

İKİ KATEGORİLİ LOJİSTİK REGRESYON ve DİSKRİMİNANT ANALİZİ

*İşletme Performans Tahmini
Üzerine Bir Uygulama*

Dr. Serap ŞAHİN

ÖZGÜR
YAYINLARI

İKİ KATEGORİLİ
LOJİSTİK REGRESYON ve
DİSKRİMİNANT ANALİZİ

*İşletme Performans Tahmini
Üzerine Bir Uygulama*

Dr. Serap ŞAHİN



Published by

Özgür Yayın-Dağıtım Co. Ltd.

Certificate Number: 45503

📍 15 Temmuz Mah. 148136. Sk. No: 9 Şehitkamil/Gaziantep

☎ +90.850 260 09 97

📞 +90.532 289 82 15

🌐 www.ozguryayinlari.com

✉ info@ozguryayinlari.com

İki Kategorili Lojistik Regresyon ve Diskriminant Analizi İşletme Performans Tahmini Üzerine Bir Uygulama

Binary Logistic Regression and Discriminant Analysis An Application on Business Performance Estimation

Dr. Scrap Şahin

Language: Turkish

Publication Date: 2023

Cover design by Mehmet Çakır

Cover design and image licensed under CC BY-NC 4.0

Print and digital versions typeset by Çizgi Medya Co. Ltd.

ISBN (PDF): 978-975-447-705-4

DOI: <https://doi.org/10.58830/ozgur.pub217>



This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc/4.0/>
This license allows for copying any part of the work for personal use, not commercial use, providing author attribution is clearly stated.

Suggested citation:

Şahin, S., (2023). *İki Kategorili Lojistik Regresyon ve Diskriminant Analizi İşletme Performans Tahmini Üzerine Bir Uygulama*. Özgür Publications.

DOI: <https://doi.org/10.58830/ozgur.pub217>. License: CC-BY-NC 4.0

“İstanbul Üniversitesi Sosyal Bilimler Enstitüsü’nce kabul edilmiş Doktora tezinden türetilmiştir.”

The full text of this book has been peer-reviewed to ensure high academic standards. For full review policies, see <https://www.ozguryayinlari.com/>



ÖN SÖZ

Bu kitap 2011 yılında İstanbul Üniversitesi Sosyal Bilimler Enstitüsü'nce kabul edilmiş Doktora tezinden türetilmiştir.

Tez çalışmamı kitap olarak yayınlama kararı almamdaki en önemli neden, çeşitli bilim dallarında araştırmacılar tarafından yürütülen çalışmalarda, bağımlı değişkenin nitel (kategorik) olduğu durumlarla karşılaşılması ve bu tür problemlerin analizinde, bağımlı ve bağımsız değişkenler arasındaki ilişkinin klasik regresyon modeli ile kurulmasının mümkün olmaması ve bu noktada problemlerin modellenebilmesi için alternatif yöntemlere ihtiyaç duyulmasıdır. Bu kitap ile amaçlanan, araştırmacılara bağımlı değişkenin kategorik olduğu problemlerle karşılaştıklarında, değişkenler arasındaki ilişkileri kurabilmek için kullanılabilecekleri birçok uygun istatistik yöntem arasından, akademik çalışmalarda genellikle en çok tercih edilen ve bu tür modellerin analizinde kullanılması uygun olan İki Kategorili Lojistik Regresyon Analizi ile İki Kategorili Diskriminant Analizi hakkında, faydalanabilecekleri ve çalışmalarına ışık tutabileceğine inandığım bilgileri sunmaktır.

Bu kitabın ilk bölümü, iki alt başlıktan oluşmaktadır. İlk alt başlıkta, Genelleştirilmiş Doğrusal Modellerden ve Doğrusal Olasılık Modeli ile Lojit ve Probit modellerden kısaca bahsedilmiştir. İkinci alt başlıkta, Lojistik Regresyon Analizi için modelin varsayımları, model kurma stratejileri, parametre tahmin yöntemi, anlamlılık ve uyum iyiliği testleri ile analiz sonucunda elde edilen model parametrelerinin nasıl yorumlanması gerektiğiyle ilgili teorik bilgilere yer verilmiştir.

İkinci bölümde, Diskriminant Analizi, analizin varsayımları, diskriminant fonksiyonun elde edilmesi, elde edilen fonksiyonun test edilmesi ve yorumlanması ile ilgili genel bilgiler paylaşılmıştır. Ayrıca bu analizden ayırıcı gücünün ve dışsal geçerliliğin nasıl ele alınabileceğiyle ilgili teorik bilgiler de bu bölüm içinde yer almaktadır.

Üçüncü bölümde, işletme performansını etkileyebilecek faktörlerin araştırılması başlığı altında, işletme performans kriterleri ve işletme performansını etkileyebilecek faktörler ile finansal ve finansal olmayan oranlar hakkında kısa bilgiler verilmiştir.

Kitabın son bölümü olan dördüncü bölümde, İstanbul Menkul Kıymetler Borsası'nda imalat sanayi sektöründe faaliyet gösteren işletmelerin 2007 yılı verilerinden elde edilen, performans göstergelerini etkileyen faktörlerin saptanması amacıyla, modellerin kurulma süreci ile diskriminant ve lojistik regresyon analiz sonuçları paylaşılmıştır. Ayrıca bu bölümde yer alan analiz sonuçlarıyla elde edilen çıktıların hem yorumlanma şekli açısından hem de elde edilen sonuçlar doğrultusunda yapılan çıkarımlar açısından bu kitabın, çalışmalarında bu yöntemleri kullanacak araştırmacılar için faydalanabilecekleri bir kaynak olacağını düşünmektedir.

Kitabımın, bu konularda çalışan akademisyenler, lisansüstü ve lisans öğrencileri ile işletme alanındaki karar vericiler tarafından hem teorik hem de uygulama açısından başvuru kaynağı olarak görülebileceği inancıyla, tüm okuyuculara katkı sağlamasını ümit ederim.

Serap ŞAHİN

Ağustos, 2023 – Ankara

İçindekiler

Ön Söz	iii
Kısaltmalar Listesi	vii
GİRİŞ	1
Bölüm 1	
<hr/>	
1. Lojistik Regresyon Analizi	3
Genelleştirilmiş Doğrusal Modeller	4
İki Kategorili Lojistik Regresyon Modelleri	15
Lojistik Regresyon Modellerinin Örnek Büyüklüğü	48
Bölüm 2	
<hr/>	
2. Diskriminant Analizi	53
İki Kategorili Diskriminant Analizi	54
Diskriminant Fonksiyonunun Dışsal Geçerliliği	74
Bölüm 3	
<hr/>	
3. İşletme Performansını Etkileyebilecek Faktörlerin Araştırılması	79
İşletme Performans Kriterleri	79
İşletme Performansını Etkileyebilecek Faktörler	91
Bölüm 4	
<hr/>	
4. İşletme Performansının Tahminine Yönelik İmalat Sanayi Sektörü Uygulaması	107
Araştırmanın Amacı ve Konusu	107
Araştırmanın Veri Seti	108
Araştırma Değişkenlerinin Seçimi	108
Verilerin Analize Hazırlanması	112

Finansal Performans Tahmin Modellerinin Uygulanması	113
Uygulama Sonuçları ve Yorumları	138
KAYNAKÇA	149
EKLER	157
ÖZGEÇMİŞ	181

Kısaltmalar Listesi

AIC	Akaike Information Criteria – Akaike Bilgi Kriteri
AR-GE	Araştırma Geliştirme
BIC	Bayesian Information Criteria – Bayes Bilgi Kriteri
CI	Conditional Index – Koşul İndeksi
D	Deviance – Sapma
DD	Defter Değeri
EKK	En Küçük Kareler
GLM	Generalized Linear Models – Genelleştirilmiş Doğrusal Modeller
IIA	Independence of Irrelavent Alternatives – İlgisiz Alternatiflerin Bağımsızlığı
İMKB	İstanbul Menkul Kıymetler Borsası
İSO	İstanbul Sanayi Odası
K.Vd.	Kısa Vadeli
LL	Log-Likelihood – Log-olabilirlik
LR	Likelihood Ratio – Olabilirlik Oranı
LRM	Linear Regression Model – Doğrusal Regresyon Modeli
MLE	Maksimum Likelihood Estimation – En çok Olabilirlik Tahmini
MNL	Multinomial Logit – Çok Kategorili Lojit
MTT test	McFadden, Train ve Tye tarafından sunulan test
OR	Odds Ratio – Odds Oranı
PD	Piyasa Değeri
ROC Eğrisi	Receiver Operating Characteristic Curves
SH testi	Small ve Hsiao testi
U.Vd.	Uzun Vadeli
VIF	Variance Inflation Factor – Varyans Artış Faktörü

GİRİŞ

İşletmelerin rekabet güçlerine bağlı olarak performanslarının sürekliliğinin sağlanmasının önemi özellikle yatırımcılar, ortaklar ve işletme yöneticileri açısından her geçen gün artmaktadır. Bu nedenle hem yatırımcı ve ortakların kararlarında hem de ülke ekonomilerinin verimliliğinin ve etkinliğinin belirlenmesinde işletme performanslarını etkileyen faktörlerin en iyi şekilde belirlenmesi amacıyla araştırmacılar tarafından birçok çalışma yapılmıştır. Bu konu ile ilgili yapılan çalışmalarda performans göstergelerinin ayrı ayrı incelenmesine karşın işletmelerin performanslarının birden çok ve farklı kriterler ile ölçülmesi ve değerlendirmesinin daha uygun olacağı düşünülmektedir.

Çok değişkenli istatistik yöntemlerden olan Lojistik Regresyon ile Diskriminant Analizi'nin ele alındığı ve bu yöntemler sonucunda işletme performansını etkileyen faktörlerin belirlenmesi amacıyla kurulan modellerin incelendiği bu çalışma dört bölümden oluşmaktadır.

Çalışmada öncelikle işletmelerin performanslarını çok boyutlu değerlendirebilmek için performans göstergelerinin farklı kombinasyonlarından meydana gelen çeşitli göstergelere göre sınıflandırılmış olan kategorik bağımlı değişkenleri modellemeye en uygun olan çok değişkenli istatistik yöntemlerden Lojistik Regresyon ve Diskriminant Analizi teorik olarak incelenmiştir.

Birinci bölümde, bağımlı değişkenin nitel olduğu durumlarda kullanılan modeller hakkında genel bilgi verilmesi açısından öncelikle Genelleştirilmiş Doğrusal Modeller başlığı altında Doğrusal Olasılık Modeli, Lojit ve Probit Modelleri'ne kısaca değinilmiştir. Çalışmanın uygulama bölümünde de

kullanılan Lojistik Regresyon Analizi'nin iki kategorili olması nedeniyle çalışmada İki Kategorili Lojistik Regresyon konusu ayrıntılı olarak ele alınmıştır.

İkinci bölümde, bağımlı değişkenin nitel, bağımsız değişkenlerin genellikle nicel ve sayıca fazla olmamak koşuluyla nitel de olduğu durumlar için kullanılabilen çok değişkenli istatistik yöntemlerden Diskriminant Analizi, Lojistik Regresyon Analizi'nde olduğu gibi yine çalışmanın uygulama kısmında bağımlı değişken iki kategorili ele alındığı için bu bölümde İki Kategorili Diskriminant Analizi ayrıntılı olarak incelenmiştir.

İşletmelerin performans göstergeleri ve bu göstergeleri etkileyebilecek faktörler hakkında bilgilere ise üçüncü bölümde yer verilmiştir. Son bölümde ise tezin ampirik çalışması yer almaktadır. İstanbul Menkul Kıymetler Borsası'nda imalat sanayi sektöründe faaliyet gösteren işletmeler performansları açısından çeşitli göstergelere göre sınıflandırıldıktan sonra hem Lojistik Regresyon hem de Diskriminant Analizi ile söz konusu göstergelerin hangi faktörlere bağımlı olduğu araştırılmıştır.

BÖLÜM 1

LOJİSTİK REGRESYON ANALİZİ

Çeşitli bilim dallarında bağımlı değişkenin nicel (nümerik) olduğu durumlarda kullanılan en temel istatistik yöntem regresyon analizidir. Ancak bağımlı değişkenin nitel olduğu durumlarda klasik regresyon modeli kullanılamamaktadır. Bunun nedeni, sosyal bilimlerdeki araştırmaların birçoğunda tutumların (tavırların), davranışların, karakteristik özelliklerin, kararların ve olayların, özünde sürekli ya da süreksiz olması, bir başka ifadeyle; kesikli, isimsel, sıralı ve sürekli olmayan yollarla ölçülebilir olmasıdır. Bu tür değişkenlerin analizi için birçok uygun istatistik model söz konusudur. Bu modellerden bazıları, “İki Kategorili Veri Analizi (Binary Data Analysis)”, “Sıralı Veri Analizi (Ordinal Data Analysis)”, Kategorik Veri Analizi (Categorical Data Analysis)”, “Kesikli Tercih Analizi (Discrete Choice Analysis)” ya da Lojit veya Probit gibi özel modellerdir. Birbiriyle ilişkili bu istatistik yöntemlerin ortak özelliği, olayların olasılıklarını yani olayın ne kadar sıklıkla ortaya çıktığını modellemeleridir. Bu nedenle, olay olasılığını analiz eden tüm istatistik yöntemlere genel olarak olasılık modelleri adı verilmektedir¹. Çalışmada olasılık modellerinden, İki Kategorili Lojit (Binary Logit), Probit ve Diskriminant modelleri ele alınacaktır.

Lojistik Regresyon Analizi, bağımlı değişkenlerinin* kategorik olarak ikili, üçlü ve çoklu kategorilerde gözleendiği durumlarda, açıklayıcı değişkenlerin

¹Tim Futing Liao, **Interpreting Probability Models Logit Probit and Other Generalized Linear Models**, Sage University Paper, 1994, pp. 1–2.

* Doğrusal Regresyon Modeli'nde “bağımlı değişken” (dependent variable) olarak ifade edilen kavram Lojistik Regresyon Modeli içi literatürde “cevap değişkeni” (response variable) veya “sonuç değişkeni” (outcome variable) olarak ifade edilmektedir.

(bağımsız değişken) bağımlı değişken ile sebep sonuç ilişkisini belirlemede yararlanılan bir yöntemdir. Ayrıca, açıklayıcı değişkenlere göre bağımlı değişkeninin beklenen değerinin olasılık olarak elde edildiği, sınıflama ve atama işlemi yapmaya yardımcı olan bir regresyon modelidir. Lojistik Regresyon Analizinin en önemli amacı, istatistikteki diğer model kurma yöntemlerinde olduğu gibi bağımsız değişkenlerle bağımlı değişken arasındaki ilişkinin tanımlanmasında en uygun modelin belirlenmesidir².

Bu bölümde Lojistik Regresyon Analizine başlamadan önce Genelleştirilmiş Doğrusal Modeller başlığı altında Doğrusal Olasılık Modeli, Lojit ve Probit modelleri hakkında genel bilgi verilecektir. Bu modeller bağımlı değişkenin nitel olduğu durumlarda kullanılan modellerdendir.

1.1 Genelleştirilmiş Doğrusal Modeller

Regresyon ve Varyans Analizi'nde kullanılan modeller, “Genelleştirilmiş Doğrusal Modeller” (Generalized Linear Models – GLM) olarak bilinen geniş bir model grubunun üyeleridir. Genelleştirilmiş Doğrusal Modeller 1972 yılında ilk olarak Nelder ve Wedderburn tarafından önerilmiş ve 1989 yılında McCullagh ve Nelder tarafından daha kapsamlı şekilde ele alınmıştır. Lojistik regresyon ise GLM'in özel bir durumu olarak kabul edilmektedir³.

GLM'in temel yaklaşımı, fonksiyonları doğrusal modelin sistematik bileşenleri olarak ifade edecek şekilde dönüştürerek Standart Doğrusal Modele oldukça benzeyen modeller elde etmektir⁴. Farkı ise doğrusal ve aralık ölçekli olmayan bağımlı değişkenlerin açıklayıcı değişkenlerle ilişkilerini göstermek bakımından uygulanabilirliğidir.

² David W. Hosmer, Stanley Lemeshow, **Applied Logistic Regression**, 2nd Edition, USA, John Wiley & Sons, 2000, p.1.

³ Ronald Christensen, **Log-Linear Models and Logistic Regression**, 2nd Edition, Springer, 1997, p. 297.

⁴ Alan Agresti, **An Introduction to Categorical Data Analysis**, John Wiley & Sons, 1996, p. 72.

GLM; verilerin sistematik bileşenleri, eşit varyanslılık ve asimptotik normallik varsayımlarını gerektirmeyen, bağımlı değişkenle açıklayıcı değişkenler arasındaki ilişkiyi tanımlamak için “bağ fonksiyonunun (link function)” kullanıldığı, gözlemlerin birbirinden bağımsız olması varsayımının ise sağlanması gereken modellerdir⁵.

