AraŐtırmada, İMKB’de iŐlem gren otuz Őirketin, 2004 yılına ait yirmi finansal oranı ve nominal TL getirileri kullanılmıŐtır. Tahmin edilecek model iin oranlar tek tek eklenmiŐ ve istatistiksel olarak anlamlı modeller tespit edilmiŐtir.

AktaŐ ve ErkuŐ (2009) “Lojistik regresyon analizi ile EskiŐehir’in sis kestiriminin incelenmesi” isimli alıŐmalarında, bađımlı deđiŐkenin iki dzeyli olması durumunda demografik, davranıŐ ve risk faktryle ilgili tahmin alıŐmalarında olduka sık kullanılan lojistik regresyon analizi teorik olarak kısaca incelenmiŐtir. Sınıflandırma analizlerinin en ok kullanılanlarından biri olan ve birok konuda uygulama alanı bulunan lojistik regresyonun parametre tahminine iliŐkin kestirim yntemi detaylı bir Őekilde incelenmiŐ, katsayıların yorumu iin bahis oranı anlatılmıŐtır. Daha sonra EskiŐehir’de sis olayı ve havacılık faaliyetleri zerine etkileri incelenmiŐtir. AraŐtırmanın uygulama blmnde ise, EskiŐehir’de sis meydana gelmesi olasılıđını verecek en uygun denklem belirlenmiŐ ve lojistik regresyon analizi kullanılarak alıŐma somut hale getirilmiŐtir.

Meulen vd. (2011) “Case Study: An Application of Logistic Regression in a Six Sigma Project in Health Care” isimli “Sađlık Hizmetlerinde Altı Sigma Projesinde Bir Lojistik Regresyon Uygulaması” alıŐmalarında kardiyoloji servisinde yatana hastaların bir rehabilitasyon programına katılma olasılıđını hastanın evinin uzaklıđı, hastanın arabasının olup olmaması gibi deđiŐkenler ile lojistik regresyon analizi kullanılarak bulunmaya alıŐmıŐlardır.

Kutlu vd. (2011) “İlkretim 5. sınıf đrencilerinin okuduđunu anlamada baŐarılı olup-olmama durumlarının kestirilmesinde etkili olan deđiŐkenlerin incelenmesi” isimli alıŐmalarında okuduđunu anlama baŐarısını sınıflamada (baŐarılı-baŐarısız) etkili olan deđiŐkenlerin modeldeki nem dzeylerini ortaya koymaya alıŐmıŐlardır. 10 ilkretim okulunun 5. sınıfında okuyan 279 đrenci zerinde yrtlmŐtr. Bu đrencilerden elde edilen veriler, lojistik regresyon analizi kul-

lanılarak incelenmiştir. Analiz sonucunda öğrencilerin okuduğunu anlamada başarılı olup olmama durumlarının kestirilmesi olasılığını; cinsiyet, öğrenciye ait kitap sayısı, evde bulunan kitap sayısı, öğretmenin okunan metinlerle ilgili özet yazdırması ve öğretmenin okumayla ilgili ev ödevi vermesi değişkenlerinin anlamlı düzeyde etkili olduğu bulunmuştur.



ÜÇÜNCÜ BÖLÜM

3.1. LOJİSTİK REGRESYON ANALİZİ: SAĞLIK VERİLERİ ÜZERİNE BİR UYGULAMA

19 Mayıs Üniversitesi Tıp Fakültesi Kadın Hastalıkları ve Doğum Kliniğinde 2005-2013 yılları arasında rahim kanseri teşhisiyle operasyon geçirmiş 177 hasta, kontrol grubu olarak kanser teşhisi konmamış ancak başka nedenlerle operasyon geçirmiş 100 hasta çalışmaya dâhil edilmiştir. Hastalar geriye dönük olarak değerlendirilmiştir. Tüm olguların yaşı, histolojik dereceleri, histopatolojik tipleri, ortalama tümör boyutu, Miyometrial invazyonu, adneksiyel tutulumu, peritoneal sitoloji, lenfovasküler tutulumu, pelvik-paraaortik lenf nodu tutulumu, cerrahi evresi değerlendirilmeye alınmıştır.

Lojistik regresyon analizi seçilirken basit lojistik regresyon analizi kullanılmıştır. Yani iki kategorili bağımlı bir değişken kanser hastası olan ya da olmayan grubu diğer bağımsız değişkenlerin nasıl ve ne şekilde etkilediği araştırılmaya çalışılmıştır. Değişken ekleme veya çıkarma yöntemleri kullanılarak arasındaki farklar ve bu farkların sonuç değişkenini nasıl etkilediği araştırılmaya çalışılmıştır. Hangi değişkenin, tek başına ya da birlikte ne tür sonuç verdiği test edilmiştir.

3.1.1. Lojistik Regresyon Analizi Öncesi Yapılan Çalışmalar ve Kullanılan Değişkenler

Verilerin analizi SPSS programıyla yapılmıştır. Kontrol grubu, erken evre ve ileri evre rahim kanseri arasındaki ilişkideki verilerin karşılaştırılması işleminde lojistik regresyon analiz yöntemi kullanıldı. İstatistiksel anlamlılık düzeyi %5 olarak alınmıştır. Çalışmamızda rahim kanserli hasta ve kontrol grubu arasındaki ilişkiyi incelemek için kullandığımız değişkenler Tablo 3.1'deki gibidir:

Bu değişkenlerden yaş sürekli değişkeni hariç diğer değişkenler kategorik değişkenlere çevrilerek hesaplamalara dâhil edilmiştir.

Tablo 3.1'de yer alan bağımsız değişkenler rahim kanserine neden olan faktörleri tespit etmek için kullanılacak değişkenlerdir.

Tablo 3.1 Araştırmada Kullanılan Değişkenlerin Simgeleri ve Açıklamaları

1-Metrik Ölçekli Olmayan (Nominal veya Ordinal) Ölçekli Değişkenler		
Simge	Değişkenin Açıklaması	Değişkene İlişkin Kodlamalar
<i>OTB</i>	Ortalama Tümör Boyutu	1=<1; 2=1-2; 3=2-4; 4=>4
<i>MİN</i>	Miyometrial İnvazyon	1=Küçük;2=Büyük
<i>HST</i>	Doku türü (histoloji)	1=Endometrium tip; 2=Endometrium olmayan
<i>SST</i>	Servikal Stromal Tutulum	1=Negatif; 2=Pozitif
<i>ADT</i>	Andeksiyal Tutulum	1=Negatif; 2=Pozitif
<i>LVİ</i>	Lenfovasküler İnvazyon	1=Negatif; 2=Pozitif
<i>PELT</i>	Pelvik Lenf Nodu Tutulumu	1=Negatif; 2=Pozitif
<i>PALT</i>	Paraaortik Lenf Nodu Tutulumu	1=Negatif; 2=Pozitif
<i>EVRE</i>	Kanser Hastalığının Evresi	1=Evre-1 (1A1B); 2=Evre-2; 3=Evre-3 (A3C13C2); 4=Evre-4B
<i>TRM3</i>	Operasyon Öncesi Yeniden Kodlanmış Trombosit Değerleri	1=<200; 2=201-400; 3=401-600; 4=>601
<i>TRM4</i>	Operasyon Sonrası Yeniden Kodlanmış Trombosit Değerleri	1=<100; 2=101-200; 3=201-300; 4=>300
<i>WBC3</i>	Operasyon Öncesi Yeniden Kodlanmış Lökosit Değerleri	1=0-2000; 2=2001-4000; 3=4001-6000; 4=6001-8000; 5=>8001
<i>WBC4</i>	Yeniden Kodlanmış Operasyon Sonrası Lökosit Değerleri	1=0-2000; 2= 2001-4000; 3=4001-6000; 4=6001-9000; 5=9001-12000; 6=>12000
<i>WBC5</i>	Danışman Hekim Tarafından Risk Değerine Göre Yeniden Kodlanan Operasyon Öncesi Lökosit Değerleri	0=<8000;1=>=8000
2-Metrik Ölçekli Değişkenler		
Simge	Değişkenin Açıklaması	
<i>YAŞ</i>	Hastanın Yaşı	
<i>TRM1</i>	Operasyon Öncesi Trombosit Değerleri	
<i>TRM2</i>	Operasyon Sonrası Trombosit Değerleri	
<i>WBC1</i>	Operasyon Öncesi Lökosit Değerleri	
<i>WBC2</i>	Operasyon Sonrası Lökosit Değerleri	

3.1.2. Araştırmada Kullanılan Önemli Değişkenlerin Tanımsal İstatistikleri

Çalışmaya 177 rahim kanseri tanısı alan hastalar dâhil edildi. Kanser ve kontrol grubu hastaların yaş dağılımı Tablo 3.2’de verilmiştir. Bağımsız iki-örneklem *t*-testi ile hasta ve kontrol grubunun yaş ortalamaları arasında anlamlı bir fark olup olmadığına bakılmıştır. Tablo 3.2 incelendiğinden hasta ve kontrol grubunun yaş ortalamaları arasındaki farkın %5 anlamlılık düzeyinde önemli olduğu sonucuna varılmıştır (*p*-değeri=0,006). Diğer bir ifadeyle denilebilir ki kanser hastalarının yaş ortalamasının kontrol grubu hastaların yaş ortalamasından daha büyüktür. Bu sonuç hasta yaşının kanser hatalığına yakalanma riskini açıklamada önemli bir değişken olabileceğini göstermektedir.

Tablo 3.2 Hasta ve Kontrol Grubunun Yaş Ortalaması ve Standart Sapma Değerleri

Grup	Sayı	Ortalama	St. Sapma	<i>p</i> -değeri
Hasta	177	59,6	10,8	0,006
Kontrol	100	56,9	5,3	
Total	277	58,6	9,3	

Hasta ve kontrol grubunun operasyon öncesi ve sonrası lökosit, trombosit değerleri ve yaş değişkeninin aritmetik ortalaması, standart sapması, en küçük ve en büyük birim değerleri Tablo 3.3’te verilmiştir. Kontrol grubunda operasyon öncesi ortalama lökosit sayısı 7039K/mm³ iken, ortalama trombosit sayısı 296K/mm³ olarak bulunmuştur(Bk. Tablo3.3). Hasta grubunda ise operasyon öncesi ortalama lökosit değeri 8100K/mm³ iken, trombosit değeri ise 302,6K/mm³ olarak tespit edilmiştir.

Hasta ve kontrol grubu lökosit, trombosit ve yaş değişkeni ortalamaları arasında anlamlı bir fark olup olmadığı bağımsız iki-örneklem *t*-testi ile araştırılmıştır (Bk. Tablo3.4).%95 güven düzeyinde yapılan T testi sonucuna göre operasyon öncesi ve sonrası trombosit değerleri ortalamalar açısından hasta ve kontrol gruplarında anlamlı bir farklılık bulunmamıştır. Yaş, operasyon öncesi lökosit ve operasyon sonrası lökosit değişkenlerine ilişkin *p*-değerleri %5’ten olduğundan sıfır hipotezi reddedilir. Diğer bir ifade ile bu değişkenlere ilişkin araştırma hipotezleri (*H*₁) kabul edilir. Yani yaş ve operasyon öncesi ve sonrası lökosit değerleri hasta

ve kontrol grubunu açıklamada anlamlı bir değişken olarak kullanılabilir görüşü desteklenmektedir.

Tablo 3.3 Hasta ve Kontrol Grubu Operasyon Öncesi ve Sonrası Lökosit ve Trombosit ile Yaş Değişkenlerinin Tanımsal İstatistikleri

		Sayı	AO	SS	SH	AO İçin %95 GA		ENK	ENB
						ASD	ÜSD		
Hastanın Yaşı	Kontrol	100	56,9	5,3	0,5	55,9	58,0	50	72
	Hasta	177	59,6	10,8	0,8	58,0	61,2	30	90
	Toplam	277	58,6	9,3	0,6	57,5	59,7	30	90
Operasyon Öncesi Trombosit Değerleri	Kontrol	100	296,0	75,6	7,6	281,0	311,0	154	543
	Hasta	177	302,6	91,8	6,9	289,0	316,2	163	742
	Toplam	277	300,2	86,2	5,2	290,0	310,4	154	742
Operasyon Öncesi Lökosit Değerleri	Kontrol	100	7039,7	1474,7	147,5	6747,1	7332,3	4050	10690
	Hasta	177	8100,4	2668,7	200,6	7704,5	8496,3	3620	16100
	Toplam	277	7717,5	2362,7	142,0	7438,0	7996,9	3620	16100
Operasyon Sonrası Lökosit Değerleri	Kontrol	100	7729,8	2221,1	222,1	7289,1	8170,5	127	12020
	Hasta	177	9297,1	2966,2	223,0	8857,1	9737,1	280	19190
	Toplam	277	8731,3	2819,4	169,4	8397,8	9064,8	127	19190
Operasyon Sonrası Trombosit Değerleri	Kontrol	100	312,0	710,5	71,1	171,0	452,9	121	7320
	Hasta	177	239,0	82,9	6,2	226,7	251,3	108	800
	Toplam	277	265,4	432,1	26,0	214,2	316,5	108	7320

Tablo 3.3 incelendiğinde örneklem içindeki hastaların genel yaş ortalaması 58,6 iken tüm hastaların %95 güven aralığında yaşları 57,5 ile 59,7 arasında değiştiği görülmektedir. Örneklem grubunda yaşı en küçük olan hastanın yaşı 30, en büyük olanın 90 olduğu anlaşılmaktadır. Diğer yandan hasta grubunun (177 hasta) yaş ortalaması 59,6 yıl iken tüm hastaların yaşlarının %95 olasılıkla 58 yıl ile 61 yıl arasında değiştiği söylenebilir. Kontrol grubunun yaş ortalaması yaklaşık olarak 56,9 yıl iken tüm kontrol grubunun ana kütle yaş ortalaması veya tüm birimlerin yaşları %95 olasılıkla 55,9 yıl ile 58 yıl arasında değiştiği söylenebilir.