GLM hakkında tartışmaya, üstel dağılım grubuyla başlanır. Bunun nedeni, GLM’in yalnızca bu kategoriye uyan parametrik formlara uygulanabilirliğidir. Bernoulli, Binom, Geometrik, Negatif Binom, Poisson, Üstel, Normal, Beta ve Gamma dağılımlarının bazıları üstel gruba aittir⁶. Burada bahsi geçen Gamma dağılımının özel iki durumu olan, Gammanın $(\rho/2, \frac{1}{2})$ parametreleri ile ρ serbestlik dereceli Ki-kare dağılımı ve Gammanın $(1, \beta)$ parametreleriyle üstel dağılıma uyan formudur⁷. İki önemli olasılık yoğunluk fonksiyonu olan Student-t ve Uniform dağılım ise üstel dağılım grubunun üyesi değildir.⁸

GLM teorisi, değişkenin dağılımının genellikle tek parametreliliğiyle üstel gruptan olmasını gerektirmektedir. Bu modeller içinde y rastgele ya da tahmini kısımdır. Bu öge bazen “olasılık dağılımı” ya da “bağımlı değişken dağılımı” olarak adlandırılmakla birlikte bazı kaynaklarda “rassal öge”⁹ olarak adlandırılmaktadır. Bağımlı değişkenin olasılık dağılımı üstel dağılım grubundan olmakta ve genel ifade şekli ile aşağıdaki gibi yazılmaktadır.

$$f(y|\theta, \phi; w) = \exp\left\{\frac{w}{\phi} [y\theta - r(\theta)]\right\} h(\phi, y, w) \quad 1.1$$

θ , ϕ ve w sayısal değerleri ifade etmektedir. Varsayıma göre w , sabit bilinen bir değerdir. Bu fonksiyon içinde ϕ farklı bir rol üstlenmektedir. Bilinmeyen sabit

⁵ Jeff Gill, **Generalized Linear Models**, USA Sage University Paper, 2001, p. 2.

⁶ Ronald Christensen, **a.g.e.**, pp. 299–304.

⁷ John E. Freund, **Mathematical Statistics**, Fifth Edition, Prentice Hall International Editions, p. 224

⁸ Jeff Gill, **a.g.e.**, p.10.

⁹ Joseph F. Hair, William C. Black, Barry J. Babin, Rolph E. Anderson and Ronald L. Tatham, **Multivariate Data Analysis**, Sixth Edition, Pearson, 2006, p. 411.

gibi davranılmakla birlikte ϕ “sorunlu parametre (nuisance parameter)” veya “dağılım parametresi (dispersion parameter)” olarak adlandırılmaktadır. Aslında ϕ , çeşitli özel durumların tanımlanması için uygun pozitif bir değerdir.¹⁰ Bazı kaynaklarda $w/\phi = \alpha(\phi)$ ölçek faktörü olarak ifade edilmektedir. Normal dağılım ve Gamma dağılımı için durum biraz farklıdır. Bu dağılımların sırası ile $\theta = (\mu, \sigma^2)$ ve $\theta = (\alpha, \beta)$ olmak üzere iki parametresi bulunmaktadır. Bu durumda, Normal dağılım için ϕ parametresi normal bir dağılımın varyansı (σ^2) ya da Gamma dağılımının p parametresi gibi özel bir tip sorunlu parametreye işaret edebilir¹¹.

Tek parametrelili üstel dağılım grubu için yukarıdaki eşitliğin en genel ifadesi $w/\phi = 1$ olacak şekilde,

$$f(y|\theta) = \exp[y\theta - r(\theta) + h(y)] \quad 1.2$$

biçiminde gösterilmektedir. Eşitlik 1.2’de $h(y)$ ile $\log[h(\phi, y, w)]$ ifade edilmektedir. Eşitliğin sağ tarafındaki y ve θ ’nin çarpımından oluşan ilk kısım etkileşim bileşeni, diğer kısım ise katkı bileşeni olarak adlandırılmaktadır. Üstel dağılım grubu ile kanonik formun* kullanılması verinin dağılımı hakkında doğrudan bilgi vermesi bakımından çalışmayı daha da kolaylaştırmaktadır¹².

1989 yılında McCullagh ve Nelder, katkı bileşeni olarak adlandırılan ikinci kısım içindeki $r(\theta)$ ’yi “kümülatif fonksiyon (cumulant function)”¹³ olarak adlandırmıştır. Bu değer ile ilgili en önemli olgu, $r(\theta)$ ’nin dağılım momentlerinin hesaplanmasında anahtar rol oynamasıdır, $r(\theta)$ ’nin θ ’ya göre

¹⁰ Ronald Christensen, **a.g.e.**, p. 300.

¹¹ J. A. Nelder, R. W. M. Wedderburn, “Generalized Linear Models”, **Journal of the Royal Statistical Society Series A**, Vol:135, No:3, 1972, pp. 370–384.

* “Kanonik form” ile kastedilen; matematikte ilgilenilen olgunun belirli bir şekilde göre gösterilmesidir. Başka bir ifadeyle ele alınan olgu için standart, genel kabul görmüş formlara kanonik form denilmektedir. Bu açıklamanın ışığında bu çalışmada kanonik form dağılımı, Eşitlik 1.2’de gösterilen ifadedir.

¹² Jeff Gill, **a.g.e.**, p. 11.

¹³ P. McCullagh, J. A. Nelder, **Generalized Linear Models**, 2nd Edition, Chapman & Hall, 1989, p. 30.

birinci mertebeden türevi bağımlı değişkenin beklenen değerini ikinci mertebeden türevi de varyansını vermektedir. Ayrıca, θ 'nın bu formu, orijinal formu ile parametre değeri olarak ifade edilen formu arasında kanonik bir bağ oluşturmaktadır¹⁴. Başka bir ifade ile modellerde, bağımlı değişken açıklayıcı değişkenlere doğrusal bir yapı ile bağlıdır. Bu yapıya da *bağ fonksiyonu* adını almaktadır.

GLM'in gelişiminde, beklenen değer (ilk moment) teorik olarak önemli rol oynamaktadır. Eşitlik 1.2'deki dağılımlar için y değişkeninin beklenen değeri θ 'ya bağlı ancak ϕ 'ye bağlı değildir. Herhangi bir fonksiyon için aşağıdaki eşitlik geçerlidir.¹⁵

$$1 = \int f(y|\theta, \phi; w)dy \quad 1.3$$

y değişkeninin kesikli olduğu durumda integral yerine toplam (sigma) kullanılır.

Her iki tarafın θ 'ya göre türevi alındığında;

$$0 = \int f'(y|\theta, \phi; w)dy. \quad 1.4$$

f' ; f 'nin θ 'ya göre birinci metbeden türevidir ve Eşitlik 1.1'deki fonksiyon Eşitlik 1.4'de yerine konulduğunda Eşitlik 1.4'ün aşağıdaki şekle dönüştüğü görülmektedir.

$$0 = \frac{w}{\phi} \int (y - r'(\theta))f(y|\theta, \phi; w)dy \quad 1.5$$

Buradan basit matematik işlemler yapıldığında y 'nin beklenen değeri,

$$E(y) \equiv \mu = \frac{\partial}{\partial \theta} r(\theta) \quad 1.6$$

¹⁴ Jeff Gill, **a.g.e.**, p. 12.

¹⁵ **A.e.**, p. 300.

olarak elde edilmektedir¹⁶. r' ; r' 'nin θ 'ya göre birinci mertebeden türevini göstermek üzere, $r'(\theta)$ tersi alınabilen fonksiyondur. Bu nedenle θ ayrıca beklenen değerin de bir fonksiyonudur, $\theta = (r')^{-1}(\mu)$. Eşitlik 1.1'deki formun dağılımlar için doğrusal yapısı genellikle, $\theta = x'\beta$ tarafından β bilinmeyen parametreler vektörü ve x bilinen ve sabit değer olarak en doğal hali ile tanımlanmaktadır. Aslında, GLM'in analizi, doğrusal yapının beklenen değerinin daha genel bir fonksiyonu ile gösterildiğinde başarılı sonuca ulaşılabilir.

Genelleştirilmiş Doğrusal Model $i = 1, 2, \dots, n$ olacak şekilde birbirinden bağımsız gözlemlerden (y_i) meydana geliyorsa,

$$\begin{aligned} y_i &\sim f(y_i | \theta_i, \phi, w_i) \\ E(y_i) &\equiv \mu_i \\ g(\mu_i) &= x_i' \beta \end{aligned}$$

olarak ifade edilmektedir. Eğer $g(\mu_i) = \theta_i$ ise model, Kanonik (Standart) Genelleştirilmiş Doğrusal Modeldir. Başka bir deyişle kanonik bir model $g(.) \equiv (r')^{-1}(.)$ demektir¹⁷.

GLM'in çeşitli bileşenleri için bazı özel ifadeler kullanılmaktadır. GLM içinde açıklayıcı değişkenlerin tanımlandığı kısım “sistemik öge”¹⁸ olarak adlandırılmakla birlikte doğrusal yapı $x'\beta$, *doğrusal tahminleyen*¹⁹ olarak adlandırılan $g(.)$ fonksiyonu ile ortalama (beklenen değer) ve doğrusal tahminleyen arasındaki ilişkiyi tanımlayan bir başka öge de $g(\mu) = x'\beta$, *bağ*

¹⁶ Nicholas T. Longford, **Handbook of Statistical Modeling for the Social and Behavioral Sciences**, Plenum Press, Edited by Gerhard Arminger, Clifford C. Clogg and Michael E. Sobel, 1995, p. 551.

¹⁷ P. McCullagh, J. A. Nelder, **a.g.e.**, pp. 28–29.

¹⁸ **A.e.**, p. 27.

¹⁹ Alvin C. Rencher, G. Bruce Schaalje, **Linear Models in Statistics**, A John Wiley & Sons, 2008, p. 513.

*fonsiyonu*²⁰ olarak bilinmektedir. Eğer $g(\mu) = \theta$ ise $g(\cdot)$ fonksiyonu *kanonik (standart) bağı fonksiyonu* olarak anılır²¹.

Dağılım parametresi tahmini ve modelin uygunluk testleri için tek parametrelili üstel dağılım varyansının hesaplanması gereklidir. Eşitlik 1.1'in θ 'ya göre ikinci mertebeden türevi alındığında, $(0 = \int f''(y|\theta, \phi; w)dy)$;

$$0 = \frac{w}{\phi} \int \frac{d[(y - r'(\theta))f(y|\theta, \phi; w)]}{d\theta} dy \quad 1.7$$

elde edilmekte ve gerekli işlemler yapıldığında Eşitlik 1.8'e ulaşılmaktadır.

$$0 = \frac{w^2}{\phi^2} \int (y - r'(\theta))^2 f(y|\theta, \phi; w) dy + \frac{w}{\phi} \int -r''(\theta) f(y|\theta, \phi; w) dy \quad 1.8$$

(Eşitliklerdeki “''” ikinci mertebeden türev için kullanılmaktadır.) Yukarıdaki eşitlik (1.8),

$$0 = [Var(y)w^2/\phi^2] - [r''(\theta)w/\phi] \quad 1.9$$

şeklini almaktadır ve $Var(y) = r''(\theta)\phi/w$.²²

Bazı yaygın olasılık fonksiyonlarının normalleştirilmiş katsayıları ve varyans fonksiyonları Tablo 1.1'de verilmekle birlikte çalışmada incelenen iki kategorili bağımlı değişkenlerin dağılım fonksiyonu olan Binom dağılımın üstel dağılım formu, beklenen değeri ve varyansı aşağıdaki gibi elde edilmektedir.

Binom dağılım, Bernoulli sürecinin özel bir hali olup, denemelerde birbirini engelleyen iki sonuçtan birinin gerçekleşmesiyle sonuçlanan sürecin çok kez yinelenmesiyle ilgilidir²³. Bu sürecin dağılımının önemli bir genellemesi olan

²⁰ Alan Agresti, **a.g.e.**, p. 72.

²¹ Ronald Christensen, **a.g.e.**, p. 301.

²² P. McCullagh, J. A. Nelder, **a.g.e.**, p. 29.

²³ Neyran Orhunbilge, **Tanımsal İstatistik Olasılık ve Olasılık Dağılımları**, Avcıol Basım, 2000, s. 192.

rassal bir denemenin birbirinden bağımsız olarak n kez yinelenmesi sonucunda ortaya çıkan başarı sayısı ile ilgilenen Binom dağılımın olasılık kütle fonksiyonu, üstel dağılım formunda şu şekilde ifade edilmektedir²⁴;

$$\begin{aligned} f(y|n, p) &= \binom{n}{y} p^y (1-p)^{n-y} \\ &= \exp \left[\log \binom{n}{y} + y \log(p) + (n-y) \log(1-p) \right] \quad \mathbf{1.10} \\ &= \exp \left[y \log \left(\frac{p}{1-p} \right) - (-n \log(1-p)) + \log \binom{n}{y} \right]. \end{aligned}$$

Binom dağılım için kanonik bağ $\theta = \log(p/(1-p))$ ile Lojit modeli oluşturmaktadır. Bu bağ fonksiyonu Lojit model başlığı altında incelenmiştir. Kanonik bağ fonksiyonunun tersi alınarak, $r(\theta)$ 'da yerine konulursa aşağıdaki eşitlik bulunmaktadır.

$$r(\theta) = [-n \log(1-p)] = n \log(1 + \exp(\theta)) \quad \mathbf{1.11}$$

Bilindiği gibi Binom dağılımın p ve n olmak üzere iki parametresi bulunmaktadır. Bunlardan p bilindiğinde ve üstel olasılık fonksiyonu n parametresi için geliştirilmek

istendiğinde, üstel olasılık fonksiyonu olarak yazılamamaktadır²⁵.

Binom olasılık fonksiyonunun ortalaması $r(\theta) = n \log(1 + \exp(\theta))$ ve kanonik bağ fonksiyonu $\theta = \log(p/(1-p))$ olarak kullanılırsa,²⁶

²⁴ Alvin C. Rencher, G. Bruce Schaalje, **a.g.e.**, p. 514.

²⁵ Jeff Gill, **a.g.e.**, p. 15.

²⁶ A.e., p. 24.

$$\begin{aligned}
\frac{\partial}{\partial \theta} r(\theta) &= \frac{\partial}{\partial \theta} (\text{nlog}(1 + \exp(\theta))) \\
&= \text{nlog}(1 + \exp(\theta))^{-1} \exp(\theta) \\
&= n \left(1 + \exp \left(\log \left(\frac{p}{1-p} \right) \right) \right)^{-1} \exp \left(\log \left(\frac{p}{1-p} \right) \right) \\
&= n(1-p) \left(\frac{p}{1-p} \right)
\end{aligned} \tag{1.12}$$

elde edilir. Buradan beklenen değer $E[y] = np$ olarak hesaplanmaktadır.

Binom olasılık kütle fonksiyonunun varyansı,²⁷

$$\begin{aligned}
\text{VAR}[Y] &= \frac{\phi}{w} \frac{\partial^2}{\partial \theta^2} r(\theta) \\
&= 1 \frac{\partial^2}{\partial \theta^2} (\text{nlog}(1 + \exp(\theta))) \\
&= \frac{\partial}{\partial \theta} \left(n(1 + \exp(\theta))^{-1} \exp(\theta) \right) \\
&= n \exp(\theta) \left[(1 + \exp(\theta))^{-1} - (1 + \exp(\theta))^{-2} \exp(\theta) \right] \Big|_{\theta = \log(p/(1-p))} \\
&= n \left(\frac{p}{1-p} \right) \left[\left(1 + \frac{p}{1-p} \right)^{-1} - \left(1 + \frac{p}{1-p} \right)^{-2} \frac{p}{1-p} \right] \\
&= np(1-p)
\end{aligned} \tag{1.13}$$

olarak gösterilmektedir.

Binom olasılık yoğunluk fonksiyonu için kanonik bağ iki şekilde ifade edilebilmektedir. Bu bağ fonksiyonlardan biri olan Lojit, üstel grup formunun tanımı gereği doğal olarak yer almaktadır. Diğer bağ fonksiyonu Probit ise Birikimli Standart Normal dağılıma dayanmaktadır. Genelde sosyal bilimler verisi için bu iki bağ fonksiyonu da (lojit ve probit) kullanılabilir ve bu bağ

²⁷ Jeff Gill, a.g.e., p. 27.

fonksiyonları benzer şekilde sonuçlar vermektedir.²⁸ Bu modellerin sonuçlarının benzerliği ile ilgili bilgi Başlık 1.1.3 içinde açıklanmaktadır.

Son olarak üstel dağılım grubunun üyesi olan dağılımlardan bazılarının normalleştirilmiş sabitleri ve varyans fonksiyonları aşağıdaki tablo ile özetlemektedir.²⁹

Tablo 1.1 Normalleştirilmiş Sabitler ve Varyans Fonksiyonları

Dağılım	$r(\theta)$	$\tau^2 = \frac{\partial^2}{\partial \theta^2} r(\theta)$
Poisson	$\exp(\theta)$	$\exp(\theta)$
Binom	$n \log(1 + \exp(\theta))$	$n \exp(1 + \exp(\theta))^{-2}$
Normal	$\frac{\theta^2}{2}$	1
Gamma	$-\log(-\theta)$	$\frac{1}{\theta^2}$
Negatif Binom	$s \log(1 - \exp(\theta))$	$s \exp(\theta)(1 - \exp(\theta))^{-2}$

1.1.1 Doğrusal Olasılık Modeli

Doğrusal Olasılık Modeli bağımlı değişkenin iki kategorili olduğu durumlarda uygulanan regresyon modelidir. Yapısal modelin gösterimi aşağıdaki gibidir.

$$y_i = x_{ik}\beta_k + \varepsilon_i \quad k = 1, 2, \dots, p \quad \mathbf{1.14}$$

x_i , i . gözlemin değerlerinden oluşan vektör; β parametre vektörü ε ise hata terimidir. y bağımlı değişkeni eğer olay gerçekleşiyorsa 1, gerçekleşmiyorsa 0 değerini almaktadır. Doğrusal Olasılık Modelini anlamak için öncelikle,

²⁸ Jeff Gill, **a.g.e.**, pp. 32–33.

²⁹ **A.e.**, p.29.

$E(y|x)$ 'in ne anlama geldiğini kavramak gerekmektedir. Olayın gerçekleşme olasılığı $E(y_i) = P(y_i = 1)$ 'dir. Koşullu değer olarak bahsedilen ise,

$$E(y_i|x_{ik}) = P(y_i = 1|x_{ik}) = x_{ik}\beta_k \text{ 'dır.} \quad \mathbf{1.15}$$

Buradan Doğrusal Olasılık Modelinin parametreleri şu şekilde yorumlanmaktadır: Model doğrusal olduğu sürece, x_k 'daki bir birimlik değişim, diğer açıklayıcı değişkenlerin sabit tutulduğu varsayıldığında, ilgilenilen durumun ortaya çıkma olasılığında β_k kadar beklenen değişime neden olmaktadır. Olasılıkları doğrusal olan modele *Doğrusal Olasılık Modeli* adı verilmektedir³⁰.

Doğrusal Olasılık Modelinde bağımlı değişkeninin iki kategorili olması durumunda parametre katsayılarının yorumu değişmemekle birlikte regresyon varsayımlarından bazılarını sağlaması gerekmemektedir. Bunlar; eşvaryanslılık, normallik ve fonksiyon formundaki bozukluklardır. Ayrıca Doğrusal Olasılık Modellerine alternatif olarak nitel bağımlı değişken için Lojit ve Probit modeller de vardır.

1.1.2 Lojit Model

Genel Log-Doğrusal Modelin özel bir formu olan Lojit Model, nitel bağımlı değişken analizlerini birleştiren önemli bir modeldir. “Lojit” in kelime anlamı kendi içinde saklı olup, araştırmada ilgilenilen durumun ortaya çıkma olasılığının çıkmama olasılığına oranını ifade eden odds'un doğal logaritmasının alınması anlamına gelmektedir. Literatürde “log-odds” olarak da yer almaktadır. Lojit Modelin genel gösterimi

$$\log \left[\frac{P(y = 1)}{1 - P(y = 1)} \right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad \mathbf{1.16}$$

³⁰ J. Scott Long, **Regression Models for Categorical and Limited Dependent Variables**, USA, Sage Publication, 1997, p. 36.

şeklindedir. Lojit dönüşüm, bağımsız değişkenler ile orjinal olasılıklarının arasındaki doğrusal olmayan ilişkileri doğrusallaştırmaktadır. Lojit Model, hatalarının ortalaması “0” ve varyansı “ $\pi^2/3$ ” olan standart lojistik dağılıma uymaktadır³¹. Bu nedenle Lojit Modeller veriye uygulanırken bağ fonksiyonu Lojit olur, $\theta = \log(p/(1 - p))$. Başka bir ifadeyle, bağımlı değişkenin lojit fonksiyonunun açıklayıcı değişkenlerin doğrusal bir bağlantısı olduğunu söyleyen bağ fonksiyonu Lojittir.

Lojistik Regresyon Analizi için kullanılan lojit dönüşümü göreceli basitliğinin avantajı nedeniyle en yaygın kullanılan modeldir ve Lojit Model olarak adlandırılmaktadır. Lojit Modeller, İki Kategorili Lojistik Regresyon Analizi bölümü içinde daha ayrıntılı olarak ele alınacaktır.

1.1.3 Probit Model

Bağımlı değişkenin iki kategorili olduğu ve Binom dağılıma uyduğu durumlarda sık kullanılan bir diğer model Probit Modeldir. Probit Modelden elde edilen dönüşüm, Probit Analizi için istenen olayın olasılıklarını, Lojistik dağılımda Lojit yerine Birikimli Normal dağılımdaki skorlara dönüştürmektedir³². Probit Modellerde GLM’deki bağımlı değişkeni açıklayıcı değişkenlere doğrusal bir yapı ile bağlayan fonksiyon “Probit”tir, $\theta = \Phi^{-1}(p)$. Olasılıklar için Probit Modeller,

$$P(y = 1) = 1 - \Phi\left(-\sum_{k=1}^p \beta'_k x_k\right) = \Phi\left(\sum_{k=1}^p \beta'_k x_k\right) \quad 1.17$$

şeklinde gösterilmektedir. Lojit Model odds’ların logaritmaları alınarak yapılan dönüşüm ve Lojistik Regresyon olarak ifade edilen, olayın olasılığını modellemek üzere iki temel şekilde ifade edilebilirken; Probit Model, ilgilenilen

³¹ J. Scott Long, a.g.e., p. 42.

³² Fred C. Pampel, **Logistic Regression A Primer**, USA, Sage Publication, 2000, pp. 54–55.

olayın olasılıklarının standart normal dağılım tablosu kullanılarak elde edilen Z değerlerinin bir Doğrusal Regresyonu olan θ ile ifade edilebilecek şekilde göstermektedir. Z değerleri ile olasılıklar arasında doğrusal olmayan ilişki söz konusudur. Olasılık değerleri uç değerlere (0 veya 1'e) yaklaştıkça Z değerlerindeki bir birimlik değişim, fonksiyonda daha az bir değişime neden olmaktadır³³.

Lojit Analizde modelin hatasının standart sapması $\sqrt{\pi^2/3}$ 'den yaklaşık olarak "1.814" iken Probit Analizde modelin hatasının standart sapması "1" dir. İki modele ait hata varyanslarının farklı olması Probit ve Lojit modellerin katsayılarının doğrudan karşılaştırılmayacağı anlamına gelmektedir. Lojit katsayıları, Probit katsayılarının yaklaşık olarak 1.8 katı olmaktadır.³⁴ Lojistik regresyon ve Probit katsayıları arasında fark modellerin hatalarının dağılımları olan Birikimli Lojistik ve Normal dağılımdan kaynaklanmaktadır.³⁵

1.2 İki Kategorili Lojistik Regresyon Modelleri

İki Kategorili Lojistik Regresyon Modellerinde, bağımlı değişkenler *nominal* veya *kategorik* olarak da adlandırılır ve ikili (binary) değer almaktadırlar. Bağımsız değişkenler ise nitel veya nicel değerler alabilirler. Bağımlı değişkenlerin ikili değer aldığı durumlarla sosyal bilimlerde oldukça yaygın olarak karşılaşmaktadır.³⁶

Bağımlı değişkenin iki kategorili olduğu durumlarda $E(y|x)$ modellenmesi için birçok dağılım fonksiyonu kullanılmakla birlikte, Lojistik fonksiyon en çok tercih edilen fonksiyondur³⁷. Cox ve Snell (1989) bu fonksiyonların bazıları üzerinde durmuş ve bunların içinde Lojistik fonksiyonun kullanılma nedeni

³³ Tim Futing Liao, **a.g.e.**, p.22.

³⁴ John H. Aldrich, Forrest D. Nelson, **Linear Probability Logit and Probit Models**, USA, Sage Publication, 1984, p. 41.

³⁵ Fred C. Pampel, **a.g.e.**, p. 60.

³⁶ J. Scott Long, **a.g.e.**, p. 34.

³⁷ David W. Hosmer, Stanley Lemeshow; **a.g.e.**, p. 6.

olarak öncelikli iki sebep öne sürmüştür. Bunlardan birincisi matematik olarak Lojistik fonksiyonun kullanımının daha esnek ve kolay olması, ikincisi ise açıklayıcı değişkenlerin daha anlamlı bir şekilde yorumlanabilmesi olarak gösterilmiştir³⁸.

Teknik olarak, Lojit ve Lojistik terimleri dönüşümler için verilen isimlerdir. Lojit dönüşümü, $[0, 1]$ arasında değer alan p değerlerine $\log [p/(1-p)]$ dönüşümünün uygulanması ile elde edilmektedir. Lojistik dönüşüm ise gerçek seri üzerinde bir x değeri alınıp bu değere $e^x/(1+e^x)$ dönüşümü uygulanması yoluyla bulunmaktadır. Lojit dönüşümü ve Lojistik dönüşüm birbirlerinden elde edilebilirler.³⁹

x_1, x_2, \dots, x_p açıklayıcı değişkenler biliniyorken y 'nin koşullu beklenen değeri basit gösterimiyle $P(y = 1|x_1, x_2, \dots, x_p) = E(y = 1|x_1, x_2, \dots, x_p)$ olarak ifade edilirse, Lojistik Regresyon Modeli basit formda;⁴⁰

$$P(y = 1|x_1, x_2, \dots, x_p) = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}} \quad \mathbf{1.18}$$

olarak ifade edilir.

Lojistik regresyonun temelini oluşturan $P(y = 1|x_1, x_2, \dots, x_p)$ dönüşümü *Lojit dönüşümü* olarak tanımlanmakta ve aşağıdaki şekilde ifade edilmektedir⁴¹.

$$g(x) = \log \left[\frac{P(y = 1|x_1, x_2, \dots, x_p)}{1 - P(y = 1|x_1, x_2, \dots, x_p)} \right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

Bu dönüşümün önemi, Doğrusal Regresyon Modelinin birçok önemli özelliğine sahip olmasından kaynaklanmaktadır. Lojit, $g(x)$, parametreleri bakımından doğrusaldır. Ayrıca p değeri arttıkça *lojit*(p) değeri de artar, p değeri $[0, 1]$

³⁸ D. R. Cox, E. J. Snell, **Analysis of Binary Data**, Second Edition, Chapman & Hall, 1989, p. 21.

³⁹ Ronald Christensen, **a.g.e.**, pp. 116–117.

⁴⁰ Fred C. Pampel, **a.g.e.**, p. 16.

⁴¹ Alan Agresti, **a.g.e.**, p. 103.

arasında değer alırken $lojit(p)$, $(-\infty, +\infty)$ aralığında değer almaktadır. Son olarak da gözlemlerin sınıflara atanmasında da kullanılan önemli bir özelliği de $p < 0,5$ olduğunda $lojit(p) < 0$ ve eğer $p \geq 0,5$ ise $lojit(p) \geq 0$ değerini alacak olmasıdır.⁴²

Lojit ile sürekli değişkenler arasındaki ilişkinin doğrusallığını tanımlamak amacıyla “tasarım değişkenleri” (design variables) ve Kesirli Polinomlar (Fractional Polynomials) olmak üzere iki yöntem önerilmektedir.⁴³

Bu yöntemlerden Royston ve Altman (1994) tarafından geliştirilen Kesirli Polinomlar bir dönüşüm yaklaşımıdır ve tanımsal istatistik bilgilerinden kartil değerleri kullanılarak çizilen grafipler üzerinden değerlendirmesi yapılan tasarım değişkenleri yöntemine göre daha analitiktir⁴⁴. Bu yaklaşımla sürekli değişkenin en iyi modelde, x 'in kaçınıcı üssü olarak yer alacağına karar verilmektedir. Kesirli Polinomlar yöntemi Çok Değişkenli Lojistik Regresyon Modellerinde de kullanılabilir ancak gösterimdeki basitliği açısından bu çalışmada, bağımsız değişkenin tek ve sürekli olduğu durumun modellenmesi için tanımlanmıştır. Lojit'in $g(x, \beta) = \beta_0 + x\beta_1$ şekilde olduğunda, bu fonksiyonu genellemenin bir yolu $g(x, \beta) = \beta_0 + \sum_{j=1}^J F_j(x)\beta_j$ 'dir. Burada $F_j(x)$ fonksiyonu bir çeşit güç fonksiyonu, ilk fonksiyonun değeri ise $F_1(x) = x^{p_1}$ 'dir. Teoride gücü ifade eden p_1 değeri bir reel sayıdır. Royston ve Altman (1994) bu gücü $P = \{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$ kümesi ile sınırlamayı önermişlerdir. $p_1 = 0$ değeri değişkenin logaritmasını ifade etmektedir⁴⁵ ve $(j = 2, 3, \dots, J)$ olmak üzere fonksiyon şu şekilde gösterilir;

⁴² Hüseyin Tatlıdil, **Uygulamalı Çok Değişkenli İstatistiksel Analiz**, Cem Ofset Ltd. Şti., Ankara, 1996, s. 293.

⁴³ David W. Hosmer, Stanley Lemeshow; **a.g.e.**, p. 99.

⁴⁴ P. Royston, D. G. Altman, “Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling”, **Applied Statistics**, Vol: 43, 1994, pp. 429–467.

⁴⁵ W. Sauerbrei, C. Meier-Hirmer, A. Benner, P. Royston, “Multivariable Regression Model Building by Using Fractional Polynomials: Description of SAS, STATA and R Programs”, **Computational Statistics & Data Analysis**, Vol: 50, 2006, p. 3466.

$$F_j(x) = \begin{cases} x^{p_j}, & p_j \neq p_{j-1} \\ F_{j-1}(x) \ln(x), & p_j = p_{j-1} \end{cases}$$

$J = 1$ olduğu durum için ($p_1 \in P$) olacak şekilde 8 farklı model yazılabilmektedir ve bu modeller içinden log-olabilirliği büyük olan seçilir. $J = 2$ için 36 farklı model oluşturulur ve yine bunlar içinden en iyi model olarak, büyük log-olabilirliğe sahip olan alınır. Daha sonra bu en iyi modelin Doğrusal Model'den anlamlı bir şekilde iyi olup olmadığını testi yapılır. $J = 1$ iken, en iyi model Doğrusal Model'in kısmi Olabilirlik Oran (Likelihood Ratio test – LR) testi yapılarak karşılaştırılır,

$$LR(1, p_1) = -2 \ln\{L(1) - L(p_1)\}.$$

Buradan $L(.)$ ile olabilirlik fonksiyonu* (likelihood function) ifade edilmektedir ve elde edilen değerler, X 'in doğrusal dağıldığı sıfır hipotezi altında 1 serbestlik dereceli Ki-kare dağılımı ile karşılaştırılmaktadır. Aynı şekilde $J = 1$ ve $J = 2$ iken elde edilen en iyi modellerin karşılaştırılmasında ise kısmi Olabilirlik Oran istatistiğinden elde edilen değer, ikinci fonksiyonun değeri sıfıra eşittir sıfır hipotezi altında 2 serbestlik dereceli Ki-kare dağılımıyla test edilmektedir. Bu şekilde değişkeni en iyi modelleyen fonksiyona karar verilir⁴⁶.

Kay ve Little (1987)⁴⁷, Copas J. B. (1983)⁴⁸ Kesirli Polinomlar yöntemini uyarlayarak, bağımlı değişken grupları içinde nicel bağımsız değişkenin dağılımına bakıp, bu değişkenlerin yaklaşık olarak nasıl modelleneceğine karar veren yeni bir yöntem önermişlerdir. Örneğin ilgili nicel değişken bağımsız değişkenlerin grupları için farklı ortalama ve farklı varyansla Normal dağılıma uyuyorsa Lojit Modelde, doğrusal terim x ve karesel terim x^2 olmalıdır. Bu da

* Olabilirlik Fonksiyonu (Likelihood Function) ile ilgili açıklama için sayfa 22'ye, Olabilirlik Oran Testi (Likelihood Ratio Test) için sayfa 26'ya bakınız.

⁴⁶ P. Royston, D. G. Altman **a.g.e.**, pp. 429–467.

⁴⁷ R. Kay, S. Little, “Transformation of the Explanatory Variables in the Logistic Regression Model for Binary Data”, **Biometrika**, Vol: 74, 1987, pp. 495–501.

⁴⁸ J. B. Copas, “Plotting p Against”, **Applied Statistics**, Vol: 32, 1983, pp. 25–31.

$p_1 = 1$ ve $p_2 = 2$ ile $J = 2$ olan Kesirli Polinomlara uyar. Eğer x değişkeni Gamma dağılıma uyuyorsa modelde x ve $\ln(x)$ 'e ihtiyaç vardır. Bu da $p_1 = 1$ ve $p_2 = 0$ ile $J = 2$ olan Kesirli Polinomlara uyar. Eğer x değişkeni bağımlı değişkenin grupları için farklı parametre değerleri ile Beta dağılımına uyuyorsa nicel değişkenin doğru bir şekilde modellenmesi için $\ln(x)$ ve $\ln(1-x)$ 'in modelde yer alması gerekmektedir. $\ln(1-x)$ Kesirli Polinomlar içinde gösterilememektedir. Bunun dışında bir modelde kullanılan lojistik olasılık değeri bir başka modelde nicel değişken olarak kullanılabilir. Eğer bağımsız değişken grupları yeterli örneklem büyüklüğüne sahipse, Kay ve Little'in yaklaşımını kullanarak bağımsız nicel değişkenin dağılımına karar vererek değişkeni modellemek daha uygun olabilir.