Operasyon öncesi trombosit değerlerinin genel ortalaması 300,2 iken bu oran hasta grubunda 302,6, kontrol grubunda ise 296,0 olarak tespit edilmiştir. Operasyon öncesi lökosit değeri aritmetik ortalaması hasta grubunda 8100,4 olarak bulunurken kontrol grubunda 7039,7 olarak tespit edilmiştir. %95 olasılıkla en küçük operasyon öncesi lökosit değeri kontrol grubunda 6747,1 iken bu oran hasta grubunda 7704,5 olarak bulunmuştur.

Diğer değişkenlere ait aritmetik ortalamalar, standart sapmalar, aritmetik ortalamaların standart hataları ve % 95 güven düzeyleri Tablo 3.3'te sunulmuştur.

Tablo 3.4 Yaş, Lökosit ve Trombosit Değerleri İçin Bağımsız İki-Örnek T-Testi

Değişken	Levene Sabit Varyans Testi		İki-Örneklem Ortalamaları İçin t-Testi						
	F	p-değeri	t	sd	p-değeri	OF	SH	OF için 95% GA	
								En Küçük	En Büyük
Hastanın Yaşı	39,0	0,000*	2,78	270,5	0,006	2,7	1,0	0,8	4,6
Operasyon Öncesi Trombosit Değerleri	1,0	0,316	,62	275	0,538	6,7	10,8	-14,6	27,9
Operasyon Sonrası Trombosit Değerleri	3,2	0,077	-1,35	275	0,178	-72,9	54,0	-179,2	33,3
Operasyona Öncesi Lökosit Değerleri	27,6	0,000*	4,26	274,9	0,000	1060,7	249,0	570,6	1550,8
Operasyon Sonrası Lökosit Değerleri	7,6	0,006*	4,98	254,0	0,000	1567,3	314,7	947,5	2187,1

*: Sabit varyans varsayımının sağlanmadığı duruma ilişkin t-testi sonucunu göstermektedir.

Endometrium kanserli hastaların 37'sinde (%21) lökositoz, 11'inde (%6,2) trombositoz tespit edilmiştir. Lökosit sayım değerlerinin 10.000 K/mm³e eşit veya büyük olması lökositoz, trombosit sayısının 450 K/mm³e eşit veya büyük olması trombositoz olarak kabul edilmiştir. Lökositoz tespit edilen hasta grubunun ortalama lökosit değeri 12.187,3 K/mm³olarak hesaplanmıştır. Bu istatistikler Tablo 3.5'te verilmiştir.

Tablo 3.5 Hasta Grubu Lökositoz ve Trombositoz Düzeyleri Frekans Tablosu

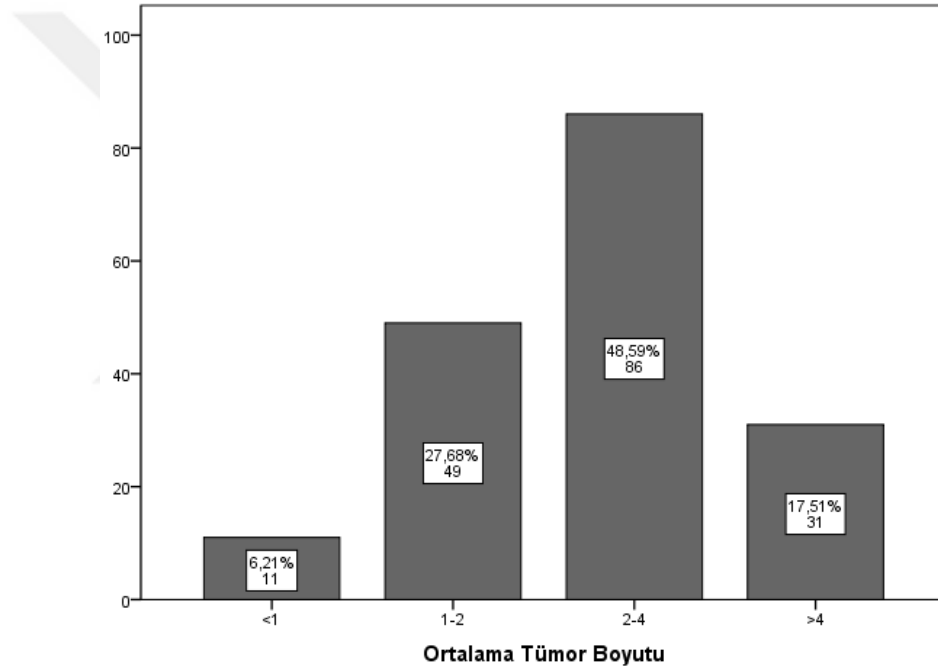
Lökositoz Düzeyi	Frekans	Yüzde	Trombositoz Düzeyi	Frekans	Yüzde
Düşük (<10.000)	140	79,1	Düşük (<450)	166	93,8
Yüksek (>=10.000)	37	20,9	Yüksek (>=450)	11	6,2
Toplam	177	100	Toplam	177	100

Hasta grubu operasyon öncesi lökosit değerleri ile lökositoz tespit edilen vakalarda ortalama lökosit değeri 12.187,3, ortancası 11.770, tepe değeri 10.000 ve standart sapması 1.848,1 olarak tespit edilmiştir. Bu sonuçlar Tablo 3.6'da sunulmuştur.

Tablo 3.6 Hasta Grubu Operasyon Öncesi Lökositöz Değişkeninin Aritmetik Ortalaması, Ortanca, Tepe Değeri ve Standart Sapması

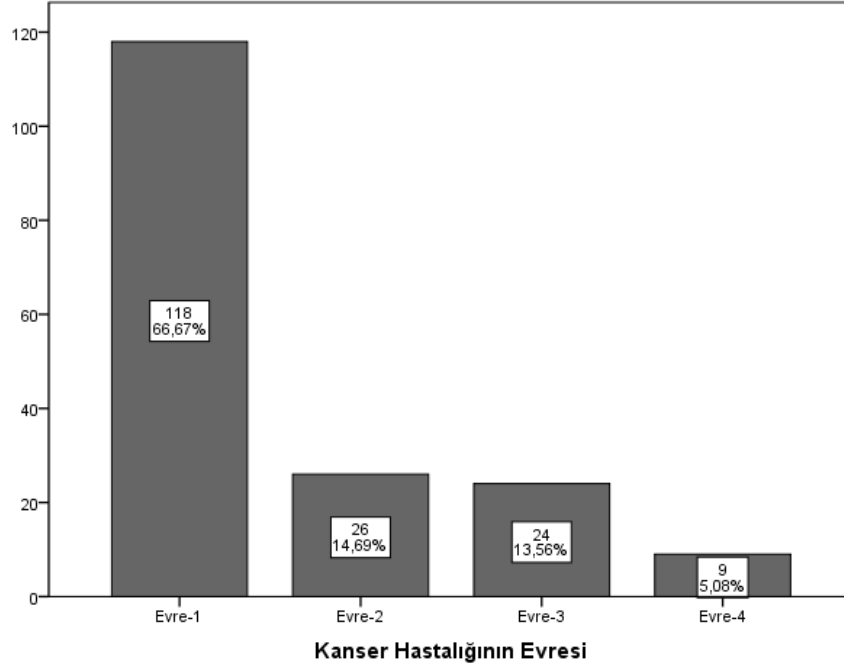
Ortalama	12.187,3
Ortanca	11.770,0
Tepe Değeri	10.000,0
Standart Sapma	1848,1

Çalışmaya alınan hastaların 11'inde (%6,21) ortalama tümör boyutu 1'den küçük, 49'unda (%26,68) ortalama tümör boyutu 1 ile 2 arasında, 86'sında (%48,59) ortalama tümör boyutu 2 ile 4 arasında, 31'inde (%17,51) 4'ten büyük ortalama tümör boyutu tespit edilmiştir. Bu sonuçlar Şekil 3.1'de verilmiştir.



Şekil 3.1 Hasta Olgularında Ortalama Tümör Boyutları

Hastaların cerrahi evrelere göre dağılımları; Şekil 3.2 incelendiğinde 118 hasta evre 1, 26 hasta evre 2, 24 hasta evre 3, 9 hasta evre 4 şeklinde bir dağılım göstermiştir. Bu sonuçlar araştırma kapsamındaki kanser hastalarının önemli kısmının başlangıç aşamasında (Evre-1) olduğunu göstermiştir.



Şekil 3.2 Çalışmaya Alınan Rahim Kanseri Hastalarının Cerrahi Evreleri

Tablo 3.7 Lökosit ve Evre, Paraaortik Lenf, pelvik Lenf Nodu Tutulumu, Andeksiyel Tutulum, Histoloji Arasında Ki-Kare Analizi Sonuçları

İstatistik	Lökosit Evre	Lökosit Paraaortik	Lökosit Pelvik	Lökosit Andeksiyel	Lökosit Histoloji
Ki-Kare	39,7	10,8	11,2	12,2	0,8
<i>p</i> -değeri	0,000	0,001	0,001	0	0,358

Çalışmaya alınan hastalardan miyometrial invazyonu %50'den küçük olan 149 (%84,2) ve % 50'ye eşit ve büyük olan 28 (%15,8) olgu tespit edilmiştir (Bk. Ek-14). Hastaların 159'unda (%89,8)adneksiye tutulum izlenmezken, sadece 18 (%10,2) hastada adneksiye tutulum gözlenmiştir (Bk. Ek-14). Hastalardan alınan peritoneal sitoloji sonuçları; 54 (%30,5) hastada malignite negatif, 27 (%15,3) hastada örnek değerlendirme için yetersiz (nondiagnostik) iken 83 (%46,9) hastada malignite pozitif izlenmiştir (Bk. Ek-14). Pelvik lenf nodu tutulumu, paraaortik lenf nodu tutulumu, histoloji, servikal stromal tutulum grafikleri ekler bölümünde verilmiştir. Elde edilen sonuçlar daha sonra ki-kare analizi ile tekrar değerlendirildiğinde *p* değeri evre, paraaortik lenf nodu tutulumu, pelvik lenf nodu tutulumu ve adneksiye tutulum için ayırt edici derecede farklı bulunmuş, histoloji ve hastalığın derecesi değişkenleri için ise farklı bulunmamıştır (Bk. Tablo 3.7).

3.1.3. SPSS ile Lojistik Regresyon Analizinin Elde Edilmesi

Lojistik modelle ilgili olarak bağımlı değişkenimiz Grup adında iki kategoride 1: Hasta, 0: Hasta olmayan vakalar olarak kodlanmıştır. Bağımsız değişken operasyon öncesi lökosit değeridir. Lökosit değeri 10000 K/mm^3 'den büyük olanlar 1, diğerleri 0 olarak yeniden kodlanarak lökositöz değişkeni adı verilmiştir. Sayısal olan operasyon öncesi trombosit ve lökosit değerleri yeniden kodlanarak kategorik hale getirilmiş ve bu şekilde analize dâhil edilmiştir. Tablo 3.8'de analizde kullanılan kategorik değişkenler ve frekans değerleri yer almaktadır. Lojistik regresyon analizine ilk olarak hiçbir değişken dâhil edilmeden hasta ve kontrol grubundan hangilerinin elimizdeki veriler ile doğru tespit edilebilme başarı oranlarının yer aldığı adım ile başlanır. İleriye doğru değişken ekleme metodu kullanılarak gerçekleştirilen lojistik regresyon analizi başlangıç adımında yer alan sınıflandırma tablosundaki veriler Tablo 3.9'da verilmiştir. Bu veriler ile hiçbir değişken dâhil edilmeden hasta olan grubun doğru tespit edilme durumunu genel yüzde de %64,1 olarak belirlemektedir. Bundan sonraki adımlarda değişkenler analize dâhil edilerek bu oranın nasıl etkilendiğine bakılmakta ve en uygun değişkenler analize dâhil edilerek işleme devam edilmektedir.

Tablo 3.8 Analizde Kullanılan Kategorik Değişkenlerin Kategorik Kodlamaları ve Frekansları

Kategorik Değişkenler	Sınıf Aralığı	Frekans	Parametre Kodu				
			(1)	(2)	(3)	(4)	(5)
WBC4 Kategorik Olarak Yeniden Kodlanmış Operasyon Sonrası Lökosit Değerleri	1=0-2000	6	1	0	0	0	0
	2=2001-4000	2	0	1	0	0	0
	3=4001-6000	28	0	0	1	0	0
	4=6001-9000	129	0	0	0	1	0
	5=9001-12000	82	0	0	0	0	1
6=12000<	29	0	0	0	0	0	
WBC3 Kategorik Olarak Yeniden Kodlanmış Operasyon Öncesi Lökosit Değerleri	2=2001-4000	3	1	0	0		
	3=4001-6000	61	0	1	0		
	4=6001-8000	108	0	0	1		
	5=>8001	104	0	0	0		
TRM3 Operasyon Öncesi Yeniden Kodlanmış Trombosit Değerleri	1=<200	26	1	0	0		
	2=201-400	218	0	1	0		
	3=401-600	29	0	0	1		
	4=>601	3	0	0	0		
TRM4 Operasyon Sonrası Yeniden Kodlanmış Trombosit Değerleri	2=101-200	85	1	0			
	3=201-300	140	0	1			
	4=>301	51	0	0			
WBC5 Danışman Hekim Tarafından Risk Değerine Göre Yeniden Kodlanan Operasyon Öncesi Lökosit Değerleri	0=<7999	113	1				
	1=>8000	163	0				

Lojistik regresyon analizi ileriye doğru değişken ekleme metodu seçilerek gerçekleştirildikten sonra elde edilen sınıflandırma tablosu her adımda tahmin oranının nasıl değiştiğini göstermektedir (Bk. Tablo 3.10). Birinci adımda hasta olan

vakaların hasta olarak tespit edilme oranının %69,5 olarak tespit edilirken 2. Adımda bu oran %90,4'e çıkmış ve bu adımda dâhil edilen değişkenin tahmin oranına oldukça yüksek oranda etki ettiği tespit edilmiştir. Sonraki adımda bu kez değişkenler değiştirilerek işlem gerçekleşmiş bu kez kontrol grubu vakalarını belirleme oranı 2. adıma göre artarak %41 olurken hasta olan vakaları tespit etme oranı %84,7 olarak tespit edilmiştir. En uygun modele 5.adımda ulaşılmıştır. Bu adımda analizde bulunan ve p değeri olarak %5'den küçük olan ve güven aralığı olarak 1'i içermeyen değişkenler bizim için anlamlı değişkenlerdir. Tablo 3.11'de her adımda dâhil edilen değişkenleri bunların birbiri ile ilişkileri, güven aralıkları ve değişimi kaç kat etkilediği gösterilmiştir.