1.2.1 Lojistik Regresyon Modellerine İlişkin Varsayımlar

Lojistik Regresyon Modelinin varsayımları aşağıdaki gibidir:⁴⁹

- İlgilenilen durum için grup üyelerinin durumları bilinmelidir. Örneğin, iki durumlu değişken için “1” değerini alma olasılığı P_1 ve “0” değerini alma olasılığı $P_0 = 1 - P_1$ olarak tanımlanmalıdır.
- Bağımlı değişkenler istatistik olarak birbirleri ile ilişkisiz olmalıdır.
- Model doğru olarak tanımlanmalıdır.
- Analiz edilen kategoriler *bağımsız ve bağdaşmayan* olaylar olmalıdır. Başka bir deyişle her olay tek bir duruma ait olmalıdır.
- Lojistik regresyon katsayılarının hipotez testlerinin uygunluğu için büyük örneklem gerekmektedir. Bunun nedeni, En Çok Olabilirlik Tahmininin katsayıların standart hatalarını büyük örneklem (bkz. sayfa 83) için hesaplanabilmesidir.

⁴⁹ Raymond E. Wright, **Reading and Understanding Multivariate Statistics**, American Psychological Association, Edited By Laurence G. Grimm and Paul R. Yarnold, 1994, p. 220.

1.2.2 Lojistik Regresyon Analizinde Model Kurma Stratejileri

Model kurarken amaç, yeterince basit ama önemli etmenleri göz ardı edecek kadar da basitleştirilmemiş bir modele ulaşmaktır. Tesadüfi değişkenlerin özelliklerine ilişkin olarak yapılan her türlü varsayım model kurmanın ilk aşamasını oluşturmaktadır. Sonraki aşama uygun tahmin yöntemleri kullanarak parametre katsayılarının tahmin edilmesidir. Üçüncü aşama ise modelin doğrulanmasıdır. Bir model ileri sürülürken, bağımlı değişkenin yapısının ve bağımsız değişkenlerin iyi belirlenmesi ile birlikte uygun fonksiyonun seçilmesi de önemlidir. Bu aşamada sorun varsa ilk adım olan model kurma aşamasına geri dönülür ve katsayılar tekrar tahmin edilerek sürece devam edilir. Son aşama olan sonuçları yorumlama ve çıkarım yapma aşamasında model kurma hataları ne kadar ciddiye, uygulanan modelden yapılan çıkarımlar da o kadar az güvenilir olacak demektir⁵⁰.

Lojistik Regresyon Modeli için değişken seçimi birkaç adımdan oluşmaktadır:⁵¹

- 1- Değişken seçimine her bir değişkenin tek başına analizi ile başlanmaktadır. Bunun için Olabilirlik Oran testi ya da onunla aynı sonucu veren Pearson Ki-kare testi kullanılabilir. Nicel değişkenler için katsayıların anlamlılığı önemlidir ve bu nedenle her bir değişkenin Wald istatistiğine bakılır. Değişkenlerin tek tek testinde bu istatistiğe denk olan t-testi de kullanılabilir. Nicel bağımsız değişkenlerin tek değişkenli olarak Lojistik regresyona uygunluğu için nicel değişkenle lojit değerleri arasında doğrusal bir ilişki olması beklenmektedir.
- 2- Değişkenlerin tek tek analizi tamamlandıktan sonra, seçilen değişkenlerin çoklu analizi yapılmaktadır. Lojistik Regresyon Analizi için genellikle α değeri 0,25 olarak alınmaktadır. Anlamlılık seviyesi için geleneksel olarak kullanılan 0,05 ve 0.01 değerleri alındığında model için önemli olan

⁵⁰ Paul Newbold, William L. Carlson, Betty Thorne, **Statistics for Business and Economics**, Sixth edition, Prentice Hall 2007, p.539.

⁵¹ David W. Hosmer, Stanley Lemeshow, **a.g.e.**, p. 92.

değişkenlerin dışarıda kalabileceğini gösteren çalışmalar bulunmaktadır⁵². Genellikle, tüm mümkün değişkenlerle modele başlama kararının uygunluğu, örnek büyüklüğüne ve aday değişkenlerin toplam sayısı ile ilişkili her bir bağımlı değişkeninin grup sayısına bağlıdır. Veri seti örnek büyüklüğü açısından uygunsuzsa analize çok değişkenli modelle başlamak uygun olabilir⁵³. Bununla birlikte, veri seti uygun değilse yani örneklem büyüklüğü yeterli değilse çok değişkenli modelle başlamak doğru değildir. Bu durumda değişken seçimi için değişkenlerin tek tek testine dayanan Wald istatistiği kullanılmamalıdır. Onun yerine, modeli oluşturacak değişkenlerin seçimi En İyi Alt Küme (Best Subsets)⁵⁴ seçimine dayandırılır.

Diğer bir değişken seçimi Adım Adım (Stepwise) yöntemidir. Uygulamada Adım Adım yöntemi, İleri Seçim (Forward Selection) ve Geriye Doğru Eleme (Backward Elimination) olmak üzere iki temel prensiple hareket etmektedir. Adım Adım, En İyi Alt Küme ve diğer mekanik seçim yöntemlerinde, modele alınan değişkenler konusunda dikkatli olunması gerekmektedir⁵⁵.

- 3- Bu adımda, çok değişkenli modelin uygunluğu, modelde bulunan her bir değişkenin önemi Wald istatistiği ile test edilerek doğrulanmalıdır. Kriteri sağlamayan değişkenler modelden çıkartılarak model yeniden oluşturulur ve yeni modelle eski modelin Olabilirlik Oran testi hesaplanarak karşılaştırılır. Ayrıca, modelde kalan değişkenlerin tahmin edilen parametreleri tam modelin parametreleri ile kıyaslanır. Bu işleme tüm önemli değişkenler modele alınana kadar devam edilir. Bu adımın sonunda

⁵² J. Mickey, S. Greenland “A Study of the Impact of Confounder-selection Criteria on Effect Estimation”, **American Journal of Epidemiology**, 129, pp. 125–137.

⁵³ David W. Hosmer, Stanley Lemeshow, **a.g.e.**, p. 95.

⁵⁴ David W. Hosmer, B. Jovanovic, Stanley Lemeshow, “Best Subsets Logistic Regression”, **Biometrics**, Vol:45, 1989, pp. 1265–1270.

⁵⁵ David W. Hosmer, Stanley Lemeshow, **a.g.e.**, p. 96.

ortaya çıkan modele *Başlangıç Temel Etki Modeli* (Preliminary Main Effects Model) adı verilmektedir.⁵⁶

- 4- Önemli değişkenlerin içinde olduğu model oluşturulduktan sonra, modeldeki değişkenler daha yakından incelenmelidir. Lojit ile nicel değişken arasındaki ilişki doğru olarak tanımlandıktan sonra (bu ilişki doğrusal, karesel veya doğrusal olmayan bir fonksiyon olabilir) nicel değişkenlerin lojit değerleri ile doğrusallık varsayımı kontrol edilmelidir. Doğrusallık varsayımının sağlandığı fonksiyon için parametre katsayıları elde edilmektedir. Dördüncü adım sonunda oluşturulan modele *Temel Etkili Modeller* (The Main Effects Model) adı verilmektedir.⁵⁷
- 5- Son olarak modeldeki değişkenlerin etkileşimlerinin kontrolü yapılır. Değişkenlerin etkileşimli halleri modele teker teker alınır ve belirli anlamlılık düzeyinde Olabilirlik Oran testi yapılır. Adım beşin sonunda elde edilen modele *Sonuç Model* (Final Model) denir.⁵⁸ Modeli yazmadan önce modelin uygunluk testleri yapılmalıdır (bkz. Başlık 1.2.5'e).

1.2.3 Lojistik Modellerde Parametre Tahmin Yöntemi

Parametre tahminleri açısından Doğrusal ve Lojistik Regresyon Analizi arasındaki farklardan iki tanesi çok önemlidir. $y = E(y|x) + \varepsilon$ şeklinde ifade edilen modelde ε 'la gösterilen hata teriminin, Doğrusal Regresyon Modelinde En Küçük Kareler (EKK) yöntemi ile parametre tahminlerinin geçerli olması için Normal dağılımlı ve eşit varyanslı olması ($\varepsilon \sim N(0, \sigma^2)$) gerekmektedir. Lojistik Regresyon Analizinde bağımsız değişkenin gözlenen değeri $y = P(x) + \varepsilon$ olarak gösterilebilir ve buradaki hata terimi; eğer $y = 1$ ise $P(x)$ olasılığı ile $\varepsilon = 1 - P(x)$ 'e ve eğer $y = 0$ ise $1 - P(x)$ olasılıkla $\varepsilon = -P(x)$ 'e eşit olacak şekilde iki mümkün sonuca sahiptir. Buradan da, bağımlı değişkenin koşullu dağılımının, koşullu beklenen değer $P(x)$ ile Binom dağılıma uyduğu

⁵⁶ David W. Hosmer, Stanley Lemeshow, a.g.e., p. 97.

⁵⁷ A.e., p. 97.

⁵⁸ A.e., p. 99.

söylenbilir.⁵⁹ Buradan da anlaşılacağı gibi Lojistik Regresyon Analizinde parametre tahminleri için, varsayımların sağlanamaması nedeniyle, EKK yönteminin kullanılması uygun değildir.

İki Kategorili Lojistik Regresyon Modellerinde parametre tahminlerinin yapılabilmesi için literatürde bazı yöntemler önerilmektedir. Bu tekniklerden uygulamada sıklıkla kullanılan En Çok Olabilirlik Tahmini'dir (Maksimum Likelihood Estimation-MLE). Bunun yanı sıra, Yeniden Ağırlıklandırılmış Tekrarlı En Küçük Kareler Tahmini (Reweighted Iterative Least Square Estimation), 1944 yılında Berkson tarafından Ağırlıklı En Küçük Kareler Tahmininden, tekrarlı veri durumunda kullanılmak üzere geliştirilmiş olan Minimum Lojit Ki-kare (Minimum Logit Chi-Square) ve Diskriminant Fonksiyonundan elde edilen Diskriminant katsayıları⁶⁰ kullanılan diğer tahmin yöntemleri arasında yer almaktadır.

Diskriminant fonksiyon tahminlerinin tutarlı olabilmesi için normallik varsayımının sağlanması gerekmektedir. Eğer modelde hem nicel hem de nitel bağımsız değişkenler birlikte bulunuyor ve nitel bağımsız değişkenlerin sayısı fazla ise, bu durum normallik varsayımının bozulmasına neden olmaktadır. Bu sebeple Diskriminant Fonksiyon Tahmin yöntemiyle elde edilen katsayı değerleri olması gerekenden daha büyük çıkacaktır⁶¹.

Bu yöntemler içinde en sık kullanılan nokta tahmin yöntemi “En Çok Olabilirlik” Tahminidir.

⁵⁹ David W. Hosmer, Stanley Lemshow, a.g.e., p. 7.

⁶⁰ Hüseyin Tatlıdil, a.g.e., s. 293.

⁶¹ David W. Hosmer, Stanley Lemshow, a.g.e., pp. 21–22.

En Çok Olabilirlik Tahmin (Maksimum Likelihood Estimation – MLE)Yöntemi*:

Anakütleyi belirleyen olasılık yoğunluk fonksiyonu ya da olasılık fonksiyonu $f(y; \theta)$, bu anakütleden çekilmiş n gözlemlili rastgele örneklem Y_1, Y_2, \dots, Y_n 'nin gerçekleşen bir alt kümesi ise y_1, y_2, \dots, y_n olarak gösterildiğinde, En Çok Olabilirlik Tahmin yöntemi belli bir örneklem değerinin gerçekleşme olasılığını en yüksek yapan parametre değerlerini bulmaya çalışmaktadır⁶². Bir rassal örneklem söz konusu olduğunda ve bunların çekildiği anakütle dağılımının bilindiği varsayıldığında; bağımsızlık özelliğinden hareketle ortak olasılık yoğunluk fonksiyonu,

$$\begin{aligned} f(y_1, y_2, \dots, y_n; \theta) &= f(y_1; \theta) \cdot f(y_2; \theta) \cdot \dots \cdot f(y_n; \theta) \\ &= \prod_{i=1}^n f(y_i; \theta) \end{aligned}$$

olarak yazılmaktadır⁶³.

Olabilirlik fonksiyonu, ortak olasılık fonksiyonuna verilen başka bir isimdir. Tek fark ortak olasılık fonksiyonunda θ 'nın bilindiği Y 'lerin bilinmediği, olabilirlik fonksiyonunda ise Y 'lerin bilindiği θ 'nın ise bilinmediği varsayılmaktadır⁶⁴. Olabilirlik fonksiyonu, rassal örneklemin belli bir sonucu olan y ($y = \{y_1, y_2, \dots, y_n\}$) biliniyorken θ 'yı bilinmeyen olarak ifade eden bir fonksiyondur ve aşağıdaki şekilde gösterilmektedir.

$$L(\theta|y_1, y_2, \dots, y_n) = L(\theta|y) = \prod_{i=1}^n f(y_i; \theta)$$

* Bazı kaynaklarda En Çok Benzerlik Tahmini olarak da kullanılmaktadır. (bkz. Neyran Orhunbilge, **Çok Değişkenli İstatistik** kitabı)

⁶² Raymond E. Wright, **a.g.e.**, p. 225.

⁶³ David W. Hosmer, Stanley Lemeshow; **a.g.e.**, pp. 7–9.

⁶⁴ Jeff Gill, **a.g.e.**, p. 21.

Bu olabirlik fonksiyonu bağımlı deęişkenin iki kategorili olduęu durum için genelleştirildiğinde; istenen durumun ortaya çıkma olasılığı $p_i = P(y_i = 1|x_i)$ ve istenmeyen durumunun ortaya çıkma olasılığı $1 - p_i = P(y_i = 0|x_i)$ olarak gösterilirse, i. gözlem için olasılık deęeri,

$$P(y_i|x_i) = p_i^{y_i}(1 - p_i)^{1-y_i} \quad 1.19$$

olarak ifade edilmektedir. Bu olasılığın birbirinden bağımsız n gözlem için olabirlik fonksiyonu aşağıdaki gibidir⁶⁵.

$$l(\beta|y, x_i) = \prod_{i=1}^n p_i^{y_i}(1 - p_i)^{1-y_i}$$

Bilindięi gibi En Çok Olabirlik Tahmin yönteminin amacı açıklayıcı deęişkenlerin katsayısı olan β deęerlerini bağımlı deęişkeni y 'nin gözlenme olabirliğini maksimum yapacak biçimde tahmin etmektir. Kısaca, $\max \ln[l(\beta|y, x_i)] = \ln[\prod_{i=1}^n p_i^{y_i}(1 - p_i)^{1-y_i}]$ olarak tanımlanmaktadır. ⁶⁶

Bu maksimizasyon probleminde kolaylık sağlamak için ortak olasılık fonksiyonunun doęal logaritması kullanılarak;

$$\begin{aligned} \max \ln L(\beta|y, x_i) &= \ln \prod_{i=1}^n p_i^{y_i}(1 - p_i)^{1-y_i} \\ &= \sum_{i=1}^n [y_i \ln p_i + (1 - y_i) \ln (1 - p_i)] \end{aligned} \quad 1.20$$

fonksiyonu elde edilir.

Lojistik Regresyon Modelinden, Eşitlik 1.18'de verilen olayın gerçekleşme olasılık deęeri Eşitlik 1.20'da yerine konularak fonksiyonun β parametrelerine göre birinci mertebeden türevi alınıp, sıfıra eşitlenerek elde edilen denklemlerin

⁶⁵ Fred C. Pampel, **a.g.e.**, p. 41.

⁶⁶ Hüseyin Tatlıdil, **a.g.e.**, s. 296.

çözümleri $\hat{\beta}$ 'leri vermektedir. Fakat bu eşitliklerdeki β parametrelerinin olabilirlik fonksiyonu içinde doğrusal olmaması nedeniyle iteratif yolla çözüme gidilmektedir.

En Çok Olabilirlik Tahmininin, iteratif çözüm yöntemleri içinde; En Hızlı Artış (Steepest Ascent), Newton-Raphson, Derecelendirme, BHHH* ve Sayısal Türevler (Numerical Derivatives) gibi yöntemler bulunmaktadır⁶⁷. Bu yöntemlerden tercih edilen ve birçok paket programın kullandığı yöntem Newton-Raphson yöntemidir.

Tahmin edilen parametrelerin yeterlilik koşulunu sağlaması için log-olabilirlik fonksiyonunun $[\ln L(\beta|y, x_i)]$ bağımsız değişkenlerin katsayılarını ifade eden β parametrelerine göre ikinci mertebeden türevlerinin (Hessian matrisini verir) sıfırdan küçük olması gerekmektedir $\left[\frac{\partial^2}{\partial \beta^2} \ln L(\beta|y, x_i) < 0 \right]$. Örnek büyüklüğü “∞”a giderken bu koşulların gerçekleşmesi ile En Çok Olabilirlik Tahminleri tutarlı, asimptotik olarak ise normal ve etkin⁶⁸ olurlar. Bu, En Çok Olabilirlik Tahmininin küçük örneklem büyüklüğü için kötü tahmin yapacağı anlamına gelmemektedir.

En Çok Olabilirlik Tahmininde örnek büyüklüğü için literatürde farklı kriterler bulunmaktadır. Genellikle 100'den küçük örnek büyüklüğü kullanmak riskli kabul edilirken, bazı kaynaklarda büyüklüğün 500'ün üzerinde olması gerektiği değişebilse de öncelikle modelde ne kadar çok değişken varsa o kadar çok gözleme ihtiyaç olduğu söylenebilir. Bazı kaynaklarda, her bağımsız değişken için en az 50 gözlemin alınması gerektiği ifade edilmektedir⁶⁹. İkinci olarak eğer veri setindeki bağımsız değişkenler arasında yüksek korelasyon varsa ya da bağımlı değişkendeki değişiklik azsa yani neredeyse gözlemlerin çoğu

* BHHH: Berndt, E. R., Hall, B. H., Hall, R. E. ve Hausman, J. A

⁶⁷ J. Scott Long, **a.g.e.**, pp. 54–57. (daha fazla bilgi için bkz.)

⁶⁸ Damodar N. Gujarati, **Temel Ekonometri**, Literatür Yayıncılık, 1999, ss. 110–113.

⁶⁹ Raymond E. Wright, **a.g.e.**, p. 221.

gruplardan birinde yer alıyorsa örneklemin yine büyük olması beklenir. Ayrıca bağımlı değişkenin sıralı (ordinal) olduğu model analizi gibi bazı modeller daha fazla örnek büyüklüğü gerektirmektedir⁷⁰.

Lojistik regresyonda parametre tahminleri için genellikle En Çok Olabilirlik Tahmin yöntemi seçilmektedir. Lojistik regresyon katsayıları “En Çok Olabilirlik parametre tahminleri” olarak bilinmektedir. Araştırmacılar tarafından genelde En Çok Olabilirlik değeri yerine Log-olabilirlik (LL) değeri kullanılmaktadır. Log-olabilirlik değerinin “0” dan küçük değerler aldığı durumlarda 2 ile çarpılarak, büyüklüğü negatiflikten kurtarmak için de “-” ile çarpılarak⁷¹ sapma değeri (-2LL) raporlanmaktadır. Log-olabilirlik (LL) genellikle negatif, sapma değeri genellikle pozitiftir. Katsayıların olabilirliğini maksimum yapmaya çalışmak aynı zamanda LL’yi maksimum yapmak ve sapmayı minimum yapmakla eşdeğerdir⁷².

1.2.4 Parametrelerin Anlamlılık Testleri

Herhangi bir modelin katsayılarının anlamlılığını test etmede amaç “bağımsız değişken modelde olduğu zaman bağımlı değişken hakkında modelde bulunmadığı duruma göre daha mı fazla bilgi veriyor” sorusunun cevabını bulmaktır. Buradan yola çıkarak test istatistikleri, bağımsız değişkenin modele dahil edildiği ve edilmediği iki durum için modelin tahminlerinin farklarının alınmasına dayanmaktadır.

En Çok Olabilirlik Tahminleri asimptotik olarak Normal dağılmaktadırlar. Bunun anlamı örneklem büyüklüğü arttıkça, En Çok Olabilirlik Tahminleyenlerinin örneklem dağılımlarının Normal dağılıma yaklaştığıdır. Tek bir parametre için bu yaklaşım $\hat{\beta}_k \sim N(\beta_k, Var(\hat{\beta}_k))$ iken parametrelerin bir vektörü için $\hat{\beta} \sim N(\beta, Var(\hat{\beta}))$ ’dir. Burada $Var(\hat{\beta})$, $\hat{\beta}$ ’nin kovaryans

⁷⁰ J. Scott Long, **a.g.e.**, p. 54.

⁷¹ Neyran Orhunbilge, **Çok Değişkenli İstatistik Yöntemler**, İstanbul Üniversitesi Basım ve Yayımevi Müdürlüğü, İstanbul, 2010, s. 191.

⁷² Raymond E. Wright, **a.g.e.**, p. 230.

matrisidir⁷³. En Çok Olabilirlikle, tahminleri yapılan Lojistik regresyon parametrelerinin anlamlılığı Wald ve Olabilirlik Oran testleri ile test edilmektedir.

Wald Testi⁷⁴

Bu test Regresyon Analizinde regresyon katsayılarının t veya Z testlerinin benzeridir.⁷⁵ Wald testinde kısıtların doğrusal formlarını göz önüne alarak sıfır hipotezi en genel hali ile $H_0: Q\beta = r$ şeklinde gösterilmektedir. Bu eşitlikte β , testi yapılacak parametreler vektörü; Q , sabit sayı matrisi; r ise sabit sayı vektörüdür. Q ve r doğrusal kısıtlardaki değişimi ifade etmektedir.

$H_0: Q\beta = r$ hipotezi Wald istatistiği ile şu şekilde test edilebilir;

$$W = [Q\hat{\beta} - r]' [Q\widehat{\text{Var}}(\hat{\beta})Q']^{-1} [Q\hat{\beta} - r]. \quad 1.21$$

W , sabitlerin sayısı kadar yani Q 'daki satır sayısına eşit olan serbestlik derecesi ile Ki-kare dağılımına uymaktadır. Wald istatistiğinin iki bileşeni bulunmaktadır. Birincisi, formüldeki tahmin ve hipotez değeri arasındaki farkı ölçen $(Q\hat{\beta} - r)$ değeridir. Farkların karesi negatif ve pozitif farkın test istatistiği üzerinde aynı etkiye sahip olmasını sağlamaktadır. İkincisi ise tahminleyenin değişkenliğini yansıtan $[Q\widehat{\text{Var}}(\hat{\beta})Q']^{-1}$ değeridir.

İki değişkenli Lojit Model ele alırsa; $\Pr(y = 1|x) = \Lambda(\beta_0 + \beta_1x_1 + \beta_2x_2)$ için $H_0: \beta_1 = \beta^*$ testi matris formunda şu şekilde ifade edilmektedir;

$$Q\hat{\beta} - r = (0 \quad 1 \quad 0) \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} - (\beta^*) = \hat{\beta}_1 - \beta^*$$

⁷³ J. Scott Long, **a.g.e.**, p. 89.

⁷⁴ **A.e.**, pp. 89–92.

⁷⁵ Neyran Orhunbilge, **Çok Değişkenli İstatistik Yöntemler, a.g.e.**, s. 192.

$$[Q\widehat{\text{Var}}(\hat{\beta})Q']^{-1} = \left[\begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \widehat{\text{Var}}(\hat{\beta}) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right]^{-1} = \frac{1}{\hat{\sigma}_{\hat{\beta}_1}^2}.$$

Bu ifade testi yapılan parametre varyansının tersidir. Varyans büyüdükçe, hipotez ve tahmin değeri arasındaki farkı veren ağırlık azalmaktadır. Eşdeğer olarak, olabilirlik fonksiyonunun eğimi $\hat{\beta}_1$ 'nin etrafındaki bölgede değişmekte ve $(\hat{\beta}_1 - \beta^*)$ farkı daha anlamlı olmaktadır. Wald istatistiği aşağıdaki şekilde formüle edilmektedir.⁷⁶

$$W = \frac{(\hat{\beta}_1 - \beta^*)^2}{\hat{\sigma}_{\hat{\beta}_1}^2} = \left(\frac{\hat{\beta}_1 - \beta^*}{\hat{\sigma}_{\hat{\beta}_1}} \right)^2 \quad 1.22$$

Wald istatistiği, Z test istatistiğinin karesidir. “1” serbestlik dereceli Ki-kare değeri normal değerın karesine eşit olmaktadır.⁷⁷

Aynı şekilde $H_0: \beta_1 = \beta_2 = 0$ hipotezi test edilmek istendiğinde Wald istatistiği aşağıdaki şekilde hesaplanmaktadır.

$$W = \sum_{k=1}^2 \frac{\hat{\beta}_k^2}{\hat{\sigma}_{\hat{\beta}_k}^2} = \sum_{k=1}^2 z_{\hat{\beta}_k}^2 \quad 1.23$$

Wald istatistiğinde Lojistik regresyon katsayıları mutlak değer olarak büyüdükçe, bu değerlerden tahmin edilen standart hatalar da çok fazla büyüyecektir.

⁷⁶ W. W. Hauck, A. Donner, “Wald’s Test as Applied to Hypotheses in Logit Analysis”, **American Statistical Association**, Vol: 72, No: 360, 1977, pp. 851–853.

⁷⁷ A. Buse, “The Likelihood Ratio, Wald and Lagrange Multiplier Tests: An Expository Note”, **The American Statistician**, Vol: 36, No: 3, 1982, pp. 153–157.

Olabilirlik Oran Testi

Olabilirlik Oran testi (Likelihood Ratio test-LR test), Lojistik Regresyon Modeli'ndeki kısıtların* testinde kullanılmaktadır. Örnek olarak aşağıdaki Lojit Modeller ele alındığında,⁷⁸

$$M_1: \Pr(y = 1|x) = \Lambda(\beta_0 + \beta_1x_1 + \beta_2x_2)$$

$$M_2: \Pr(y = 1|x) = \Lambda(\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3)$$

$$M_3: \Pr(y = 1|x) = \Lambda(\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_4x_4)$$

$$M_4: \Pr(y = 1|x) = \Lambda(\beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_4x_4)$$

model M_1 'in, $\beta_3 = 0$ kısıtının geçerli olduğu M_2 'nin ve aynı zamanda $\beta_4 = 0$ kısıtının geçerli olduğu M_3 'ün formunda olduğu görülmektedir. Bir model kısıtlarla birlikte diğer bir modelden elde edilebiliyorsa, kısıtlanmış model kısıtlanmamış modelin içinde yer almaktadır (nested – iç içe) denmektedir. O halde M_1, M_2 ve M_3 'ün içinde yer almaktadır. Bununla birlikte; ne M_2 'nin M_3 'ün içinde, ne de M_3 'ün, M_2 'nin içinde olduğu görülmektedir.

Olabilirlik Oran testi, β_U parametreleri ile kısıtlanmamış model M_U ve bu model içinde yer alan β_R parametreleri ile kısıtlanmış M_R modelini karşılaştırmaktadır. Sıfır hipotezi, “kısıtlanmış modeli ifade eden M_R modeli doğrudur” olarak kurulmaktadır. Kısıtlanmamış model için En Çok Olabilirlik Tahminlerinden elde edilen olabilirlik fonksiyonunun değeri $L(M_U)$ ve kısıtlanmış tahminlerin değeri $L(M_R)$ ise, Olabilirlik Oran (LR) istatistiği,

$$LR = -2 \ln[L(M_R) - L(M_U)] = 2 \ln L(M_U) - 2 \ln L(M_R)$$

* Bu bölümde kullanılan “kısıt” kelimesi parametre değerinin “0” olması durumunu ifade etmektedir. Genellikle kaynaklarda Olabilirlik Oranı; değişkenin modelde yer almadığı olabilirlik değerinin, değişken modelde bulunduğu olabilirlik değerine oranı olarak verilmektedir.

⁷⁸ J. Scott Long, **a.g.e.**, p. 93.

olarak hesaplanmaktadır. LR , bağımsız kısıtların sayısı kadar serbestlik derecesi ile Ki-kare dağılımına uymaktadır. LR istatistiği iç içe model çiftlerinin karşılaştırılmasında kullanılmaktadır. Bu testin sonucunda hatalı karar verme riskini azaltmak için, karşılaştırılacak modellerin tamamının aynı örneklem kümesi için kullanıldığından emin olmak gerekmektedir. Herhangi bir modelin testine başlamadan önce, modelde kullanılan tüm değişkenler için gözlenmemiş değerlerin olduğu gözlem birimlerinin dışında bırakılarak veri kümesinin tekrar yapılandırılması önemlidir⁷⁹.

Standart istatistik paket programları genellikle LR için iki test değeri hesaplamakta ve bu hesap değerlerine ML tarafından modellenen sonuçların sunulduğu tablonun içinde yer vermektedirler. İlk test, verilen modelde tüm bağımsız değişken katsayılarının sıfıra eşit olduğunu öngören kısıtlanmış modelin karşılaştırılmasıdır. Bu testten genellikle Olabilirlik Oran Ki-kare testi olarak bahsedilmektedir. Model M_β , sabit ve bağımsız değişken katsayılarının içinde bulunduğu kısıtlanmamış modeli; M_α ise tüm bağımsız değişken katsayılarının model dışı bırakıldığı ve sadece sabitin modelde olduğu kısıtlı model temsil etmektedir. “Bağımsız değişken katsayılarının tamamı eş zamanlı olarak sıfıra eşittir” hipotezini incelemek üzere kullanılacak test istatistiği aşağıdaki gibi yazılmaktadır.

$$LR = 2\ln L(M_\beta) - 2\ln L(M_\alpha)$$

İkinci test olan D istatistiği, 1989 yılında McCullagh ve Nelder tarafından sapma (deviance) olarak adlandırılmıştır ve bu istatistik değeri GLM’in iskeletinin oluşturulmasında geniş bir şekilde yer almıştır⁸⁰. Aynı zamanda bu değer, modellerin uyum iyiliği testleri için geliştirilen bazı yaklaşımlarda da önemli bir rol üstlenmektedir. Bununla birlikte sapma değeri, ifade olarak Doğrusal Regresyon Analizi içindeki hata kareleri toplamının benzeridir.⁸¹ Sapma, verilen

⁷⁹ J. Scott Long, **a.g.e.**, p. 93.

⁸⁰ P. McCullagh, J. A. Nelder, **a.g.e.**, pp. 33–34.

⁸¹ David W. Hosmer, Stanley Lemeshow, **a.g.e.**, p. 13.

modelle (M_β) tam modeli (M_F) karşılaştırmaktadır. Gözlemlenmiş veri kümesi değerleri tam olarak tahmin edildiğinde, M_F 'nin olabilirliği "1" ve log-olabilirlik "0"dir. Sapma aşağıdaki gibi yazılmaktadır.

$$\begin{aligned} D(M_\beta) &= D(M_\beta|M_F) = 2\ln L(M_F) - 2\ln L(M_\beta) \\ &= -2\ln L(M_\beta) \\ &= LR(M_\beta|M_F) \end{aligned}$$

$D(M_\beta)$ istatistiği Ki-kare dağılımına uygun olmakla birlikte, gözlemlerin sayısı ile tam modeldeki parametrelerin sayısının doğrudan artması nedeniyle $D(M_\beta)$ asimptotik olarak Normal dağılıma da uymaktadır⁸². Hosmer ve Lemeshow (2000), seyrek veriler* söz konusu olduğunda modelin uygunluk ölçüsü olarak $D(M_\beta)$ 'nin kullanılmasının uygun olmadığını vurgulamaktadır⁸³.

LR ve Wald testleri asimptotik olarak eşit oldukları halde, sonlu örneklem, özellikle de küçük örneklem için farklı sonuçlar vermektedirler. Genellikle hangi testin tercih edileceği çok açık değildir. Rothenberg (1984) testlerden hiçbirinin diğerine göre üstün olmadığını ifade etmiştir⁸⁴. Hauck ve Donner (1977) yaptıkları çalışmada LR ve Wald testini karşılaştırmış ve LR testinin daha güçlü olduğunu söylemişlerdir⁸⁵. Uygulamada, hangi testin kullanılacağına seçimi sıklıkla, kullanılan bilgisayar programının uygunluğu tarafından belirlenmektedir. Bununla birlikte, birçok istatistikçi nispeten kolaylığı nedeniyle LR testini tercih etmektedir⁸⁶.

⁸² P. McCullagh, "The Conditional Distribution of Goodness-of-fit Statistics for Discrete Data", **Journal of the American Statistical Association**, Vol:81, No:393, 1986, pp. 104–107.

* Seyrek veri ile bağımsız değişkenlerin değerlerinin her bir kombinasyonunun örnekte sadece bir kere ortaya çıktığı durumdan bahsedilmektedir.

⁸³ David W. Hosmer, Stanley Lemeshow, **a.g.e.**, pp. 143–147.

⁸⁴ T. J. Rothenberg, "Hypothesis Testing in Linear Models When the Error Covariance Matrix is Nonscalar", **Econometrica**, 1984, Vol: 52, pp. 827–842.

⁸⁵ W. W. Hauck, A. Donner, **a.g.e.**, pp. 851–853.

⁸⁶ J. Scott Long, Simon Cheng, **The Handbook of Data Analysis**, Edited by Hardy and Bryman pp. 261–262.

1.2.5 Uyum İyiliği Testleri

Uygunluk ölçüsü bir modelin yeterli olup olmadığı hususunda kabaca bir indeks (gösterge) sağlayabilir. Ancak uygunluk ölçüleri model hakkında bazı bilgileri elde etmemizi sağlasa dahi, bu değerlerin geçmiş araştırma kapsamında ve dikkate alınan modelin tahmin parametreleri içinde değerlendirilmesi gereken sadece kısmi bilgiler olduğu unutulmamalıdır.