Tablo 3.9 Lojistik Regresyon Analizi İlk Adım Sınıflandırma Tablosu

Gözlenen Hasta ve Kontrol Grubu (<i>GRP</i>)	Tahmin Edilen Seçilmiş			Tahmin Edilen Seçilmemiş		
	Hasta ve Kontrol Grubu(<i>GRP</i>)		Başarı Oranı	Hasta ve Kontrol Grubu(<i>GRP</i>)		Başarı Oranı
	1=Hasta	0=Kontrol		1=Hasta	0=Kontrol	
1=Hasta	177	0	100	0	0	.
0=Kontrol	99	0	0	1	0	0
Toplam Yüzde			64,1			0

İleriye doğru değişken ekleme metodu kullanılarak yapılan lojistik regresyon analizinde 1.adımda sadece WBC5 değeri ile analiz yapılmış ve p küçük %5 ($p=0,000$) olarak oldukça anlamlı bulunmuştur. 2. adımda WBC5 ile birlikte WBC3 değişkeni analize alınmış kategorik olarak kodlanan WBC3 değişkeninin 3. alandaki değerleri test için anlamlı bulunmuştur. Analiz bu şekilde ilerleyerek en son 5. adımda tüm değişkenler analize dâhil edilmiştir. Bizim için anlamlı olan son adımdır. Bu adımda hangi değişkenlerin analizde etkili olduğu sonucuna ulaşılmaktadır. İleriye doğru değişken ekleme metodunda YAŞ değişkeni $p=0.03$, WBC5(1) $p=0.006$, WBC3 $p=0.01$, TRM4 $p=0.02$ ile anlamlı bulunurken, TRM4(2) $p=0.934$ olarak bulunmuş ve anlamlı olmadığı tespit edilmiştir. Bulunan bu değişkenler kanser vakalarını tespit etme açısından anlamlı birer değişken olarak sonraki çalışmalarda da kullanılabileceği bilgisini vermektedir.

Tablo 3.10 Lojistik Regresyon Analizi İleriye Doğru Değişken Ekleme Metodu Sınıflandırma Tablosu

Adım	Gözlenen		Tahmin Edilen		Başarı Yüzdesi
			Hasta ve Kontrol Grubu		
			Kontrol	Hasta	
Adım 1	Hasta ve Kontrol Grubu	Kontrol	59	41	59,0
		Hasta	54	123	69,5
	Toplam Yüzde				65,7
Adım 2	Hasta ve Kontrol Grubu	Kontrol	31	69	31,0
		Hasta	17	160	90,4
	Toplam Yüzde				69,0
Adım 3	Hasta ve Kontrol Grubu	Kontrol	41	59	41,0
		Hasta	27	150	84,7
	Toplam Yüzde				69,0
Adım 4	Hasta ve Kontrol Grubu	Kontrol	42	58	42,0
		Hasta	26	151	85,3
	Toplam Yüzde				69,7
Adım 5	Hasta ve Kontrol Grubu	Kontrol	48	52	48,0
		Hasta	31	146	82,5
	Toplam Yüzde				70,0

Aynı değişkenler kullanılarak lojistik regresyon analizi bu kez geriye doğru değişken eleme metodu kullanılarak gerçekleştirildiğinde sıfırıncı adımda analizden herhangi bir değişken atılmadan yani tüm değişkenler analize dâhil edildiğinde sınıflandırma bulguları Tablo 3.12’de verilmiştir. Sabit değişken analizde iken herhangi bir değişken dâhil edilmeden önceki grupta doğru tahmin etme genel yüzdesi %63,9 olarak tespit edilmiştir. Bu değer sonraki adımlarda analize eklenen değişkenlerin bu başarı oranını ne kadar arttırabildiği için önemli bir bilgidir.

Tablo 3.11 İleriye Doğru Değişken Ekleme Metodu Lojistik Regresyon Analizi Adımlar ve Güven Aralığı Değerleri

Adım	B	SH.	Wald	Sd	Anl.	Exp(B)	Exp(B) için 95% Güven Aralığı		
							Alt	Üst	
1	WBC5(1)	-1,2	0,2	20,7	1	0,000	0,3	0,2	0,5
	Sabit	1,1	0,2	37,1	1	0,000	3,0		
2	WBC3			9,2	3	0,027			
	WBC3(1)	20,7	22509,9	0,0	1	0,999	927518098,4	0,0	.
	WBC3(2)	-0,0	0,4	0,0	1	0,966	1,0	0,5	2,1
	WBC3(3)	-0,8	0,3	6,8	1	0,009	0,44	0,2	0,8
	WBC5(1)	-1,2	0,3	17,0	1	0,000	0,3	0,2	0,5
	Sabit	1,4	0,2	33,6	1	0,000	4,2		
3	WBC1	0,0	0,0	7,1	1	0,008	1,0	1,0	1,0
	WBC3			11,5	3	0,009			
	WBC3(1)	23,1	22494,9	0,0	1	0,999	11206941176,2	0,0	.
	WBC3(2)	1,9	0,8	5,7	1	0,016	6,4	1,4	29,0
	WBC3(3)	0,4	0,5	0,5	1	0,491	1,4	0,5	3,8
	WBC5(1)	-1,1	0,3	14,6	1	0,000	0,3	0,2	0,6
	Sabit	-2,7	1,5	3,2	1	0,075	0,1		
4	WBC1	0,0	0,0	7,9	1	0,005	1,0	1,0	1,0
	TRM4			9,3	2	0,009			
	TRM4(1)	1,0	0,4	5,6	1	0,018	2,8	1,2	6,4
	TRM4(2)	,1	0,4	0,0	1	0,896	1,1	0,5	2,2
	WBC3			10,6	3	0,014			
	WBC3(1)	22,6	22437,5	0,0	1	0,999	6678317836,3	0,0	.
	WBC3(2)	1,8	0,8	5,1	1	0,024	6,1	1,3	29,1
	WBC3(3)	,3	0,5	0,4	1	0,539	1,4	0,5	3,8
	WBC5(1)	-1,2	0,3	15,7	1	0,000	0,3	0,1	0,5
Sabit	-3,3	1,7	4,0	1	0,044	0,0			
5	WBC1	0,0	0,0	8,0	1	0,005	1,0	1,0	1,0
	TRM4			7,8	2	0,020			
	TRM4(1)	0,9	0,4	4,5	1	0,034	2,5	1,071	6,0
	TRM4(2)	0,0	0,4	0,0	1	0,934	1,0	0,5	2,2
	WBC3			10,2	3	0,017			
	WBC3(1)	23,0	21313,9	0,0	1	0,999	9400122649,8	0,0	.
	WBC3(2)	1,8	0,8	4,9	1	0,027	6,0	1,2	29,2
	WBC3(3)	0,3	0,5	0,3	1	0,541	1,4	0,5	3,9
	WBC5(1)	-1,2	0,3	15,1	1	0,000	0,3	0,1	0,6
YAS	0,0	0,0	4,7	1	0,030	1,0	1,0	1,1	

Tablo 3.12 Geriye Doğru Değişken Eleme Metodu Lojistik Regresyon Analizi Sınıflandırma Tablosu

Gözlenen	Tahmin Edilen Hasta ve Kontrol Grubu		Başarı Yüzdesi
	Kontrol	Hasta	
Kontrol	0	100	0,0
Hasta	0	177	100,0
Genel Yüzde			63,9

Analiz tüm değişkenler dâhil edilerek gerçekleştirildiğinde Tablo 3.13'deki sonuçlara ulaşılmıştır. Birinci adımda tüm değişkenler dâhil edildiğinde hasta olan grubu hasta olarak doğru tahmin etme olasılığı %84,7, kontrol grubunu doğru tahmin etme oranı %47, genel toplam %71.1 olarak tespit edilmiştir. Bu şekilde değişkenler analizden çıkarılarak işlem 5 adım gerçekleştirilmiştir. Tablo 3.14'de verilen beşinci adım sonuçları analiz edildiğinde, değişkenlerin p değerlerine bakılarak TRM4 $p=0.005$, TRM4(1) $p=0.007$, WBC3 $p=0.04$, WBC4(3) $p=0.03$, WBC5(1) $p=0.02$ olarak tespit edilmiş olup %5 anlamlılık değerinin altında oldukları için anlamlı değişkenler olarak kabul edilmiştir.

Tablo 3.13 Geriye Doğru Değişken Eleme Metodu Lojistik Regresyon Analizi Sınıflandırma Tablosu

Adım	Gözlenen	Tahmin Edilen		Başarı Yüzdesi
		Kontrol	Hasta	
Adım-1	Kontrol	47	53	47,0
	Hasta	27	150	84,7
	Genel Yüzde			71,1
Adım-2	Kontrol	49	51	49,0
	Hasta	27	150	84,7
	Genel Yüzde			71,8
Adım-3	Kontrol	48	52	48,0
	Hasta	27	150	84,7
	Genel Yüzde			71,5
Adım-4	Kontrol	44	56	44,0
	Hasta	27	150	84,7
	Genel Yüzde			70,0
Adım-5	Kontrol	46	54	46,0
	Hasta	28	149	84,2
	Genel Yüzde			70,4

Tablo 3.14 Geriye Doğru Değişken Eleme Metodu Lojistik Regresyon Analizi Adımlar ve Güven Aralığı Değerleri

Değişken	B	SH	Wald	sd	Anl.	Exp(B)	Exp(B) için 95% Güven Aralığı	
							Alt	Üst
<i>TRM4</i>			10,6	2	0,005			
<i>TRM4(1)</i>	1,2	0,5	7,2	1	0,007	3,3	1,4	8,0
<i>TRM4(2)</i>	0,2	0,4	0,2	1	0,637	1,2	0,6	2,6
<i>WBC3</i>			7,9	3	0,049			
<i>WBC3(1)</i>	22,5	22219,9	0,0	1	0,999	876,6	0,0	.
<i>WBC3(2)</i>	1,5	0,8	3,4	1	0,064	4,6	0,9	23,1
<i>WBC3(3)</i>	0,2	0,5	0,2	1	0,697	1,2	0,4	3,5
<i>WBC4</i>			4,9	5	0,443			
<i>WBC4(1)</i>	-2,3	1,4	2,5	1	0,115	0,1	0,0	1,7
<i>WBC4(2)</i>	18,9	27190,9	0,0	1	0,999	8481,0	0,0	.
<i>WBC4(3)</i>	-2,6	1,2	4,7	1	0,030	0,1	0,0	0,8
<i>WBC4(4)</i>	-2,1	1,1	3,7	1	0,054	0,1	0,0	1,0
<i>WBC4(5)</i>	-2,0	1,1	3,5	1	0,063	0,1	0,0	1,1
<i>WBC5(1)</i>	-0,9	0,4	5,4	1	0,020	0,4	0,2	0,9
<i>WBC1</i>	0,0	0,0	4,7	1	0,031	1,0	1,0	1,0
Sabit	-0,8	2,1	0,1	1	0,717	0,5		

Anlamlılık düzeyleri %5'in altında bulunan değişkenlerin güven aralıkları 1 değerini içermediğinden analize dâhil edilebilirler. TRM4, TRM4(1), WBC3(1), WBC4(3) ve WBC5(1) değerleri analiz sonucu anlamlı olarak bulunmuştur. İleriye doğru değişken ekleme ve geriye doğru değişken eleme metotları sonuçları değerlendirildiğinde her iki durumda da WBC3, WBC5 ve TRM4 değerlerinin hasta grubunu belirlemede ayırıcı etkenler olarak kullanılabilceği sonucuna ulaşılmıştır.

3.1.3.1. Model Seçim Yöntemleri ve Uygunluğunun Değerlendirilmesi

Model seçim yöntemlerindeki asıl amaç en uygun değişkeni ve en iyi temsil eden modeli belirlemektir. Seçim stratejileri araştırmanın amacına bağlı olarak değişiklik gösterebilir. Bu çalışmada geriye dönük olarak değerlendirilen rahim kanseri tanısı almış ve operasyon geçirmiş hastalar ile herhangi başka bir sebepten operasyon geçirmiş ancak kanser teşhisi konmamış hastalar arasında lojistik regresyon analizi kullanılarak eldeki verilerin rahim kanseri tanısını belirlemede bir kriter olarak kullanılıp kullanılmayacağı durumu incelenecektir. Adımsal seçim yöntemlerinden öncelikle ileriye doğru değişken ekleme yöntemi kullanılarak hangi de-ğiş-

kenlerin model belirlemede daha anlamlı olduğuna bakılmıştır. Öncelikle hiçbir de-ğişken dâhil edilmeden yapılan analize sırayla deęişkenler eklenir ve en uygun mo- del belirlenmeye çalışılır.