Pearson Ki-Kare Testi

Lojistik regresyon için gözlenen ve beklenen değer arasındaki farklar yani hatalar, İki Kategorili Model yani $P_i = E(y_i|x_i) = \Pr(y_i = 1|x_i)$ için, $Var(y_i|x_i) = P_i(1 - P_i)$ ile $y_i - P_i$ sapmaları farklı varyanslıdır ve hatalar şu şekilde hesaplanmaktadır,

$$e_i = \frac{y_i - \hat{P}_i}{\sqrt{\hat{P}_i(1 - \hat{P}_i)}}.$$

e_i 'nin büyük değerleri eldeki gözlemlerin modele uygun olmadığını göstermektedir. Bu hesaplanan hatalara dayanan özet istatistik Pearson Ki-kare istatistiğidir ve formülasyonu aşağıdaki gibidir,

$$\chi^2 = \sum_{i=1}^N e_i^2$$

χ^2 , Ki-kare dağılımına sahiptir.

$Var(y_i - P_i) = P_i(1 - P_i)$ iken, $Var(y_i - \hat{P}_i) \neq \hat{P}_i(1 - \hat{P}_i)$ 'dir. Sonuç olarak e_i 'nin varyansı "1"e eşit değildir. Standartlaştırılmış hataların hesaplanmasında şapka (hat)⁸⁷ matris veya önkestirim (predictive)⁸⁸ matrisi olarak adlandırılan H

⁸⁷ J. Scott Long, **a.g.e.**, p.99.

⁸⁸ Sampri Chatterjee, Ali. S. Hadi, Bertram Price, **Regression Analysis by Example**, Third Edition, John Wiley & Sons, 2000, p.36.

matrisi önemli bir rol oynamaktadır⁸⁹. Tahmin edilen hataların varyansları hesaplanırken, önkestim matrisi H 'nin ne olduğunun bilinmesi gerekmektedir. İkili regresyon modeli için Pregibon (1981) H matrisi aşağıdaki şekilde türetmiştir,

$$H = \hat{V}X(X'\hat{V}X)^{-1}X'\hat{V}.$$

\hat{V} , $\sqrt{\hat{P}_i(1 - \hat{P}_i)}$ 'nin köşegenleri üzerindeki değerlerden oluşan köşegen matrisi; X , örnek büyüklüğü n ve modeldeki bağımsız değişken sayısı p ile ifade edilmek üzere $n \times p$ boyutunda veri matrisidir. H 'nin sadece köşegen elemanlarına ihtiyaç duyulduğundan, hesaplama için daha basit olan

$$h_{ii} = \hat{P}_i(1 - \hat{P}_i)x_i\widehat{Var}(\hat{\beta})x_i'$$

formülü kullanılmaktadır. x_i , i . gözlem için bağımsız değişkenlerin değerlerinden oluşan satır vektörü ve $\widehat{Var}(\hat{\beta})$, En Çok Olabilirlik tahminleri $\hat{\beta}$ 'den elde edilen tahminlerin kovaryans değeridir⁹⁰. e_i 'nin varyansını tahmin etmek için $1 - h_{ii}$ kullanılır. Standartlaştırılmış Pearson hataları ise,

$$e_i^{std} = \frac{e_i}{\sqrt{1 - h_{ii}}}$$

şeklinde ifade edilir. Uygulamada genellikle bu iki hata değeri birbirine benzemekle birlikte e^{std} , e yerine tercih edilmektedir.

Bununla birlikte McCullagh (1986), Pearson Ki-kare ve daha önce bahsedilen sapma testinin (D), modelin uygunluğunu test etmek için özellikle kapsamlı ve seyrek veri kümesi söz konusu olduğunda kullanışlı istatistikler olmadığını ve χ^2 'nin hesaplanması zor olan ortalama ve varyans ile asimptotik Normal

⁸⁹ Erling B. Andersen, **Introduction to the Statistical Analysis of Categorical Data**, Springer, 1997, p. 30, 177.

⁹⁰ D. Pregibon, "Logistic Regression Diagnostic", **Annals of Statistics**, Vol:9, 1981, pp. 705–724.

dağılıma uyduğunu ispat etmiştir⁹¹. McCullagh'ın kapsamlı ve seyrek veri kümesi ile kastettiği bağımsız gözlemlerin sayısı veya “hücreler” çok büyük olduğu halde, her bir hücrenin toplam içindeki katkısının miktarı bakımından küçük olduğu durumlardır. Ayrıca McCullagh ve Nelder (1989), χ^2 'nin kesin uyum için kullanılmaması gerektiğini ve küçük örneklem için Olabilirlik Oran testinin Pearson Ki-kare testinden daha güçlü olduğu iddiasının da tersini söylemişlerdir⁹².

Hosmer-Lemeshow Testi

Hosmer ve Lemeshow (2000), seyrek veri kümesi için alternatif test olarak grup verilerini yapılandırarak test etmeyi önermişlerdir. n birime ait tahmin edilen olasılıkları küçükten büyüğe sıraladıktan sonra, iki farklı yöntem kullanılabilir. Yöntemlerden ilki $g = 10$ 'u grup sayısı olarak kullanarak $n'_1 = n/10$ adet En Küçük Tahmin olasılıklarına sahip birimi birinci gruba yerleştirip, son grup olan $n'_{10} = n/10$ birimi de en büyük tahmin olasılık değerine sahip grup olarak yazmaktır. İkinci yöntem ise kesim noktaları $k/10$ ile $k = 1, 2, \dots, 9$ olarak tanımlanan $g = 10$ grup kullanmaktır. Her iki gruplama stratejisi için de Hosmer-Lemeshow uyum iyiliği istatistiği, \hat{C} , gözlemlenen ve beklenen sıklıklardan oluşan $g \times 2$ tablosundan Pearson Ki-kare istatistiğinin hesaplanması ile elde edilir,

$$\hat{C} = \sum_{k=1}^g \frac{(o_k - n'_k \bar{P}_k)^2}{n'_k \bar{P}_k (1 - \bar{P}_k)}$$

n'_k , k . grup içindeki birimlerin toplamı, $o_k = \sum_{j=1}^{c_k} y_j$ ile $\bar{P}_k = \sum_{j=1}^{c_k} \frac{\hat{P}_k}{n'_k} \cdot \hat{C}$, $(g - 2)$ serbestlik derecesi ile χ^2 dağılımına uymaktadır.

⁹¹ P. McCullagh, **a.g.e.**, pp. 104–107.

⁹² P. McCullagh, J. A. Nelder, **a.g.e.**, pp. 112–122.

\hat{C} gibi özet uyum iyiliği ölçülerinin avantajı, tek bir değer ile uygunluğun kolayca yorumlanabilmesidir. Hosmer-Lemeshow testinin en büyük dezavantajı, gruplama sürecinde bireysel veri noktalarının az sayıda olması nedeniyle uyumda önemli bir sapmanın atlanmış olabileceğidir. Bu nedenle modelin uygunluğunu kabul etmeden önce, bireysel hataların analizinin ve tanımsal istatistiklerin yapılması uygun olur⁹³.

Bireysel hataların analizi ve tanımsal istatistikler incelenmeye, gözlem sayılarına karşı standartlaştırılmış artıkların *endeks grafiği* çizilerek, aykırı değerlerin araştırılmasıyla başlanır. Bu grafikten elde edilen değerler yanlış kodlanmış veya model eksik tanımlandığı için bazı değerler dışarıda kalmış olabilir. Çok büyük pozitif veya negatif hataların neden kaynaklandığı araştırılmadan, basit bir şekilde analizin dışarında bırakılmamalıdır. Ancak bu sorunun kaynağı, modelin doğru yapılandırıldığından emin olunması amacıyla ayrıca incelenmelidir. Büyük hatalar gözlemlerin pek de uygun olmadığını gösterirken, tahmin edilen β 'lar üzerinde büyük etkiye sahip olan veya olmayan gözlemleri göstermemektedir. Verilerin merkezine yakın olmak, her bir bağımsız değişken için bir gözlemin değerlerinin örneklemdaki değişkenin beklenen değerine yakın olması demektir. Diğer taraftan, uç gözlemler çok büyük artık değerlere sahip olmasalar da tahminleri etkilemektedirler. Bu tür gözlemleri bularak, dışarıda bırakıldıkları durumda $\hat{\beta}$ 'lerde meydana gelen değişimin hesaplanması uygun bir yoldur⁹⁴. Örneklemdaki her bir gözlemin tek tek dışarıda bırakılarak n kez modelin tahmin edilmesinin işlem yükü çok fazla olduğu için, Pregibon (1981), modeli tek seferde yaklaşık olarak tahminlemeyi amaçlamıştır. Eğer i . gözlem dışarıda bırakılırsa, $\hat{\beta}$ 'deki beklenen değişiklik ($\Delta_i \hat{\beta}$) yaklaşık olarak,

$$\Delta_i \hat{\beta} = \widehat{Var}(\hat{\beta}) x_i' \frac{y_i - \hat{P}_i}{1 - h_{ii}}$$

⁹³ David W. Hosmer, Stanley Lemeshow, **a.g.e.**, pp. 147–156.

⁹⁴ J. Scott Long, **a.g.e.**, p.100.

ifadesine eşittir. Çıkarılan x_i 'ye uygun olan β_k 'daki standartlaştırılmış değişim, DFBETA olarak bilinmektedir,

$$DFBETA_{ik} = \frac{\Delta_i \hat{\beta}_k}{\sqrt{\widehat{Var}(\hat{\beta}_k)}}.$$

$DFBETA_{ik}$ büyük ise, i . gözlemin β_k 'nın tahmini üzerinde büyük bir etkiye sahip olduğunu göstermektedir⁹⁵.

Vektör $\hat{\beta}$ 'nın tamamı üzerinden çıkarılan i . gözlemin etkilerini özetleyen Doğrusal Regresyon Modelinde (Linear Regression Model-LRM) Cook's uzaklığına denk gelen ikinci ölçüm

$$C_i = (\Delta_i \hat{\beta})' \widehat{Var}(\hat{\beta}) (\Delta_i \hat{\beta}) = \frac{r_i^2 h_{ii}}{(1 - h_{ii})^2}$$

şeklinde ifade edilmektedir.

Tek bir gözlemin etkisinin diğer bir ölçüsü de i . gözlem çıkartıldığı zaman χ^2 'deki değişim değeri ile hesaplanmaktadır.⁹⁶

$$\Delta_i \chi^2 = \frac{r_i^2}{1 - h_{ii}}$$

Uygunluğun Sayısal Ölçümleri

Modelin uyum iyiliğinin tek bir değer ile özetlenmesi, değişkenlerin tek tek test edilmesinden daha iyi olabilmektedir. Bu tür ölçümler, alternatif modellerin karşılaştırılmasında yardımcı olmaktadır. Doğrusal Regresyon Modeli'nin (LRM) standart uygunluk ölçüsü, belirlilik katsayısı R^2 iken, kategorik sonuçların modelleri için kesin bir tercih yoktur. LRM'deki R^2 'nin benzerlerini

⁹⁵ D. Pregibon, **a.g.e.**, pp. 705–724.

⁹⁶ J. Scott Long, **a.g.e.**, p.101.

oluşturmak için çok fazla girişim vardır. Ancak hiçbir ölçümün açık üstünlüğü veya açıklanan değişkenlik açısından net bir yorumlama avantajı bulunmamaktadır. Kategorik ve sınırlı bağımlı değişken modelleri için uygunluğun birçok sayısal ölçüsü LRM içindeki belirlilik katsayısı R^2 'den yaklaşık olarak yapılandırılmıştır.

Kategorik bağımlı değişkenli modellerin uygunluk ölçümleri için birçok *Pseudo (Yaklaşık) - R²*'ler önerilmiştir. Aldrich ve Nelson (1984), *Pseudo - R²* değerini tam model için hesaplanan Ki-kare istatistik değerini kullanarak,

$$R^2 = \frac{\chi^2}{\chi^2 + n}$$

şeklinde hesaplamayı önermişlerdir⁹⁷.

Efron's (1978) iki alternatifli sonuçlar için \hat{y} değeri yerine $\hat{P} = \widehat{Pr}(y|x)$ değeri kullanıldığında *Pseudo - R²* değerini; LRM'de açıklanan değişkenliğin yüzdesi şeklinde hesaplanan R^2 değerine benzer şekilde kullanmıştır.

$$R_{Efron}^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{P}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$$

İki alternatifli sonuçlar için $\sum_{i=1}^N (y_i - \bar{y})^2 = (n_0 n_1)/n$ olarak tanımlanır. n_0 , örneklemdeki "0"ların sayısını, n_1 ise örneklemdeki "1"lerin sayısını ifade etmektedir⁹⁸.

McFadden (1973) ise En Çok Olabilirlik yaklaşımı ile tahmin edilen herhangi bir modelde uygulanabilecek LRM içinden, açıklanan değişim yüzdesine benzer farklı bir yol önermiştir. Bu popüler ölçüm aynı zamanda "olabilirlik oran

⁹⁷ John H. Aldrich, Forrest D. Nelson, **a.g.e.**, p. 57.

⁹⁸ Bradley Efron, "Regression and ANOVA with Zero-One Data: Measures of Residual Variation", **Journal of the American Statistical Association**, Vol: 73, 1978, pp. 113–121.

indeksi” olarak da adlandırılmaktadır. Bu ölçüm içinde bağımsız değişkenler olmaksızın modellenen M_α ’nın log-olabilirlik değeri, toplam kareler toplamı olarak; bağımsız değişkenler ile modellenen M_β ’nin log-olabilirlik değeri ise hata kareler toplamı olarak düşünülebilir,

$$R_{MCF}^2 = 1 - \frac{\ln \hat{L}(M_\beta)}{\ln \hat{L}(M_\alpha)}.$$

Eğer $M_\alpha = M_\beta$ ise R_{MCF}^2 değeri “0”a eşittir. Fakat R_{MCF}^2 değeri asla “1” olmaz. Yeni değişkenler modele ilave edildikçe R_{MCF}^2 değeri artmaktadır.

Bunu düzenlemek için Ben-Akiva ve Lerman, modeldeki parametre sayısı için düzeltilmiş R_{MCF}^2 , \bar{R}_{MCF}^2 ’yi önermiştir.

$$\bar{R}_{MCF}^2 = 1 - \frac{\ln \hat{L}(M_\beta) - K}{\ln \hat{L}(M_\alpha)}$$

Modele ilave edilen her parametre için, $\ln \hat{L}(M_\beta)$ değeri “1”den daha büyük olduğu zaman \bar{R}_{MCF}^2 değeri artmaktadır⁹⁹. Cramer (1999) da Ben-Akiva ve Lerman tarafından dikkate alınmayan dengesiz (sonuç değişkenin aldığı alternatif değerlerin sayısı bakımından) örneklemde; genellikle sıklığı az olan bağımlı değişkeninin çok kötü tahmin edildiği durumları göz önüne alarak düzeltilmiş ölçüm önermiştir¹⁰⁰.

LRM’de belirlilik katsayısını hesaplamının başka bir yolu da parametre değerlerinin En Çok Olabilirlikle hesaplandığı model için Olabilirlik Oran dönüşümünü kullanmaktır. Buradan yola çıkarak M_α ’yı sadece sabiti olan model

⁹⁹ J. Scott Long, **a.g.e.**, p. 104.

¹⁰⁰ J. Cramer, “Predictive Performance of the Binary Logit Model in Unbalanced Samples”, **Journal of the Royal Statistical Society**, Series D, Vol: 88, 1990, pp. 85–94.

ve M_β 'yi da sabit ve bağımsız değişken katsayılarını içeren model olarak tanımlayıp *Pseudo* – R^2 şu şekilde hesaplanabilir:

$$R_{ML}^2 = 1 - \left[\frac{L(M_\alpha)}{L(M_\beta)} \right]^{2/N}$$

Maddala (1983), yukarıdaki eşitliği $G^2 = -2\ln [L(M_\alpha)/L(M_\beta)]$ değerini alacak şekilde aşağıdaki gibi ifade etmiştir;¹⁰¹

$$R_{ML}^2 = 1 - \exp(-G^2/N)$$

Diğer bir ifadeyle, uygunluk ölçüsü olarak gösterilen R^2 ve çeşitli *Pseudo* – R^2 'ler genellikle hipotez testleri ile yakından ilişkilidir. Magee (1990), Wald ve skor testlerine dayanan uygunluk ölçülerini,¹⁰²

$$R^2 = \frac{W}{n + W} \equiv R_{W}^2, \quad R^2 = 1 - \exp(-LR/n) \equiv R_{LR}^2$$

olarak bulmuştur. McFadden'ın haricinde, hesaplanan *Pseudo* – R^2 'yi uygulamalarında kullanan bulunmamaktadır. Bunun nedeni, *Pseudo* – R^2 'nin “0” ve “1” dışındaki değerleri için açık bir yorumlanma imkanı olmamasıdır. Bu değer sadece belirlilik katsayısı büyüdükçe modelin uygunluğunun gelişmekte olduğunu göstermektedir.

Sınıflama Tabloları

Kategorik sonuçların gözlemlenen ve beklenen değerlerini karşılaştırmak, modellerin uyum iyiliğini değerlendirmede diğer bir yaklaşımdır. Gözlemlenen

¹⁰¹ G. S. Maddala, **Limited-Dependent and Qualitative Variables in Econometrics**, Cambridge University Press, 1983, p.39.

¹⁰² Lonnie Magee, “ R^2 Measures Based on Wald and Likelihood Ratio Joint Significance Tests”, **The American Statistician**, Vol: 44, 1990, pp. 250–253.

ve beklenen değerlere dayanan sınıflandırma tablolarında doğru tahmin oranları “*Count (Sayılan) – R²*” ile ifade edilmektedir.

$$R_{Count}^2 = \frac{1}{N} \sum_j n_{jj}$$

En büyük satır marjına göre R_{Count}^2 değerinin

$$R_{AdjCount}^2 = \frac{\sum_j n_{jj} - \max_r(n_{r+})}{N - \max_r(n_{r+})}$$

şekilde düzeltilmesi gerekmektedir. n_{r+} , r . satırın marjinal değerini; $\max_r(n_{r+})$ ise satır marjı en büyük olanı belirtmektedir¹⁰³.

Bilgi Ölçüleri

Bilgi ölçüleri, İç İçe Modellerin karşılaştırılmasında kullanışlı olan ve en çok bilinen Akaike Bilgi Kriteri (Akaike Information Criteria-AIC) ile popülaritesi artmakta olan Bayes Bilgi Kriteridir (Bayesian Information Criteria-BIC).

Akaike 1973 yılında bilgi kriterini şu şekilde tanımlamıştır:¹⁰⁴

$$AIC = \frac{-2\ln\hat{L}(M_\beta) + 2P}{n}$$

P ; modeldeki parametrelerin sayısı; $\hat{L}(M_\beta)$, verilerin modeldeki olabilirliklerini gösterir ki, büyük değer daha iyi uyum anlamına gelmektedir. $-2\ln\hat{L}(M_\beta)$ $[0, +\infty)$ arasında değer almakta ve küçük değerler daha iyi uyum olduğunu göstermektedir. Modeldeki parametre sayıları arttığında, $-2\ln\hat{L}(M_\beta)$ küçülür. AIC’ye $2P$ ’nin ilave edilmesi parametre sayısındaki artış için bir ceza değeridir.

¹⁰³ J. Scott Long, **a.g.e.**, p. 106–108.

¹⁰⁴ H. Akaike, “Information Theory and an Extension of the Maximum Likelihood Principle”, **Perspective in Statistics Selected Paper of Hirotugu Akaike**, Springer, 1998, p. 209.

Gözlemlerin sayısı $-2\ln\hat{L}(M_\beta)$ değerini etkilediği için düzeltilmiş $-2\ln\hat{L}(M_\beta)$, her bir gözlemin katkısını elde etmek amacıyla örnek büyüklüğüne bölünmektedir.

AIC genellikle, farklı örnek büyüklüklerini veya Olabilirlik Oran (LR) testleri ile karşılaştırılmayacak, İç İç Modelleri karşılaştırmada kullanılmaktadır. Diğer her şey aynı iken, küçük AIC değerli model en uygun model olarak kabul edilmektedir¹⁰⁵.

Bayes Bilgi Kriteri (BIC) 1995 yılında Raftery tarafından bir modelin uygunluğunu kapsamlı bir şekilde değerlendirmek ve hem İç İç (Nested) hem de İç İç Olmayan (Nonnested) Modellerin karşılaştırılmasına olanak vermek amacıyla önerilmiştir.

BIC istatistiği, Bayes faktör tahmini kullanılarak hesaplanması uygun bir yaklaşımdır. BIC ölçüsü, model M_1 model M_2 içinde yer alıyorsa yani İç İç Modeller ise şu şekilde hesaplanmaktadır,

$$BIC_{12} = \chi_{12}^2 - df_{12} \log n$$

Model M_1 'e göre model M_2 karşılaştırılırken hesaplanan standart Olabilirlik Oran test istatistik değeri χ_{12}^2 , $df_{12} = df_2 - df_1$ ile testin serbestlik derecesine eşittir.

Genellikle modeller karşılaştırılırken, temel model olarak M_D ile gösterilen Doymuş (Saturated) Model ya da sadece sabit terimin yer aldığı M_α modeli kullanılmaktadır.

Temel model olarak Doymuş Model (M_D) alınır, model M_k için hesaplanan BIC değeri BIC_k ile gösterilmektedir ve

$$BIC_k = D_k - df_k \log n$$

¹⁰⁵ J. Scott Long, **a.g.e.**, p. 109.

ifadesine eşittir. Burada D_k , model M_k için sapma değeri; df_k ise sapma istatistiğine ait serbestlik derecesidir. Doymuş Model'e ait BIC değeri "0" a eşit olduğu için, $BIC_k > 0$ olduğunda Doymuş Model kullanılmaktadır. $BIC_k < 0$ olduğunda ise, veri seti için daha uygun olan, model M_k tercih edilmektedir.

Temel model olarak sadece sabitin yer aldığı M_α modeli alındığında ise, BIC yerine BIC'_k kullanılır. $LR(M_k)$ 'nin diğer hesaplama yolunda LR 'deki df'_k , modeldeki bağımsız değişken sayısı ile Ki-kare dağılımına uymaktadır.¹⁰⁶

$$BIC'_k = -LR(M_k) - df'_k \log n$$

$LR(M_k)$, sadece sabit değer α 'nın olduğu model M_α 'ya karşı model M_k 'nin Olabilirlik Oran test istatistik değeridir. df'_k , model M_k 'daki bağımsız değişken sayısıdır. $BIC'_\alpha=0$ yokluk hipotezi altında, $BIC'_k > 0$ olduğu zaman M_α modeli, $BIC'_k < 0$ olduğu zaman ise M_k modeli geçerli olmaktadır.

Bayes Bilgi Kriteri olarak hesaplanan BIC veya BIC' değerlerinden hangisinin kullanılacağına, model M_α 'ya karşı diğer model karşılaştırılırken, "sapma istatistik değeri"nin mi yoksa "olabilirlik oran test istatistik değeri"nin mi hesaplandığına göre karar verilmektedir. Eğer sapma değeri kullanılıyorsa BIC , olabilirlik oran değeri kullanılıyorsa BIC' değeri kullanılmalıdır.

M_j ve M_k gibi herhangi iki model karşılaştırılırken, BIC_{jk} değeri iki modelin BIC değerlerinin farkı alınarak hesaplanmaktadır.

$$BIC_{jk} = BIC_k - BIC_j$$

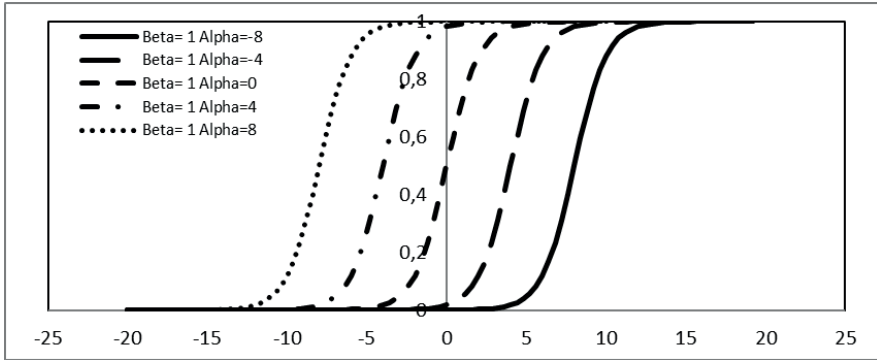
Model seçiminde BIC değeri küçük olan model; yani $BIC_k - BIC_j < 0$ ise model M_k , $BIC_k - BIC_j > 0$ ise model M_j tercih edilmektedir.¹⁰⁷

¹⁰⁶ J. Scott Long, **a.g.e.**, p. 111.

¹⁰⁷ Adrian E. Raftery, "Bayesian Model Selection Social Research", **Sociological Methodoloji**, Vol: 25, 1995, pp. 111–165.

1.2.6 Model Parametrelerinin Yorumu

Lojistik regresyon eğrisi S şeklindedir. Ancak “sabit” ve “eğim (bağımsız değişken)” katsayılarının farklı değer çiftleri için farklı lojistik regresyon eğrileri oluşmaktadır. Grafik olarak modeldeki sabit ve bağımsız değişkenlere ait parametre katsayılarının, olasılık değeri üzerindeki etkisini görmek için Lojistik Modeli tek bağımsız değişkenli $P(y = 1|x) = F(\alpha + \beta x)$ olacak şekilde ifade ettiğinde, olasılık eğrisi üzerindeki “sabit”in etkisi Şekil 1.1’de görülmektedir. Burada eğim katsayısı β ’ya sabit bir değer verildiğinde, α değeri büyüdükçe lojistik regresyon eğrisinin sola, küçüldükçe de sağa doğru yer değiştirdiği görülmektedir.

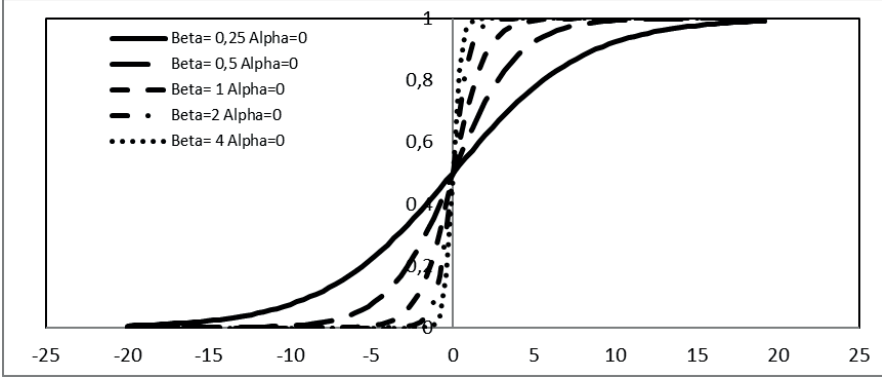


Şekil 1.1 Modeldeki “sabit” in olasılık üzerindeki etkisi.¹⁰⁸

Şekil 1.2’de ise modeldeki eğim katsayısındaki değişimin etkisi görülmektedir. Modelde α ’nın olmadığı durumlar için, β ’nin aldığı değer küçüldükçe lojistik regresyon eğrisi esnemektedir. Örneğin, $\beta = 0.25$ olduğunda, x ’in -20 ile +20 arasında muntazam bir şekilde arttığı görülürken, $\beta = 0.5$ için eğrideki değişimin hızı başlarda daha yavaşken x değeri “0”a yaklaştıkça eğimdeki değişimin hızının arttığı gözlemlenmektedir.¹⁰⁹

¹⁰⁸ J. Scott Long, *a.g.e.*, pp. 62–63.

¹⁰⁹ *A.e.*, p. 63.



Şekil 1.2 Modeldeki “eğim” katsayısının olasılık eğrisi üzerindeki etkisi.¹¹⁰

İki Kategorili Lojistik Regresyon Modelinde birden fazla bağımsız değişkenin olması bu değişkenlere ait parametrelerin ilgilenilen olayın olasılığı üzerindeki etkilerini grafikte görmeyi güçleştirecektir.

Doğrusal Olasılık Modelinde, “diğer bağımsız değişkenlerin değeri sabit¹¹¹ olduğu varsayımı altında”* x_k ’daki bir birimlik değişimin $P(y = 1)$ olasılığı üzerindeki etkisi $\hat{\beta}_k$ kadar olacaktır ve x_k ’nın tüm değerleri için bu etki sabittir. Lojit Model’de $P(y = 1)$ ile her bir x_k arasındaki doğrusal olmayan ilişki nedeniyle, $P(y = 1)$ üzerindeki x_k ’daki değişimin etkisini yorumlamak çok kolay olmamaktadır. Bu etkiyi yorumlamanın bir yolu, eğer ilgilenilen x_k bağımsız değişkeni nicel ise bu değişkendeki değişimin, $P(y = 1)$ olasılığı üzerindeki değişim oranını ölçmek için kısmi türevini almaktır.

¹¹⁰ J. Scott Long, a.g.e., p. 63.

¹¹¹ J. Scott Long, Simon Cheng, a.g.e., p. 260.

* Modeldeki parametre katsayılarının yorumunda, herhangi bir bağımsız değişkendeki değişimin olasılık değeri üzerindeki etkisinden bahsedildiğinde, her zaman diğer değişkenlerin modele sabit bir değer olarak alındığı kabul edilmektedir. Sabit değerleri elde etmenin iki yolu vardır; ya ilgilenilen değişkenler dışındaki bağımsız değişkenlerin değerlerinin kendi ortalamaları modele sabit değer olarak alınmakta veya doğrudan her bir gözlemin ilgilenilen değişken dışındaki diğer değişkenlerin değerleri tek tek modelde yerine konularak marjinal değişimler hesaplanıp sonra tüm gözlemler için ortalaması hesaplanmaktadır.

$$\frac{\partial P(y = 1|x)}{\partial x_k} = \frac{\exp(\beta_0 + \sum_{k=1}^p \beta_k x_k)}{1 + \exp(\beta_0 + \sum_{k=1}^p \beta_k x_k)} \frac{1}{1 + \exp(\beta_0 + \sum_{k=1}^p \beta_k x_k)} \hat{\beta}_k$$

$$= P(y = 1|x)[1 - P(y = 1|x)]\hat{\beta}_k$$

Yukarıdaki formülde $\hat{\beta}_k$ değeri çarpımsal faktör olarak eşitlikte yer almaktadır ve $\hat{\beta}_k$ 'nin işareti x_k 'daki değişimin $P(y = 1)$ üzerindeki etkisinin yönünü tayin etmektedir. Etkinin büyüklüğü ise x_k 'ların tümünü içine alan $(\beta_0 + \sum_{k=1}^p \beta_k x_k)$ değerinin büyüklüğüne bağlı olmaktadır. Doğrusal olmayan $(\beta_0 + \sum_{k=1}^p \beta_k x_k)$ değeri x_k 'daki değişimin $P(y = 1)$ üzerindeki etkisinin değerini düşürmektedir. $(\beta_0 + \sum_{k=1}^p \beta_k x_k) = 0$ iken etkinin değerindeki düşüş minimumdur. Bu değer “0”dan (-) veya (+) tarafa uzaklaştıkça etkinin değerindeki düşüş de artmaktadır.¹¹²

x_k bağımsız değişkeni nitel ve “0” ile “1” olmak üzere iki değer alan bir kukla değişken ise; x_k 'daki değişimin $P(y = 1|x)$ üzerindeki etkisi şu şekilde hesaplanmaktadır.

$$\frac{\Delta P(y = 1|x)}{\Delta x_k} = P(y = 1|x, x_k = 1) - P(y = 1|x, x_k = 0)$$

Parametre katsayılarının yorumlanmasında bir başka yol da önce bağımlı ve bağımsız değişkenler arasındaki fonksiyonel ilişkinin tanımlanarak bağ fonksiyonunun oluşturulması ve sonra da her bir bağımsız değişkendeki bir birimlik değişikliğin bu tanımlanan bağ fonksiyonu üzerindeki etkisinin uygun bir şekilde ifade edilmesidir¹¹³. Lojistik Regresyon Modeli şu şekilde ifade edilmektedir,

$$P(y = 1|x) = \frac{\exp(\beta_0 + \sum_{k=1}^p \beta_k x_k)}{1 + \exp(\beta_0 + \sum_{k=1}^p \beta_k x_k)}$$

¹¹² John H. Aldrich, Forrest D. Nelson, **a.g.e.**, pp.41–44.

¹¹³ David W. Hosmer, Stanley Lemeshow, **a.g.e.**, pp. 47–48.

Bu ifade ile Doğrusal Olmayan Olasılık Modeli'nde, olasılık değerleri odds şekline dönüştürülebilmektedir. Lojistik regresyon katsayılarının düzgün yorumlanabilmesi için “odds” kavramının iyi anlaşılması gerekmektedir. İki Kategorili Lojistik Regresyon Modeli için odds değerini $\Omega(x)$ ile gösterildiğinde,

$$\Omega(x) = \frac{P(y = 1|x)}{1 - P(y = 1|x)} = e^{\beta_0 + \sum_{k=1}^p \beta_k x_k} = e^{\beta_0} \prod_{k=1}^p e^{\beta_k x_k}$$

eşitliğin sol tarafı odds, sağ tarafı da $\exp(\beta_k)$ üsteli ile odds üzerindeki x_k 'nin marjinal etkisini vermektedir. Odds değeri olasılık değerinden farklıdır. Olasılık $[0, 1]$ aralığında değerler alırken, odds değeri $[0, +\infty)$ aralığında değerler almaktadır.¹¹⁴

Lojistik Regresyon Modeli'nde katsayıların yorumlanmasında diğer önemli bir kavram da Odds Oranı (Odds Ratio-OR)'dir. OR, yorumlamadaki kolaylığı nedeniyle Lojistik regresyonda genellikle ilgilenilen bir parametredir. Odds değeri $\Omega(x, x_k) = \exp(x_k \beta_k) = e^{\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \dots + \beta_p x_p}$ olarak tanımlandığında OR değeri, örneklemdaki herhangi bir gözlemin, modeldeki diğer değişkenlerin sabit olduğu varsayımı altında, bağımsız değişken x_k 'daki δ birimlik artışın ilgilenilen olayın ortaya çıkmasını $\exp(\beta_k \delta)$ kadar daha mümkün kıldığını ifade etmektedir.¹¹⁵

$$OR = \frac{\Omega(x, x_k + \delta)}{\Omega(x, x_k)} = \exp(\beta_k \delta)$$

Buradaki sorun, OR tahmininin (\widehat{OR}) çarpık bir dağılıma sahip olmasıdır. Bu sorunun kaynağı, OR parametre değerinin “1”e eşitliği test edilirken, odds değerlerinin $[0, +\infty)$ arasında değerler alıyor olmasıdır. Teoride yeterince büyük örneklem için \widehat{OR} 'nin dağılımı normaldir, ancak birçok çalışma, \widehat{OR} değerlerinin

¹¹⁴ Raymond E. Wright, **a.g.e.**, pp. 222–224.