Öncelikle modelimiz belirlenmelidir. Veri setimizdeki WBC2, WBC1, TRM2, TRM1 ve kategorik deęişken olarak yeniden kodlanmış olan WBC4 deęiş- kenlerinin rahim kanseri üzerinde anlamlı bir etkiye sahip olduğunu düşünölmek- tedir. Bu şekilde oluşturmuş olduğumuz WBC5 deęişkenimizi Ki-Kare risk analizi ile hasta ve kontrol grupları üzerinde nasıl bir etkiye sahip olduğu görölebilir. De- ğişkenimizi belirlerken belli bir risk deęeri seçilir bu bizim deęişkenimiz için 8000'den küçük deęerler 0, 8.000'den büyük ve eşit deęerler 1 olarak kodlanmış ve rahim kanseri için bu deęer risk olarak belirlenmiştir.

Model seçimi yapılırken bazı kriterlere oldukça dikkat etmek gerekmektedir. Tablo 3.15'de -2loglikelihood deęerine bakıldığında 5. adımda 304,1'e kadar düştüğü görölmüştür. Bu deęer ne kadar küçük olursa analizimiz için modelinin o kadar uygun olduğu anlamına gelmektedir. Cox Snell R-kare ifadesindeki deęer be- şinci adımda 0,19(%19) olarak bulunmuştur. Bu deęer bizim modelimizin, oluştur-duğumuz hipotez de rahim kanseri tahmin deęerini göstermektedir.

Tablo 3.15 İleri Doğru Deęişken Ekleme Model Özeti

Adım	-2 Loglikelihood	Cox Snell R-Kare	Nagelkerke R-Kare
1	340,9	0,074	0,102
2	328,0	0,116	0,159
3	318,8	0,145	0,199
4	308,9	0,175	0,240
5	304,1	0,190	0,260

Birden fazla deęişken analize dâhil edileceęi zaman kullanılacak yöntemde farklı alternatifler olabilmektedir. Tam regresyon metodu deęişkenlerin birbiri üze- rinde herhangi bir üstünlüğü olmadığı bir dięer anlatımda deęişkenler arasında bir ayırım yapılmadan analize alınacağı durumda kullanılacak bir deęişken seçim yön- temidir. İleriye doğru deęişken ekleme yönteminde sırayla deęişkenler analize alı- narak hangi deęişkenlerin analiz için daha önemli olduğuna bakılır. Geriye doğru deęişken eleme yönteminde önce tüm deęişkenler analize alınır sırayla denenerek

en anlamsız deęişkenler analizden çıkarılarak deęişken elemeye devam edilir. Geriye doęru deęişken eleme ve ileriye doęru deęişken ekleme yöntemleri çok deęişkenli modellerde zaman açısında kayıp oluşturmaktadır buda modellerin dezavantajıdır.

Tablo 3.16 Geriye Doęru Deęişken Eleme Metodu Model Özeti

Adım	-2 Loglikelihood	Cox Snell R-Kare	Nagelkerke R-Kare
1	294,337	0,218	0,298
2	294,449	0,217	0,298
3	294,606	0,217	0,297
4	295,402	0,215	0,294
5	299,207	0,204	0,279

Geriye doęru deęişken eleme metodunda loglikelihood istatistięi deęerleri Tablo 3.16'da verilmiştir. Birinci ve beşinci adım deęerleri arasında çok büyük farklar görülmemektedir. Cox-Snell deęerlerine bakıldığında birinci adımda %21 olarak tespit edilen tahmin deęeri beşinci adıma %20 olarak bulunmuştur. Model seçim yöntemlerinden kullanılacak yöntemin iyilięi için loglikelihood deęerlerine bakılır. Bu deęer hangi analizde daha düşük olarak tespit edilmiş ise o yöntem deęişken seçimi için kullanılabilir doęru yöntem olarak tercih edilmelidir.

Model seçimindeki R^2 deęerleri bize anlamlı bir model seçmeyi gösterirken, seçmiş olduğumuz modelin uyum iyilięi Hosmer-Lemeshow testi ile deęerlendirilmektedir.

BULGULAR

Tüm olguların; yaşı, histolojik dereceleri, histopatolojik tipleri, ortalama tümör boyutu, Miyometrial invazyonu, adneksiyel tutulumu, peritoneal sitoloji, lenfovasküler tutulumu, pelvik-paraaortik lenf nodu tutulumu, cerrahi evresi değerlendirilmeye alındı. Yeniden adlandırılarak değişken isimleri tespit edildi. WBC3, WBC4, TRM3 değişkenleri WBC1, TRM1 değişkenleri kullanılarak kategorik olarak yeniden oluşturuldu.

Hastaların cerrahi evrelemesi, FIGO 2009 sınıflamasına göre yapıldı. Trombosit ve lökosit sayımı için ameliyat öncesi 14 günlük aralıkta alınan kan örnekleri çalışmaya dâhil edildi. Lökosit sayım değerlerinin büyük veya eşit 10.000K/mm³ olması lökositoz, trombosit sayısının büyük veya eşit 450K/mm³ olması trombositoz olarak kabul edildi. Bu kabulden sonra lökosit değerlerinin yükselmeye başladığı istatistik olarak anlamlı bir sonuç oluşturup oluşturmadığının tespiti için lökositoz değişkeni ile büyük veya eşit 10.000 K/ mm³ değerine sahip olanlar 1, küçük 10.000 K/mm³ değerine sahip olanlar 0 olarak yeniden kodlandı.

Hasta ve kontrol grubu binary lojistik regresyon analizi lökosit ve trombosit değerleri bakımından anlamlı bir ilişki içinde olup olmadığının tespiti için grup değişkeni adı altında hasta olan olgular 1, kontrol grubundaki olgular 0 olarak yeniden adlandırıldı.

Birinci adımda lökositoz ile belirtilen değişken ile grup değişkeninde bulunan iki kategorili hasta ve kontrol olguları arasında anlamlı bir ilişki olup olmadığına bakıldı. Tablo 4.2’de hasta olup lökositozu olmayan 140 vaka toplam sayının %79.1’lik kısmını oluştururken, lökositoz olan hasta sayısı 37 ile %20.9’luk kısmını oluşturuyordu. Kontrol grubunun lökositozu olmayan olgu sayısı 94 ile %94’lük kısmını, lökositoz olan olgu sayısı 6 ile %6’lık kısmını oluşturuyordu. Grup değişkeni ile lökositoz değişkeni binary lojistik regresyon analiziyle test edildiğinde lökositoz olan hastaların anlamlılık değeri $p < 0.05$ ($p = 0,002$) ile anlamlı bir ilişki olduğunu bize gösterdi (Bkz. Tablo 4.1, Tablo 4.2, Tablo 4.3).

Tablo 4.1 Lökositöz Modele Dâhil Edildiğinde Oluşan Sonuçlar

-2 Loglikelihood	Cox Snell R-Kare	Nagelkerke R-Kare
350,045 ^a	0,043	0,059

Bağımlı ve bağımsız değişken sayıları değiştirilerek testin anlamlılığı sorgulanmaya devam edildiğinde grup değişkeni bağımlı değişken olarak kalırken bağımsız değişkenler arttırıldı.

Tablo 4.2 Lökositöz ve Grup Değişkeni Arasında Lojistik Regresyon

	B	S.E.	Wald	Sd	Anl.	Exp(B)	EXP(B) İçin %95	
							Alt	Üst
Lökositöz(1)	-1,4	0,5	9,5	1	0,002	0.2	0.1	0,6
Sabit	1.8	0,4	17.1	1	0,000	6.2		

Lökositöz, WBC1 ve kategorik olarak yeniden kodlanan WBC3 ile değişken seçim yöntemi olarak tam metodu kullanılarak analiz yapıldığında yeni -2LogL istatistiği değerimiz 333,6 olarak bulundu. Sayısal olan WBC1 değişkeni ile kategorik olan WBC3 değerleri Tablo 4.5'te görüldüğü gibi her iki durumda da anlamlı bir ilişki göstermiştir.

Tablo 4.3 Lökositöz Olan ve Olmayan Hasta ve Kontrol Grubu Çapraz Tablosu

Grup	Lökositöz		Toplam
	0=(Lökositöz Olmayan)	1=(Lökositöz)	
Hasta	140 (%79)	37 (%21)	177 (%100)
Kontrol	82 (%94)	18 (%16)	100 (%100)

Değişken ekleme yöntemimiz Forward-Stepwise LR olarak seçilip işlem tekrar gerçekleştirildiğinde adım adım sisteme değişkenler dâhil edilerek uyum iyiliği ve testin başarısı ölçülmüştür. Testin uyum iliği analizi Hosmer and Lemeshow yöntemi kullanılarak yapılmıştır. Adım-1'den Adım-2'ye geçerken p değeri 0,429 olup büyük bir değere sahiptir. Bu sonuç bize elde edilen modelin uyumlu olduğunu göstermiştir (Bk. Tablo 4.4).

Tablo 4.4 Lökositoz, WBC1 ve WBC3 Değişkenleri ile Lojistik Regresyon Analizi Sonucu Oluşan Durumlar (1.Adım)

Değişken	B	SH	Wald	Sd	Anl.	Exp(B)	EXP(B) için 95% Güven Aralığı	
							Alt	Üst
WBC3			8,2	3	0,042			
WBC3(1)	23,2	23192,3	0,0	1	0,999	11970424657,6	0,0	.
WBC3(2)	1,7	0,9	3,9	1	0,047	5,5	1,0	29,7
WBC3(3)	0,3	0,5	0,4	1	0,538	1,4	0,5	3,9
WBC1	0,0	0,0	6,9	1	0,009	1,0	1,0	1,0
Lökositoz(1)	0,5	0,7	0,4	1	0,519	1,6	0,4	6,5
Sabit	-4,6	2,4	3,7	1	0,056	0,0		

Testin gücüne baktığımızda hasta ve kontrol sayılarının tespitinin oldukça başarılı olduğu sonucuna ulaşmış bulunmaktayız. Başlangıçta modelimizde herhangi bir değişken eklenmeden %63.9 olarak belirlenen hasta ve kontrol tespit oranı Tablo 4.4’de tespit edildiği gibi %89.8’lik hasta doğru tespit oranı, %66.1 toplam doğru tespit oranı ile başarılı bir artış göstermiştir

Tablo 4.5 Lökositoz, WBC1 ve WBC3 Değişkenleri ile Lojistik Regresyon Analizi Test Uyum İyiliği (Hosmer-Lemeshow Testi)

Adım	Ki-Kare	Anl.
1	15,531	0,050
2	9,046	0,429

Değişkenler Tablo 4.5 ile elde edilen sonuçlarda görüldüğü gibi 1. adımda dâhil edilen WBC3 değeri $p < 0,05$ ($p = 0,000$) ile sayısal bir değişken olarak lojistik regresyon analizinde anlamlı bir ilişki göstermiştir. İkinci adımda dâhil edilen kategorik olarak kodlanmış WBC3 değişkeni de $p < 0,05$ ($p = 0,03$) ile anlamlı olduğunu göstermiştir.

Tablo 4.6 Sınıflandırma Tahmin Edilebilirliği

	Gözlem Grubu	Tahmin Grubu		Yüzdelerik Başarı
		Hasta	Kontrol	
Adım-1	Hasta	159	18	89,8
	Kontrol	93	7	7,0
	Genel Yüzde			59,9
Adım-2	Hasta	155	22	87,6
	Kontrol	72	28	28,0
	Genel Yüzde			66,1

Tablo 4.7 Grup Değişkeni ile Lökosit, WBC1, WBC3 Lojistik regresyon Analizi

	B	SH	Wald	Sd	Anl.	Exp(B)	EXP(B) İçin 95GA		
							Alt	Üst	
Adım-1	WBC1	0,0	0,0	12,2	1	0,000	1,0	1,0	1,0
	Sabit	1,1	0,5	5,1	1	0,024	3,0		
Adım-2	WBC3			8,5	3	0,036			
	WBC3(1)	-22,8	23196,2	0,0	1	0,999	0,0	0,0	
	WBC3(2)	-1,5	0,8	3,7	1	0,054	0,2	0,1	1,0
	WBC3(3)	-0,2	0,5	0,2	1	0,656	0,8	0,3	2,1
	WBC1	0,0	0,0	8,9	1	0,003	1,0	1,0	1,0
	Sabit	3,4	1,5	5,1	1	0,024	30,7		

Bu kez WBC3 değeri kategorik olarak kodlandığında ve WBC2, TRM1, TRM2, WBC3 ve WBC5 değişkenleri yeniden analiz yapıldığında toplam tahmin edilebilirlik %69 değerine yükselmiştir (Bk. Tablo 4.6.).

Tablo 4.8 Lökositöz ile WBC1, WBC2, TRM1, TRM2, WBC3 ve WBC5 Değişkenleri ile Sınıflandırma Sonuçları

	Gözlem Grubu	Tahmin Grubu		Yüzdelerik Başarı
		Hasta	Kontrol	
Adım-1	Hasta	123	54	69,5
	Kontrol	41	59	59,0
	Genel Yüzde			65,7
Adım-2	Hasta	160	17	90,4
	Kontrol	69	31	31,0
	Genel Yüzde			69,0

WBC5 değeri $p < 0,05$ ($p = 0,000$) ile olduğunda anlamlı bir ilişki göstermiştir. Tablo 4.9'da belirtilen WBC3(3) değeri kategorik olarak kodlanmış ve 3.basa-

mak değeri WBC1 değerinin yükseldiği bu basamakta yine anlamlı ilişkiyi göstermiştir. Lökositöz değeri baz olarak 10000 K/mm³ alındığında tek başına anlamlı bir ilişkiyi bize göstermiştir. WBC5 değeri ile baz aldığımız değeri 8000 K/mm³'ün altına çektiğimizde $p < 0.05$ ($p = 0,000$) oldukça anlamlı hale gelmiştir. Lökositöz ile beraber, WBC1, WBC2, TRM1, TRM2, WBC3 ve WBC5 değişkenleri seçilerek yeniden lojistik regresyon analizi yapılarak Tablo 4.9'daki sonuçlara ulaşıldı.

Tablo 4.9 WBC3 Değişkeni ile WBC5 Arasında Lojistik Regresyon Analizi

		B	SH	Wald	Sd	Anl.	Exp(B)	EXP(B) İçin 95% GA	
								Alt	Üst
Adım-1	WBC5	1,2	0,3	20,7	1	0,000	3,3	2,0	5,5
	Sabit	-1,1	0,2	37,1	1	0,000	0,3		
Adım-2	WBC3			9,2	3	0,027			
	WBC3	-20,6	22509	0,0	1	0,999	0,0	0,0	
	WBC3	0,0	0,4	0,0	1	0,966	1,0	0,5	2,2
	WBC3	,8	0,3	6,8	1	0,009	2,3	1,2	4,2
	WBC5	1,2	0,3	17,0	1	0,000	3,3	1,9	5,9
	Sabit	-1,4	0,2	33,6	1	0,000	0,2		

Kanserli hastalarda 177 olgunun 37'sinde (%21) lökositözis tespit edildi. Kontrol grubu olarak çalışmaya dâhil edilen olgularda lökositözis oranı ise %6 olarak tespit edilmiştir. İki kategorili olarak belirlenen grup değişkeni altında binary lojistik regresyon analizi yapıldığında yukarıdaki tablolarımızda belirtmiş olduğumuz gibi bir bağımlı ve bir bağımsız değişken kullanılarak analizi uygulandığında %63,9 olarak tespit edilen model doğruluk oranı daha sonra uygulanan adımlar ile teste eklenen değişkenlerin varlığında doğruluk oranının %69'a kadar çıktığı tespit edilmiştir. $p < 0.05$ baz alınan anlamlılık değerinde lökositöz, WBC5 ve WBC3 değişkenleri ile anlamlı bir ilişki içinde olduğu tespit edilmiştir. İki farklı grup ile yaptığımız binary lojistik regresyon analizi bu kez de hasta grubunun kendi içinde tanımlamış olduğumuz iki kategorili değişkenler ile yeniden gerçekleştirildi. Lökositözu olan gruptaki hastaların lökositözu olmayanlara göre, daha büyük bir yüzdesinde adneksiyel tutulum, lenfovasküler invazyon ve paraaortik lenf nodu tutulumu tespit edildi. Hastalık evresi kategorik olarak kodlandığında ileri evre olan 3 ve 4 numaralı gruptaki hastaların daha büyük çoğunda lökositöz varlığı tespit edildi. İleri evre kategorik değişkeni ile lökositöz değişkeni teste tabi tutulduğunda değişken

ilave dilmeden önce belirlenen doğruluk tespit oranı %80,8'dir. Evre kategorik değişkeni ikili kategorik lökositöz değişkeni ile teste girdiğinde doğruluk oranı %84,7 olarak tespit edilmiştir bu değerler Tablo 4.10'da belirtilmiştir.

Lökositöz değişkeni Evre kategorik değişkeni arasında tespit edilen p değeri, $p < 0,05$ ($p = 0,000$) olduğundan oldukça anlamlı bir ilişki bulunmuştur (Bk. Tablo 4.11).

Tablo 4.10 Lökositöz Sınıflandırma Tahmin Tablosu

		Lökosit Durumu		Yüzdelerik Tahmin
		0=Lökosit Olmayan	1= Lökosit Olan	
Adım-1	0=Lökosit Olmayan	142	1	99,3
	1=Lökosit Olan	26	8	23,5
Genel Tahmin (%)				84,7

Tablo 4.11 Lökositöz ve Evre Değişkeni Arasında Lojistik Regresyon Analizi Sonuç

Adım-1	B	SH	Wald	Sd	Anl.	Exp(B)	EXP(B) İçin %95 GA	
							Alt	Üst
EVRE			21,2	3	0,000			
EVRE-1	-3,9	1,1	12,9	1	0,000	0,0	0,0	0,2
EVRE-2	-5,3	1,5	13,0	1	0,000	0,0	0,0	0,1
EVRE-3	-2,6	1,1	5,2	1	0,023	0,1	0,0	0,7
Sabit	2,1	1,1	3,8	1	0,050	8,0		

Miyometrial invazyonu olan hastalarda evre değişkeni arasında anlamlı bir ilişkinin olup olmadığına bakıldığında sıfırcıncı adımda %84,2 olarak tespit edilen Miyometrial invazyon tespit oranı değişkenin dâhil edilmesi ile %88,1 olarak tespit edilmiştir. Anlamlılık durumu incelendiğinde Miyometrial invazyon ile evre değişkeni arasındaki ilişki $p < 0,05$ ($p = 0,000$) tespit edilmiş olup anlamlı bulunmuştur.

Tablo 4.12 Miyometrial İnvazyon ve Evre Arasında Lojistik Regresyon Analizi

Adım-1	B	SH	Wald	Sd	Anl.	Exp(B)	EXP(B) İçin %95 GA	
							Alt	Üst
EVRE			27,5	3	0,000			
EVRE-1	-4,8	1,1	18,4	1	0,000	0,0	0,0	0,1
EVRE-2	-3,8	1,2	10,1	1	0,001	0,0	0,0	0,2

EVRE-3	-2,6	1,1	5,2	1	0,023	0,1	0,0	0,7
Sabit	2,2	1,1	3,8	1	0,050	8,0		

Uygulama sonucunda oluşturulan grup değişkeni ile seçim yöntemleri çok değişkenli analiz ileri doğru (Forward-LR) seçim yöntemi kullanılmış olup tek değişken ile analiz yapılırken herhangi bir üstünlük durumu söz konusu olmadığı için tam regresyon metodu kullanılarak lojistik regresyon analizi yöntemi kullanılmıştır. Tablo 4.13’de değişken seçme yöntemi değiştirilerek bu kez geriye doğru eleme yöntemi kullanılarak yeniden analiz yapılmış ve adımların uyum analizleri ve testin gücüne bakılmıştır. Tablo 4.13’de maksimum olabilirlik oranları verilmiştir. Birinci adımda tüm değişkenler dâhil iken 320,2 çıkmış beşinci adımda WBC3 ve WBC5 değişkenleri varken bu oran 328,0’e yükselmiştir. Tablo 4.14’de sınıflandırma başarı oranları verilmiştir. Birinci adımda beş değişken dâhil edilmişken hasta grubunu doğru tahmin etme oranı %85,3 olarak, toplam doğru tahmin oranı %67,5 olarak bulunmuştur. Beşinci adımda bizim için önemli olan WBC3 değeri ve yine kategorik olarak WBC5 değişkenleri dâhil edildiğinde hasta grubunu doğru tahmin etme oranı %90,4’e çıkmıştır.

Tablo 4.13 Geriye Doğru Eleme Yöntemi İle Lojistik Regresyon Analizi Sonuçları

Adım	-2 Loglikelihood	Cox-Snell-Kare	Nagelkerke-Kare
1	320,2	0,141	0,193
2	321,2	0,138	0,189
3	323,1	0,132	0,181
4	325,5	0,124	0,171
5	328,0	0,116	0,159

Tablo 4.14 Tüm Değişkenler Dâhil İken Geriye Doğru Değişken Eleme Metodu Sınıflandırma Sonuçları

	Gözlem Grubu	Tahmin Grubu		Başarı Oranı
		Hasta	Kontrol	
Adım-1	Hasta	151	26	85,3
	Kontrol	64	36	36,0
	Genel Yüzde			67,5
Adım-2	Hasta	151	26	85,3
	Kontrol	63	37	37,0
	Genel Yüzde			67,9
Adım-3	Hasta	152	25	85,9
	Kontrol	65	35	35,0
	Genel Yüzde			67,5
Adım-4	Hasta	155	22	87,6
	Kontrol	68	32	32,0
	Genel Yüzde			67,5
Adım-5	Hasta	160	17	90,4
	Kontrol	69	31	31,0
	Genel Yüzde			69,0

Tablo 4.15’de lojistik regresyon oranı anlamlılık değerleri ve %95 güven aralığında tüm adımlar için anlamlı olan değişkenlerin başarı oranları belirtilmiştir. Birinci adımda tüm değişkenler dâhil edildiğinde bizim için anlamlı olan bir değişken bulunamamıştır. Üçüncü adımdan itibaren kategorik olarak kodlanan WBC3 ve WBC5 değişkenleri anlamlı sonuçlar belirtmiştir. Üçüncü adımda WBC3 değişkeni %95 güven aralığında grup değişkenimizi 1,2 ile 4,3 kat arasında etkilemektedir. Beşinci adımda kalan iki değişken ile testin anlamlılık düzeyi $p < 0,05$ ($p = 0,000$) WBC5 değişkeni ile anlamlı bir ilişki olduğunu göstermiştir.

Tablo 4.15 Geriye Doğru Değişken Eleme Yöntemi Lojistik Regresyon Analizi

Adım-5	B	SH	Wald	Sd	Anl.	Exp(B)	EXP(B) İçin %95	
							Alt	Üst
<i>WBC3</i>			9,2	3	0,027			
<i>WBC3-1</i>	-20,6	22509,9	0	1	0,999	0	0	
<i>WBC3-2</i>	0,0	0,4	0,0	1	0,966	1,0	0,5	2,2
<i>WBC3-3</i>	0,8	0,3	6,8	1	0,009	2,3	1,2	4,3
<i>WBC5-1</i>	1,2	0,3	17,0	1	0	3,3	1,9	5,9
<i>Sabit</i>	-1,4	0,2	33,6	1	0	0,2		

İleriye doğru deęişken ekleme metodunda elde ettiđimiz bulgular gibi geriye doğru deęişken eleme metodunda da WBC3 ve WBC5 deęişkenleri çalışmamızda bulmaya çalıştığımız hastalık teşhisi için oldukça anlamlı deęişkenler olarak tespit edilmiştir.



SONUÇ VE ÖNERİLER

Çok değişik konularda uygulama alanlarına sahip olan ve anket türü verilerin analizinde çok kullanılan lojistik regresyon analizi incelenmeye çalışılmış ve sonuçları yorumlanmıştır.

Bağımlı değişkenin iki kategorili olduğu bir durumda çok değişkenli lojistik regresyon analizi yönteminin kullanıldığı sağlık verileri üzerine bir uygulama yapılmıştır.

19 Mayıs Üniversitesi hastanesine başvuran ve geriye dönük teknikler ile hastane veri tabanı taranarak elde edilen rahim kanseri teşhisi konmuş hasta ve sadece operasyon ancak rahim kanseri teşhisi konmamış toplam 277 adet vaka analize dâhil edilmiştir. Bağımlı değişken iki düzeyli hasta grubu 1, hasta olmayan grup 0 değeri ile belirtilmiştir. Analize dâhil edilen tüm değişkenlerin anlamlı olup olmadıkları ileri doğru değişken ekleme ve geriye doğru değişken ekleme metotları ile değerlendirilmiştir. On dört adet değişken analize dâhil edilmiş bu değişkenlerden beş tanesinin modelde yer aldığı görülmüştür.

Hasta ve kontrol grupları yaş değişkeni açısından değerlendirilmiş hasta grubun yaş ortalamasının anlamlı derecede yüksek olduğu bulgusuna ulaşılmıştır.

Operasyon öncesi lökosit ve trombosit değerlerinin rahim kanseri için anlamlı bir değişken olup olmadığının tespiti için yapılan lojistik regresyon analizi tekniğinde kategorik olarak kodlanmış operasyon öncesi lökosit değişkeninin ve danışman hekim tarafından baz değer olarak belirlenen operasyon sonrası lökosit risk değerinin hasta grubunda anlamlı düzeyde yüksek olduğu tespit edilmiştir.

Lökositozu olan gruptaki hastaların lökositozu olmayanlara göre daha büyük bir yüzdesinde adneksiyel tutulum, lenfovasküler invazyon ve paraaortik lenf nodu tutulumu izlenmiştir.

Hastalık evresi kategorik olarak kodlandığında ileri evre olan 3 ve 4 numaralı grupta yer alan hastaların büyük çoğunluğunda lökositoz tespit edilmiştir.

Analiz sonuçları sınıflandırma değerleri ileri doğru değişken ekleme ve geriye doğru değişken eleme metotlarında hasta grubunu doğru tahmin etme oranı maksimum %90,4 olarak bulunurken, genel başarı yüzdeleri %69 olarak tespit edilmiştir.

Çapraz sorgular ve uygulanan diğer istatistikler ile farklı açılardan sorgulamalar yapılarak testin önemi denenmiş farklı sorgular ile değişkenler arasındaki ilişkiler tespit edilmiştir.

İleriye doğru değişken ekleme veya geriye doğru değişken eleme model seçim yöntemleri uygulanırken kullanılan değişkenlerin aynı olmasına rağmen iki farklı yöntem için farklı sonuçlar çıkabileceği bilgisine ulaşılmıştır. İleriye doğru değişken ekleme metodunun daha kapsamlı bir araştırma tekniği olarak kullanılabilceği, geriye doğru eleme yönteminin ise başta tüm değişkenlerin dâhil edilmesi ile karışık bir hal almasına sebep olduğu gözlemlenmiştir. Ancak olabilirlik fonksiyonu sonuçları değerlendirildiğinde geriye doğru değişken eleme metodu daha düşük sonuçlar bildirmiştir.

Lojistik regresyon analizinin birçok alanda olduğu gibi sağlık alanında da eldeki mevcut veriler ile daha sonra dâhil edilecek veriler hakkında yorum yapılabilme imkânı tanınması sebebiyle oldukça önemli bir test olduğu bir kez daha tespit edilmiştir.

KAYNAKLAR

- Akar, M., Şahinler, S. İstatistik. Çukurova Ün. Ziraat Fakültesi Yayınları No:74, Ders Kitaplar No:17, Ç.Ü.Z.F. Ofset Atölyesi, Adana , 1997.
- Akkuş, Z., ve Çelik, M. Y. Lojistik regresyon ve diskriminant analizi yöntemlerinde önemli ölçütler. VII. Ulusal Biyoistatistik Kongresinde sunulan bildiri. Mersin Üniversitesi, Tıp Fakültesi, Biyoistatistik Anabilim Dalı, Mersin, Eylül-Ekim 2004.
- Aktas, C. ve Yılmaz V. “Eskişehir’de LPG Kullanan Özel Araç Sürücülerinin Sınıflandırılmasında Lojistik Regresyon Analizi” İstanbul Kent içi Ulaşım Sempozyumu, İstanbul, 251-256, 2001
- Aktaş, D. ve Giray, S. Mali Başarısızlığın Çok Değişkenli İstatistiksel Yöntemlerle Belirlenmesi: Tekstil Sektörü Örneği: Anadolu Üniversitesi Sosyal Bilimler Dergisi. Cilt.05, Sayı.2,2005
- Aktaş, C. ve Erkuş, O. Lojistik Regresyon Analizi ile Eskişehir’in Sis Kestiriminin İncelenmesi Investigating of Fog Forecasting of Eskişehir Using Logistic Regression Analysis: İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, Cilt 8, Sayı 16,2009.
- Alan, A., “An Introduction to Categorical Data Analysis”, Barnett, V., Bradley, R.,A., Fisher, N.,I., Hunter J.,S., Kadane, J.,B., Kendall, D.,G., Scott, D.,W., Smith, A.,F.,M., Teugels, J.,L., Watson, G.,S., Wiley Series In Probability And Stastisytics, Canada, 2,2-3,22,103, 1996.
- Albayrak, A. Uygulamalı Çok Değişkenli İstatistik Teknikleri, Asil Yayın Dağıtım, 2006.
- Alpar, R., “Uygulamalı Çok Değişkenli İstatistiksel Yöntemlere Giriş-1”, Kültür Ofset, Ankara, 275 ,1997.
- Alpar, R. Uygulamalı Çok Değişkenli İstatistiksel Yöntemler Detay Yayıncılık Ankara,2013.

- Arıcı, H., İstatistik Yöntemler ve Uygulamaları. METEKSAN A.Ş. Ankara 2001.
- Arıcı, H. 2001 Akar ve Şahinler, 1997.
- Atakurt, Y.: Lojistik Regresyon Analizi Ve Tıp Alanında Kullanımına İlişkin Bir Uygulama, Atatürk Üniversitesi, 1999.
- Atasoy, D. Lojistik regresyon analizinin incelenmesi ve bir uygulaması. Yayınlanmamış yüksek lisans tezi, Cumhuriyet Üniversitesi, Sosyal Bilimler Enstitüsü, Sivas, 2001.
- Ayhan, S. Sıralı Lojistik Regresyon Analiziyle Türkiye'deki Hemşirelerin İş bırakma Niyetini etkileyen Faktörlerin Belirlenmesi.(Yayınlanmamış Yüksek Lisans Tezi). Eskişehir: Osmangazi Üniversitesi Fen Bilimleri Enstitüsü, 2006.
- Bartz, 2001; Moore, 1997; Couch ve Stoloff, 1989.(<http://www.ex.ac.uk/~SEGL/ea/multvar2/disclogi.html>).
- Başarır, G. Çok Değişkenli Verilerde Ayrımsama Sorunu ve Lojistik Regresyon Analizi. (Uygulamalı istatistik doktora tezi.) H.Ü., 1-36, Ankara, 1990.
- Bonney, G.E. (1987), Logistic Regression for Dependent Binary Observations, Biometrics, 43: 951-973.
- Bircan, H. "Lojistik Regresyon Analizi: Tıp Verileri Üzerine Bir Uygulama", Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 2, 185-208, 2004.
- Canbaz, S. Tıp fakültesi son sınıf öğrencilerinin kaygı düzeyi ve etkileyen faktörler, Genel Tıp Dergisi, 17:15-19,2007
- Cornfield, J. Joint Dependence of the Risk of Coronary Heart Disease on serum Cholesterol and Systolic Blood Pressure: A Discriminant Function Analysis, Federation Proceedings, 21: 58-61.(1962)

- Cox, B.D., Whichelow, H.J., Prevost, A.T. The Development of Cardiovascular disease in relation to anthropometric indices and hypertension in British adults. *International Journal of Obesity*, 22: 97330-10966, 1998.
- Çok, G. İleri evre küçük hücreli dışı akciğer karsinomunda sisplatin-etoposid ile mitomisin-ifosfamid-sisplatin kombinasyonlarının karşılaştırılması *Tüberküloz ve Toraks Dergisi*:54:161-167,2006
- Discriminant function analysis. <http://userwww.sfsu.edu/~efc/classes/biol710/discrim/discrim.pdf> adresinden 22 Kasım 2008 tarihinde edinilmiştir.
- Düzenli, H. Kronik solunum yetmezliği olan olgularda uzun süreli oksijen tedavisinin yaşam süresi üzerine etkisi *Tüberküloz ve Toraks Dergisi*; 56(2): 179-186,2008.
- Düzgüneş, O., Kesici, T., Kavuncu, O. ve Gürbüz, F., *Araştırma ve Deneme Metotları (İstatistik Metotları - II)*. 381 , Ankara Üniversitesi, Ziraat Fakültesi Yayınları, Ankara. 1987.
- Ediz, B.: *Lojistik Regresyon: Ayırma Analizi*. Doktora Tezi, 1997.
- Ege, İ. ve Bayrakdaroğlu, A. İMKB Şirketlerinin Hisse Senedi Getiri Başarılarının Lojistik Regresyon Tekniği İle Analizi: *ZKÜ Sosyal Bilimler Dergisi*, Cilt 5, Sayı 10, ss. 139–158, 2009.
- Elhan, Atilla H.: *Lojistik regresyon analizin incelenmesi ve tıpta bir uygulaması*, Ankara Üniversitesi Yüksek Lisans Tezi, 1997.
- Ergün, M. *Bilimsel Araştırmalarda Bilgisayarlarla İstatistik Uygulamaları, SPSS for Windows*, Ankara: Ocak yay. 1995.
- Ersan, Y. vd. Mide Kanserli Hastalarda Klinikopatolojik Prognostik Faktörler: *Cerrahpaşa Tıp Dergisi*, Cilt 34,Sayı 1, 2003.

- Ersin, M. Lise Öğrencisi Ergenlerde Depresyonun Yaygınlığı ve İlişkili Olduğu Etmenler. *Türk Psikiyatri Dergisi*; 19(4):382-389 ,2008.
- Field, A. *Discovering statistics using SPSS (2nd ed.)*. London: Sage. 2005.
- Foreman, K. K. *Design and evaluation of computer-assisted instruction in the health sciences*. Unpublished Doctoral Dissertation. The University of Utah, ABD, 2005.
- Girginer, N. ve Cankuş Tramvay Yolcu Memnuniyetinin Lojistik Regresyon Analiziyle Ölçülmesi: Estram Örneği: *Yönetim ve Ekonomi* 15/1 , 181-193:2008.
- Grimm, L.G. & Yarnold, P. R. (Eds.). *Reading and understanding multivariate statistics*. Washington D.C.: American Psychological Association, 1995.
- Gürcan, M. *Lojistik regresyon analizi ve bir uygulama*. Yayınlanmamış yüksek lisans tezi, Ondokuz Mayıs Üniversitesi, Fen Bilimleri Enstitüsü, Samsun. 1998.
- Hosmer, D. W., Lemeshow, S.: *Applied Logistic Regression*, John Wiley & Sons, 1989.
- Kalaycı, Ş. (Ed.) *SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri*. Ankara: Asil Yayın Dağıtım,2006.
- Kalaycı, Ş. (Ed.). *SPSS uygulamalı çok değişkenli istatistik teknikleri*. Ankara: Asil Yayın Dağıtım, 2005.
- Kaşko Y., “Çoklu Bağlantı Durumunda İkili (Binary) Lojistik Regresyon Modelinde Gerçekleşen I. Tip Hata ve Testin Gücü”, *Yüksek Lisans Tezi*, Ankara Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 20,25, 2007.
- Kaya, M. *Tıp Fakültesi ve Sağlık Yüksekokulu Öğrencilerinde Depresif Belirteç Yaygınlığı, Stresle Başa Çıkma Tarzları ve Etkileyen Faktörler*, *Türkiye Psikiyatri Dergisi* ,18:137-146,2007.

- Kılıç, S. Lojistik regresyon analizi ve pazarlama arařtırmalarında bir uygulama. Yayınlanmamıř yüksek lisans tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul. 2000.
- Kleinbaum, D. G., Kupper, L. L., Muller, K. E., and Nizam, A., Applied Regression Analysis and Other Multivariable Methods, 798, Duxbury Press, 511 Forest Lodge Road Pacific Grove, CA 93950 USA, 1998.
- Kleinbaum, David G.: Logistic Regression: A self-Learning Text, 1992.
- Koçak, İ.: Temel bileřenler analizi ve uygulaması, İnönü üniversitesi, Yüksek lisans tezi, 1997.
- Kutlu, Ö. vd. İlköğretim 5. Sınıf Öğrencilerinin Okuduğunu Anlamada Başarılı Olup-Olmama Durumlarının Kestirilmesinde Etkili Olan Değışkenlerin İncelenmesi. Eğitimde ve Psikolojide Ölçme ve Değıerlendirme Dergisi, 2(1): 132-139.2011.
- Lee, C.T, Logistic Models for Cross-over Designs. Biometrika, 71: 216-217, 1984.
- Leech, N. L., Barrett, K. C., & Morgan, G. A. SPSS for intermediate statistics: Use and interpretation (2nd ed). Mahwah, NJ: Lawrence Erlbaum Associates.; Poulsen, J., & French, A. 2008. London: Sage 2005.
- Maral, I. Depresyon Yaygınlığı ve Risk Etkenleri: Huzurevinde ve Evde Yařayan Yařlılarda Karşılařtırılmalı Bir Çalıřma; 12(4): 251-259,2001.
- Maulen, F. Case Study: An Application of Logistic Regression in a Six Sigma Project in Health Care, Quality Engineering, 23: 2, 113-124,2011.
- Mertler C.A. & Vannatta, R. A. Advanced and multivariate statistical methods: Practical application and interpretation (3rd ed.). Glendale, CA: Pyrczak Publishing 2005.
- Önder, H., Cebeci, Z. Lojistik Regresyonlarda Değışken Seçimi, Çukurova Üniversitesi Ziraat Fakültesi Dergisi, 17 (2):105-114, 2002.

- Özdamar, K. Paket Programlar ve İstatistiksel Veri Analizi, Kaan Kitapevi, Eskişehir, 2009.
- Özkan, S. vd. Bebek Ölüm Nedenleri Üzerine Bir Araştırma: Gazi Tıp Dergisi, Cilt 20: Sayı 4: 169-172:2009
- Pelitli Gürlü, V. Açık göz küresi yaralanmalarında sonuç görme keskinliği üzerine etkisi olan faktörler:2007.
- Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. An introduction to logistic regression analysis and reporting. The Journal of Educational Research, 96 (1), 3-14., 2002.
- Powers Daniel A. ve Xie Yu Statistical Methods for Categorical Data Analysis, Academic Press, ABD. 2000.
- Sanfort, W.: Applied Linear Regression, University of Minnesota, 1985.
- Saraçbaşı, O.: Lojistik Regresyon, Lisansüstü eğitim için hazırlanmış ders notu, Hacettepe Üniversitesi Tıp Fakültesi Biyoistatistik Bilim Dalı, 1994.
- Sharma, S. Applied Multivariate Techniques, New York, 1996.
- Seven, Z, Değişken seçimi yöntemi olarak adımsal lojistik regresyon ile adımsal diskriminant analizinin karşılaştırılması, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi, 1997.
- SPSS'te Çözümleri İstatistik Yöntemler 2 http://www.baskent.edu.tr/~matemel/courses/deney_tasarimi.pdf adresinden 23 aralık 2015 tarihinde edinilmiştir.
- Sulock, M. An Application of Binary Logistic Regression to College Admissions Data. Masters of Sciences in Statistics, Montana State University, 2009.
- Şenoğlu, B. T.C. Osmangazi Üniversitesi Fen Edebiyat Fakültesi İstatistik Bölümü İstatistiksel Çözümleme Teknikleri I Regresyon Analizi, 2003.

- Tabachnick, B. G., Fidell, L. S. Using Multivariate Statistics (3rd ed.), New York, USA: Harper Collins College Publishers, 1996.
- Taşdelen, A.: Lojistik regresyon modelinin yaşam verilerine uygulanması, Ege Üniversitesi, 1997.
- Taşkın, V.: Lojistik regresyon yöntemi ile Angına pektoris tanısına çeşitli ölçütlerin katkısının incelenmesi, İstanbul üniversitesi, 1997.
- Tatlıdil, H. Uygulamalı çok değişkenli istatistiksel analiz. Ankara: Engin Yayınları, 1996.
- Tuzcu, M. Keklikoğlu, M.: Çeviri Editörleri. The Merck Manual, Cilt: 1-2, İstanbul 1995.
- Ünal, A.: Lojistik Regresyon Analizi ve Uygulaması, Gazi Üniversitesi, Yüksek Lisans.

EKLER

Ek-1 Hastalığın Derecesi

Hastalığın Derecesi	Frekans	Geçerli Yüzde	Birikimli Yüzde
1	88	49,7	49,7
2	68	38,4	88,1
3	21	11,9	100,0
Toplam	177	100,0	

Ek-2 Ortalama Tümör Boyutu

Tümör Boyutu	Frekans	Geçerli Yüzde	Birikimli Yüzde
<1	11	6,2	6,2
1-2	49	27,7	33,9
2-4	86	48,6	82,5
>4	31	17,5	100,0
Toplam	177	100,0	

Ek-3 Miyometrial İnvazyon

Miyometrial İnvazyon	Frekans	Geçerli Yüzde	Birikimli Yüzde
Küçük	149	84,2	84,2
Büyük	28	15,8	100,0
Toplam	177	100,0	

Ek-4 Doku Türü (Histoloji)

Doku Türü (histoloji)	Frekans	Geçerli Yüzde	Birikimli Yüzde
Endometrioid tip	145	81,9	81,9
Endometrioid olmayan	32	18,1	100,0
Toplam	177	100,0	

Ek-5 Servikal Stromal Tutulum

Servikal Stromal Tutulum	Frekans	Geçerli Yüzde	Birikimli Yüzde
Pozitif	135	76,3	76,3
Negatif	42	23,7	100,0
Toplam	177	100,0	

Ek-6 Adneksiyal Tutulum

Adneksiyal Tutulum	Frekans	Geçerli Yüzde	Birikimli Yüzde
Pozitif	159	89,8	89,8
Negatif	18	10,2	100,0
Toplam	177	100,0	

Ek-7 Lenfovasküler İnvazyon

Lenfovasküler İnvazyon	Frekans	Geçerli Yüzde	Birikimli Yüzde
Pozitif	148	83,6	83,6
Negatif	29	16,4	100,0
Toplam	177	100,0	

Ek-8 Pelvik Lenf Tutulumu

Pelvik Lenf Tutulumu	Frekans	Geçerli Yüzde	Birikimli Yüzde
Pozitif	155	87,6	87,6
Negatif	22	12,4	100,0
Toplam	177	100,0	

Ek-9 Paraaortik Lenf Nodu Tutulumu

Tutulum	Frekans	Geçerli Yüzde	Birikimli Yüzde
Pozitif	161	91,0	91,0
Negatif	16	9,0	100,0
Toplam	177	100,0	

Ek-10 Kanser Hastalığının Evresi

Evre	Frekans	Geçerli Yüzde	Birikimli Yüzde
Evre-1	118	66,7	66,7
Evre-2	26	14,7	81,4
Evre-3	24	13,6	94,9
Evre-4	9	5,1	100,0
Toplam	177	100,0	

Ek-11 Peritoneal Sistoloji

Peritoneal Sistoloji	Frekans	Geçerli Yüzde	Birikimli Yüzde
Nondiagnostik	27	15,3	15,3
Malignite Negatif	54	30,5	45,8
Malignite Pozitif	83	46,9	92,7
3	5	2,8	95,5
4	8	4,5	100,0
Toplam	177	100,0	

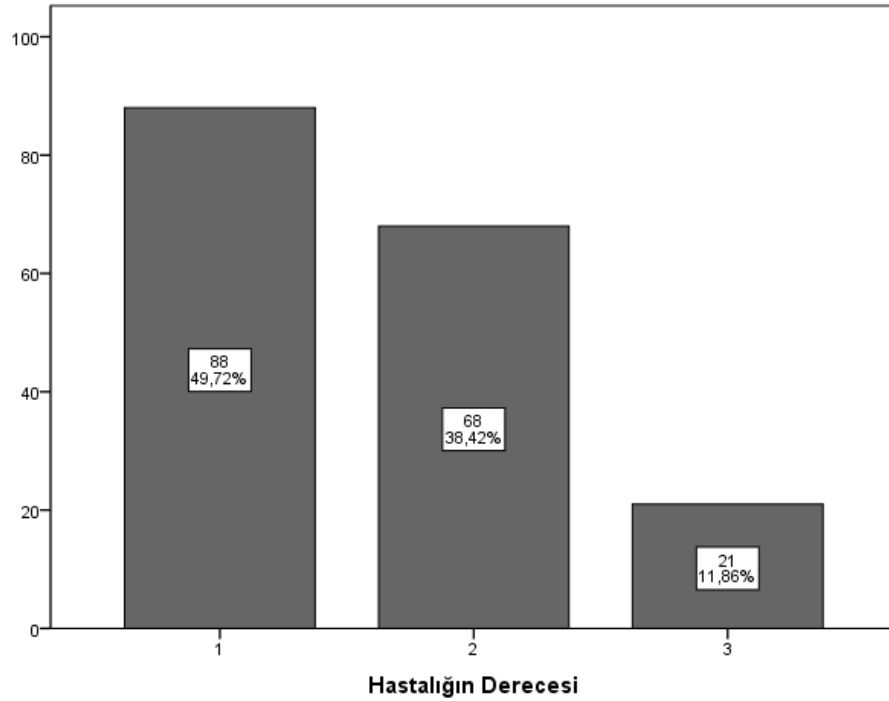
Ek-12 Lökosit Düzeyi

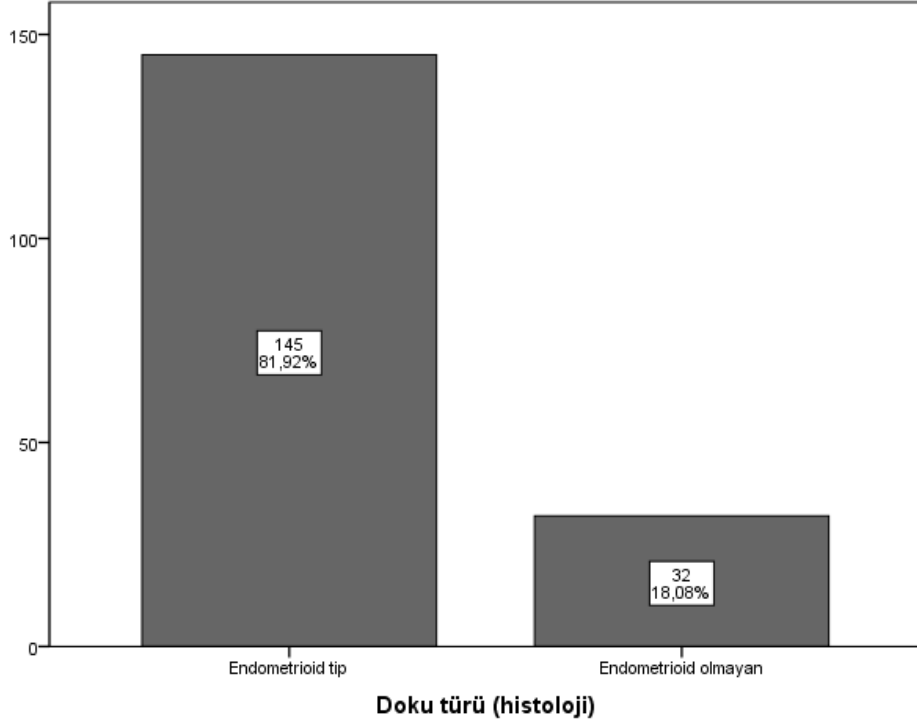
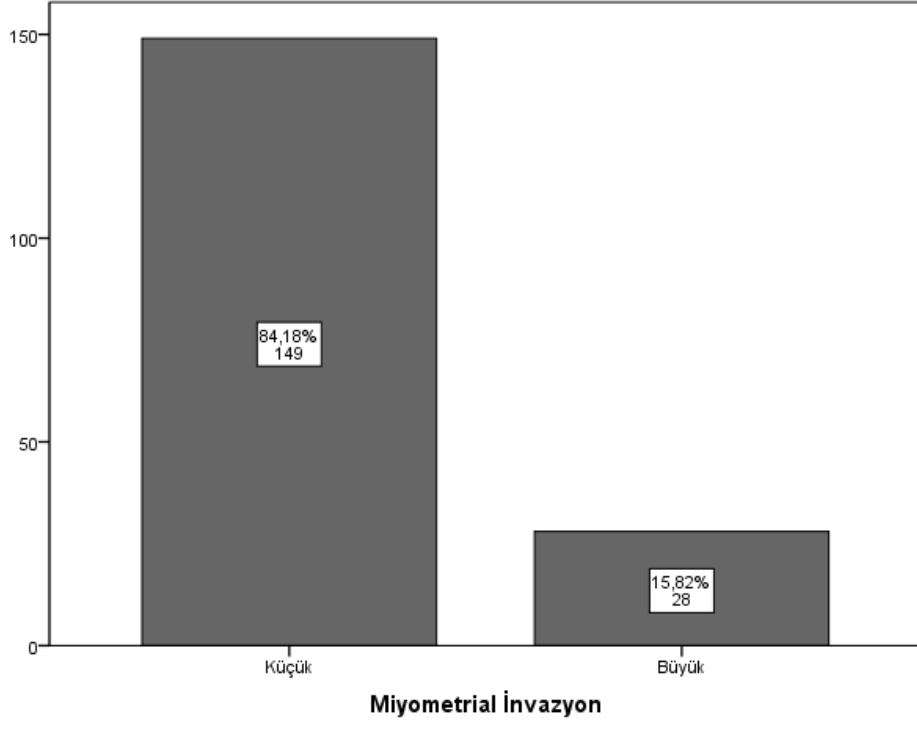
Lökosit Düzeyi	Frekans	Geçerli Yüzde	Birikimli Yüzde
Düşük (<10.000)	143	80,8	80,8
Yüksek (>=10.000)	34	19,2	100,0
Toplam	177	100,0	

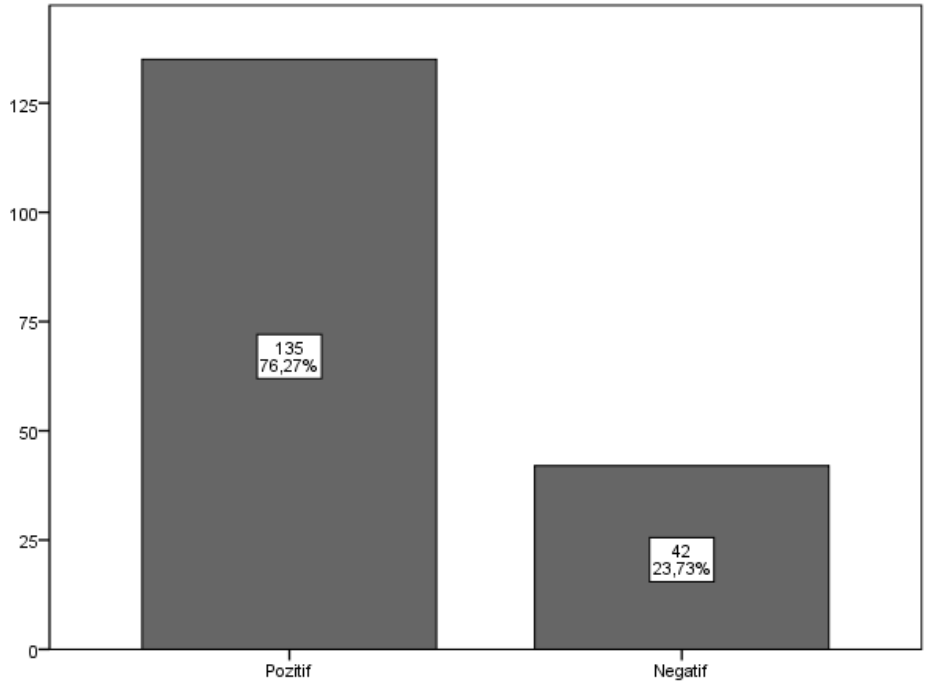
**Ek-13 Danışman Hekim Tarafından Risk Değerine Göre Yeniden Kodlanan
Operasyon Öncesi Lökosit Değerleri (WBC5)**

WBC5	Frekans	Geçerli Yüzde	Birikimli Yüzde
<7999	54	30,5	30,5
>8000	123	69,5	100,0
Toplam	177	100,0	

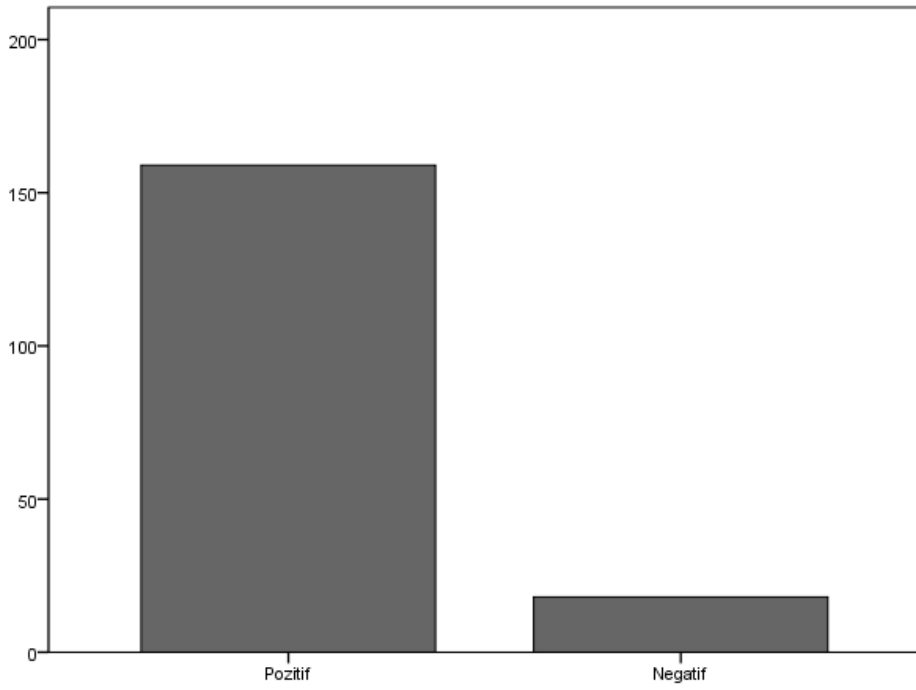
Ek-14 Tanımlayıcı Grafikler





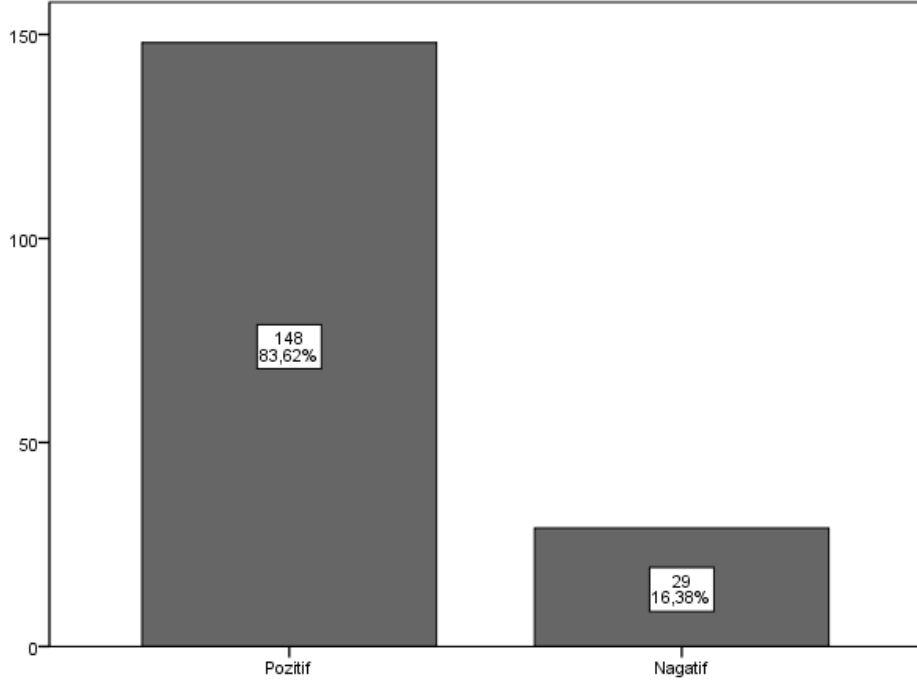


Servikal Stromal Tutulum

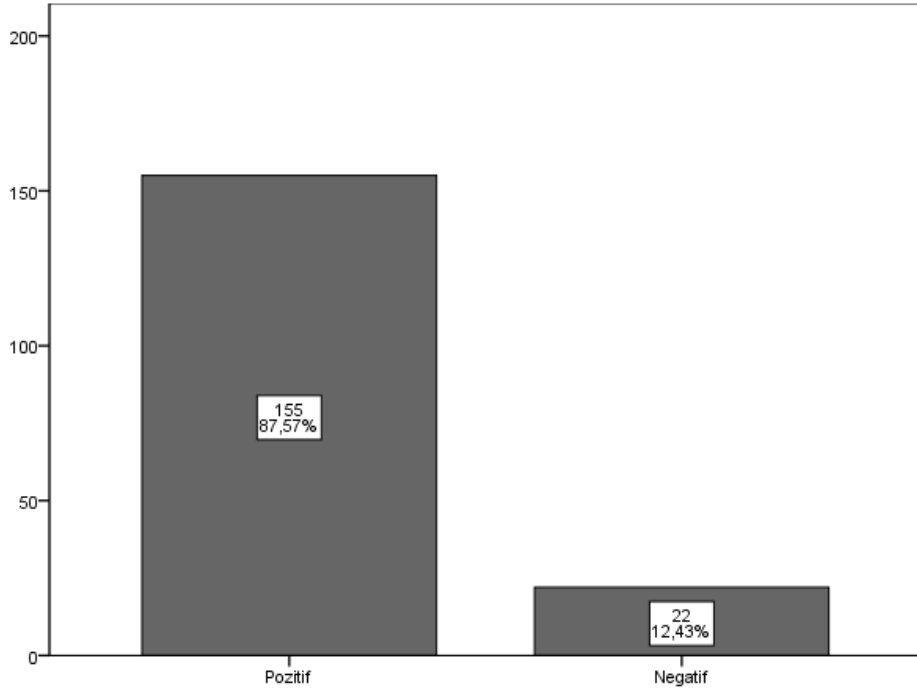


Adneksiyal Tutulum

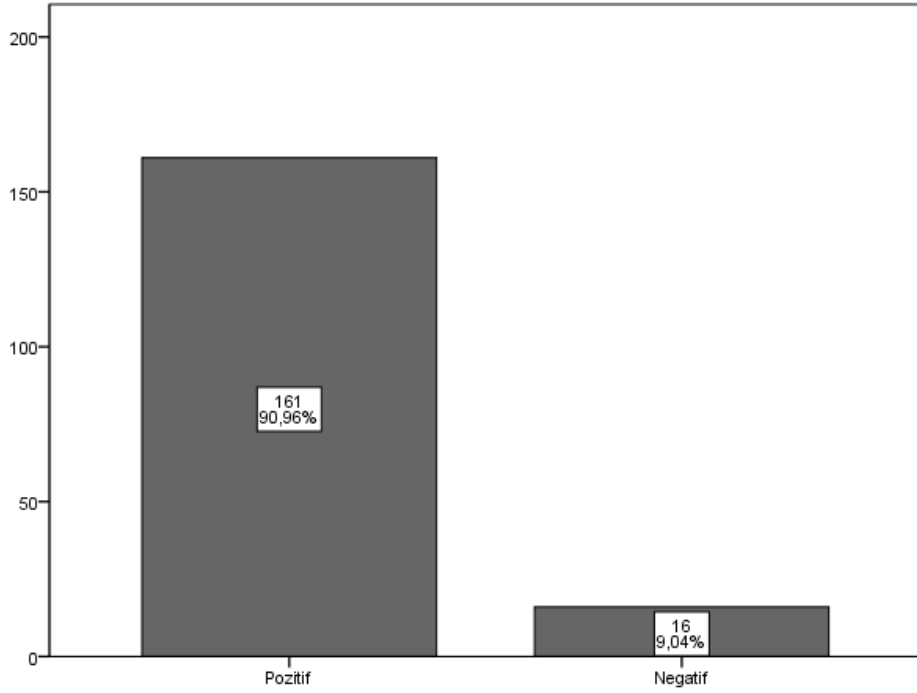
»



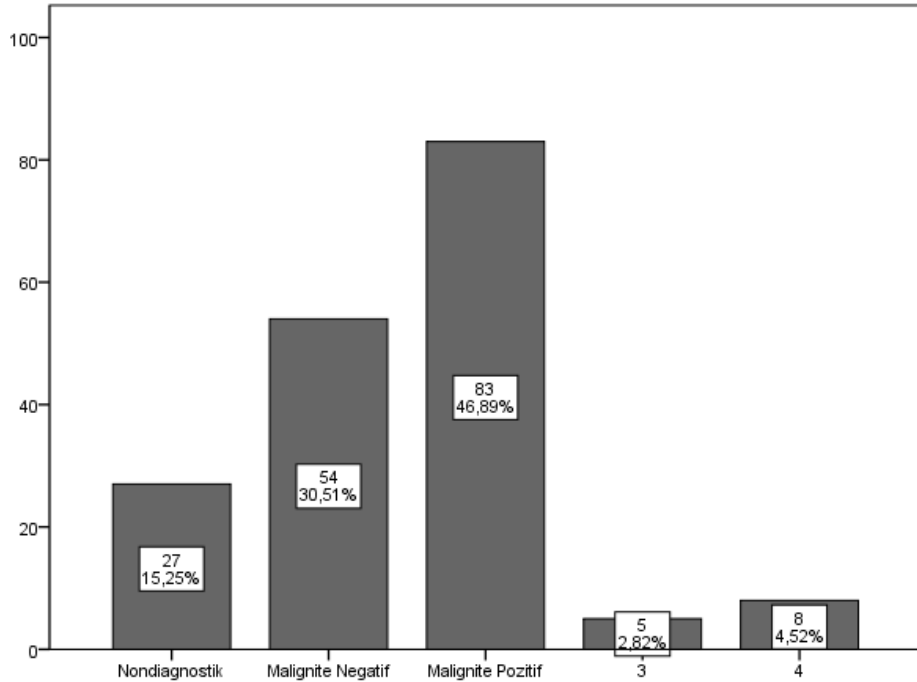
Lenfovasküler İnvazyon



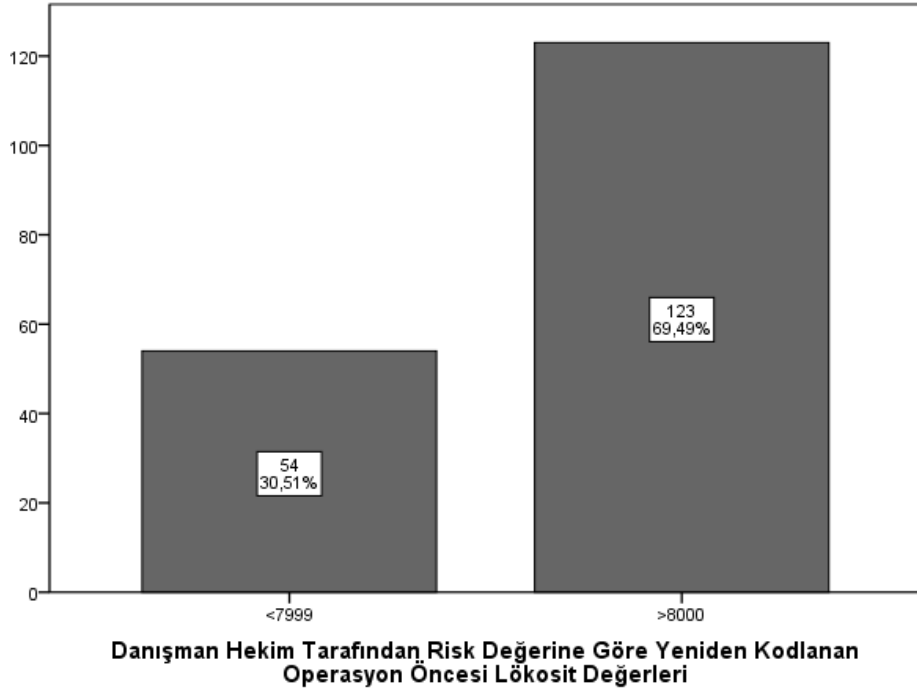
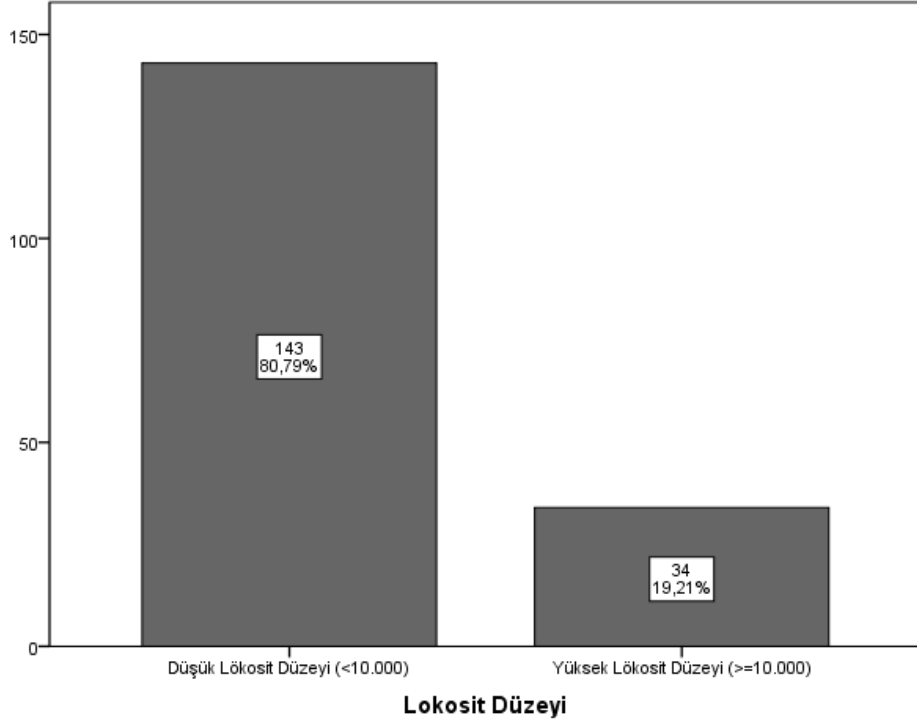
Pelvik Lenf Tutulumu



Paraaortik Lenf Nodu Tutulumu



Peritoneal Sistoloji



ÖZ GEÇMİŞ

Adı, Soyadı	Hilal Polat		
Doğum Yeri ve yılı	Göksun-1983		
Medeni Durumu	Evli		
Bildiği Yabancı diller ve Düzeyi	İngilizce-Orta		
Öğrenim Durumu	Başlama- Bitirme Yılı		Kurum Adı
Lisans	2002	2007	Fırat Üniversitesi
Yüksek Lisans	2012	2017	Recep Tayyip Erdoğan Üniversitesi
Çalıştığı Kurumlar		Başlama- Ayrılma Yılı	
1	ING BANK A.Ş.	2009	2010
2	Milli Eğitim Bakanlığı	2011	2012
3	Recep Tayyip Erdoğan Üniversitesi	2013	Devam
İletişim (eposta)	hilal.polat@erdogan.edu.tr		