¹¹⁵ J. Scott Long, **a.g.e.**, p.80.

Normal dağılıma uygun olmasını sağlayacak örnek büyüklüğüne sahip değildir. Bu nedenle çıkarımlar, genellikle çok daha küçük olan örneklem büyüklüğü için Normal dağılıma sahip olma eğiliminde olan $\ln(\widehat{OR})$ 'nin örneklem dağılımına dayanmaktadır.¹¹⁶ Buradan OR parametre değeri için nokta tahmini yapılabildiği gibi $\ln(\widehat{OR}) = \hat{\beta}_k$ için $\%(1 - \alpha)$ düzeyinde uç değer tahmini de yapılabilmektedir.

$$\exp[\hat{\beta}_k \pm z_{1-\alpha/2} \widehat{SE}(\hat{\beta}_k)]$$

Bu aralık “1” değerini içermiyorsa aralık değerlerine bakılıp örnekleme ait OR 'nin sağa ya da sola çarpık olduğuna karar verilir.

1.3 Lojistik Regresyon Modellerinin Örnek Büyüklüğü

Lojistik Regresyon Modelleri için örnek büyüklüğü belirlemekle ilgili literatürde çok az çalışma bulunmaktadır. Lojit Modellerin değerlendirilmesinde temel unsurlardan biri, her bir bağımsız değişken için yeterli sayıda olayın gerçekleşmesinin beklenmesidir. Örnek büyüklüğünü belirlemek için yöntemlerden bahsederken, aynı zamanda her bir bağımsız değişkendeki olayların yeterli sayıda olmasının öneminin de farkında olmak oldukça önemlidir.

Örnek büyüklüğü belirlenirken cevaplanması gereken “Herhangi bir parametre katsayısının sıfıra eşit olduğunu ileri süren sıfır hipotezine karşı bu katsayının belli bir değere eşit olduğunu ifade eden karşı ön savı test etmek için ne kadar örnek büyüklüğüne gereksinim olduğu?” sorusudur. İki kategorili tek bağımsız değişkenli Lojistik Regresyon Modellerinde örnek büyüklüğünün hesaplanması

¹¹⁶ David W. Hosmer, Stanley Lemeshow, **a.g.e.**, p. 52.

için 1998'de Hsieh ve arkadaşları tarafından önerilen formülasyon şu şekildedir;¹¹⁷

$$n = \frac{(Z_{1-\alpha/2} + Z_{1-\theta})^2}{P_0(1 - P_0)\beta^{*2}}$$

Burada n , gereken örnek büyüklüğü; $\beta^* = \Delta/\sigma$ ifadesindeki Δ , hesaplanan iki grup ortalamasının farkıdır ve P_0 , standartlaştırılmış X 'in veya $X = 0$ iken olasılık değeri ve $(1 - \theta)$, testin gücü iken; $Z_{1-\theta}$ ve $Z_{1-\alpha/2}$ sırası ile θ ve $\alpha/2$ 'in Standart Normal dağılım tablosundaki kuyruk olasılık değerleridir.

Bunun dışında, bağımlı değişkenin iki kategorisine ait oran eşitliğinin test edilmesine dayanan geleneksel örnek büyüklüğü yöntemi kullanılabileceği gibi alternatif bir yol da 1981 yılında Whittemore tarafından önerilmiştir.¹¹⁸

İki kategorili tek bağımsız değişkenli Lojistik Regresyon Modeli için oran eşitliğinin test edilmesine dayanan örnek büyüklüğü, sıfır hipotezi $H_0: P_0 = P_1$ 'in α anlamlılık düzeyindeki alternatif hipotez $H_a: P_0 \neq P_1$ olacak şekilde, testin gücü $1 - \theta$ için aşağıdaki eşitlik tarafından hesaplanabilmektedir,¹¹⁹

$$n = \frac{\left(z_{1-\alpha} \sqrt{2\bar{P}(1-\bar{P})} + z_{1-\theta} \sqrt{P_0(1-P_0) + P_1(1-P_1)} \right)^2}{(P_1 - P_0)^2}$$

Burada $\bar{P} = (P_0 + P_1)/2$ ve $z_{1-\alpha}$ ve $z_{1-\theta}$ sırası ile Standart Normal dağılımda α ve θ yüzde değerinin üst kısmını göstermektedir yani Standart Normal dağılım

¹¹⁷ F. Y. Hsieh, Daniel A. Bloch, Michael D. Larsen, "A Simple Method of Sample Size Calculation for Linear and Logistic Regression", **Statistics in Medicine**, Vol:17, 1998, pp. 1623-1634.

¹¹⁸ Alice S. Whittemore, "Sample Size for Logisitic Regression with Small Response Probability", **Journal of the American Statistical Association**, Vol: 76, 1981, pp. 27-32.

¹¹⁹ David W. Hosmer, Stanley Lemeshow, **a.g.e.**, p. 339.

tablosundaki kuyruk olasılık değerleridir. Yukarıdaki örnek büyüklüğü bağımlı değişkenin her bir grubu için olması gereken örnek büyüklüğü değeridir.

Whitemore çalışmasında örnek büyüklüğünü bulurken, farklı anlamlılık düzeyleri ve testin güç düzeyleri için bağımsız değişkenin Normal, Üstel, Poisson ve Bernoulli dağılıma uyduğu durumlar için ayrı ayrı örnek büyüklüğü değerleri hesaplamıştır¹²⁰. İki kategorili tek bağımsız değişken içeren Lojistik Regresyon Modelinde bağımsız değişkenin Bernoulli dağılımına uyduğu durumlar için (x , “0” ve “1” olarak kodlanmaktadır) hipotez testleri $H_0: \beta_1 = 0$ 'a karşı $H_a: \beta_1 = \beta_1^*$ için gereken toplam örnek büyüklüğü şu şekilde hesaplanmaktadır,¹²¹

$$n = (1 + 2P_0) * \frac{\left(z_{1-\alpha} \sqrt{\frac{1}{1-\pi} + \frac{1}{\pi}} + z_{1-\theta} \sqrt{\frac{1}{1-\pi} + \frac{1}{\pi e^{\beta_1^*}}} \right)^2}{P_0 \beta_1^{*2}}$$

$\pi = P(X = 0)$. Buradan hesaplanan örnek büyüklüğü bağımlı değişkenin her grubu için eşit olacak şekilde yuvarlanmaktadır.

Whitemore ayrıca, bağımsız değişkenin nicel olduğu durumlar için örnek büyüklüğünü hesaplarken, bağımsız değişkenin ortalaması “0” ve varyansı “1” olacak şekilde standartlaştırılmış değerlerin kullanılmasını önermiştir. α anlamlılık düzeyi ve $1 - \theta$ testin gücü ile tek taraflı test değerinin kullanılarak $H_0: \beta_0 = 0$ karşı $H_a: \beta_1 = \beta_1^*$ hipotezinin testini yapabilmek için örnek büyüklüğü şu şekilde hesaplanmaktadır,

$$n = (1 + 2P_0 \delta) * \frac{\left(z_{1-\alpha} + z_{1-\theta} e^{-0.25 \beta_1^{*2}} \right)^2}{P_0 \beta_1^{*2}}$$

¹²⁰ Alice S. Whitemore, **a.g.e.**, pp. 27–32.

¹²¹ David W. Hosmer, Stanley Lemeshow, **a.g.e.**, p. 341.

Bu eşitlikte $\delta = \frac{1+(1+\beta_1^{*2})e^{1.25\beta_1^{*2}}}{1+e^{-0.25\beta_1^{*2}}}$ ve $P_0 = \frac{e^{\beta_0}}{1+e^{\beta_0}}$ yani P_0 standartlaştırılmış bağımsız değişkenin beklenen değerinin lojistik olasılığı olarak hesaplanan değeridir.

Hsieh ve arkadaşlarının 1998 yılında yaptıkları çalışmadaki Çok Değişkenli Lojistik Regresyon için önerdikleri örnek büyüklüğü yukarıdaki denkleme uyarlandığında aşağıdaki eşitlik elde edilmektedir,

$$n = \frac{(1 + 2P_0\delta)}{1 - \rho^2} * \frac{(z_{1-\alpha} + z_{1-\theta}e^{-0.25\beta_1^{*2}})^2}{P_0\beta_1^{*2}}$$

ρ^2 , ilgilenilen değişkenin (örneğin, x_k), modelde yer alan diğer $p - 1$ değişkene göre hesaplanan kısmi korelasyon katsayısının karesidir.¹²²

2010'da Alam v.d.¹²³ tarafından yapılan çalışmada, Whittmore (1981) ile Hsieh v.d. (1998) tarafından önerilen örnek büyüklüğü hesaplama yöntemlerinin birleşiminden elde edilen yeni bir yöntem önerilmiştir. Bu simülasyon çalışmasında, bağımsız değişken Bernoulli dağılımına uyduğu durumlarda, Hsieh ve arkadaşlarının uyarladığı yöntemle yeni önerilen yöntemin eşit düzeyde iyi sonuç verdiği elde edilmekle birlikte güç bakımından yeni yöntemin daha iyi sonuçlar verdiği gözlemlenmiştir.

Lojistik Regresyon Modellerinde regresyon katsayılarının tahminlerinin güvenilirliğini sağlamak için her bir bağımsız değişken için kaç tane birime gerek olduğu konusu 1996'da Peduzzi v.d. tarafından Monte Carlo çalışması olarak incelenmiştir¹²⁴. Söz konusu çalışmada, her bir parametre için birim sayısı

¹²² F. Y. Hsieh, Daniel A. Bloch, Michael D. Larsen, **a.g.e.**, pp. 1623–1634.

¹²³ M. Khorshed Alam, M. Bhaskara Rao, Fu-Chih Cheng, "Sample Size Determination in Logistic Regression", **The Indian Journal of Statistics**, Vol: 72-B, 2010, pp. 58-75.

¹²⁴ P. N. Peduzzi, J. Concato, E. Kemper, T. R. Holford, A. Feinstein, "A Simulation Study of the Number of Events per Variable in Logistic Regression Analysis", **Journal of Clinical Epidemiology**, Vol: 49, 1996, pp. 1373–1379.

2, 5, 10, 15, 20 ve 25 olacak şekilde simülasyon çalışması yapılmıştır. Yapılan çalışma sonucunda, her bir parametre için birin sayısının en az 10 olmasının, varyansın yüksek ya da düşük tahmin edilmesinden dolayı, Wald tarafından önerilen katsayıların testlerinde meydana gelebilecek problemleri ortadan kaldırmak için yeterli olduğu gösterilmiştir. Veri setinin yeterli olması için modeldeki parametre sayısı $p + 1 \leq \min(n_1, n_0) / 10$ 'dan daha çok olmamalıdır. Bununla beraber, model temelli tahminlerin performansı genellikle değişkenlerin birim sayısı yerine toplam örnek büyüklüğü tarafından belirlenmektedir.¹²⁵

¹²⁵ David W. Hosmer, Stanley Lemeshow, **a.g.e.**, p. 343.

BÖLÜM 2

DİSKRİMİNANT ANALİZİ

Bağımlı değişkenin nitel, bağımsız değişkenlerin çoğunun nicel olduğu durumlar için kullanılması uygun olan çok değişkenli istatistik yöntemlerden biri de Diskriminant Analizi'dir.

İlk olarak 1936'da Fisher¹²⁶ tarafından iki grubu sınıflamak için geliştirilmiş olan Diskriminant Analizi, 1948 yılında Rao¹²⁷ tarafından ikiden fazla grup için geliştirilmiştir. 1949'da Hoel ve Peterson¹²⁸ Diskriminant Analizi'nin, grupların farklarının kaynağını saptamak ve birimlerin hangi gruba ait olduğuna karar vermek gibi iki temel amaç için kullanıldığına dikkat çekmişlerdir. Sonrasında yapılan bazı çalışmalarda Diskriminant Analizi, uygulama amacına göre farklı isim almaktadır. Eğer çalışmada birbirinden farklı iki veya daha fazla grup üyesini en iyi ayırt etme yolu ile değişkenlerin tanımlanması ve fonksiyonun elde edilmesi amaçlanıyorsa *Tanımlayıcı Diskriminant Analizi* (Descriptive Discriminant Analysis), amaç hangi gruptan geldiği bilinmeyen birimlerin ait olduğu grubun saptanması ise *Tahmin Edici Diskriminant Analizi* (Predictive Discriminant Analysis) adı verilmektedir.¹²⁹ Bununla birlikte birçok çalışmada

¹²⁶ R. A. Fisher, "The Use of Multiple Measurements in Taxonomic Problems", *Annals of Eugenics*, 7(2), 1936, pp. 466–475.

¹²⁷ C. R. Rao, "The Utilization of Multiple Measurements in Problem of Biological Classification", *Journal of the Royal Statistical Society, Series B (Methodological)*, 1948, Vol. 10, No. 2, pp. 159–203.

¹²⁸ P. G. Hoel, R. P. Peterson, "A Solution to the Problem of Optimum Classification", *The Annals of Mathematical Statistics*, 1949, Vol. 20, No. 3 (Sep.), p. 433.

¹²⁹ A. Pedro Durate Silva, Antonie Stam, *Reading and Understanding Multivariate Statistics*, Edited by Laurence G. Grimm, Paul R. Yarnold, American Psychological Association Washington, DC., 1994, p. 277.

bu iki amaç iç içe girmiş durumda olup birlikte ele alınmaktadır. Bu çalışmada Diskriminant Analizi denildiğinde bu iki durumun da birlikte ele alındığı analizden bahsedilmektedir.

Kısaca Diskriminant Analizi, bağımsız değişkenler açısından birbirleri içine girmiş ortak özelliklere sahip grupları birbirinden ayırmak için grup ortalama vektörlerini kullanarak fonksiyonlar geliştiren ve geliştirdiği fonksiyonların sınıflandırmadaki gücünü test eden bir yöntemdir¹³⁰.

2.1 İki Kategorili Diskriminant Analizi

İki Kategorili Diskriminant Analizi'nde, iki grup arasındaki ayrımı en iyi yapacak faktörlerin doğrusal bileşimini belirlemek amacıyla diskriminant fonksiyonunun yazılması amaçlanmaktadır. Diskriminant fonksiyonunu yazarken, gruplararası değişimin grupları içi değişime oranından elde edilen özdeğer veya karakteristik kök (λ) değerini maksimum yapacak fonksiyonun yazılmasına çalışılmaktadır.¹³¹

Tüm Diskriminant Analizi problemleri aşağıdaki dört ortak belirleyici özellik ile tanımlanmaktadır:¹³²

- Diskriminant Analizi'nde incelenen birimler iki gruptan sadece birine ait olması ve grupların da bağıdaşmaz ve bağımsız olması gerekmektedir.
- Her bir Diskriminant Analiz grubu iyi tanımlı olmalı ve gruplar birimler arasındaki gerçek farkları yansıtmalıdır. Ayrıca gruplar ikiden fazla ve sıralı ise Diskriminant Analizi yerine bu durum için geliştirilen Sıralı Lojit Analiz tekniğinin kullanılması tercih edilmelidir.

¹³⁰ Kazım Özdamar, **Paket Programlar ile İstatistiksel Veri Analizi**, Cilt 2, Yenilenmiş 5. Baskı, 2004, s. 355.

¹³¹ Subhash Sharma, **Applied Multivariate Techniques**, John Wiley & Sons, Inc., 1996, p. 241.

¹³² A. Pedro Durate Silva, Antonie Stam, Edited by Laurence G. Grimm, Paul R. Yarnold, **a.g.e.**, p. 277.

- Diskriminant Analizi'nde, gruplar veri seti toplanmadan önce tanımlanmış olmalıdır. Bunun zıttı durumlar için, veri setinin belirleyici özelliklerine dayanarak grupları belirlemede ise Kümeleme Analizi kullanılmaktadır.
- Son olarak da bağımsız değişkenler seti mümkün olduğunca tam ve birimleri tanımlamada doğru ayırım yapacak şekilde oluşturulmalıdır. Ancak bazı belirleyici özellikler bağımlı değişkeni açıklamada anlamlı iken, grupları ayırma veya sınıflandırmada etkili olmayabilir. Model yorumlanırken bu durum göz önünde bulundurulmalıdır.

Uygulamada sıklıkla yapılan, bağımlı değişken gruplarının ayırımını sağlamada önemli olabileceği düşünülen tüm değişkenlerden oluşan büyük bir set oluşturmak ve analiz sırasında bu set içinde bağımlı değişkeni açıklamada anlamlı olanları saptamaktır. Diskriminant Analizi'nde önemli ve zor olan bir süreçtir. Bunun için, araştırmacının kullanacağı istatistik teknikle birlikte konu hakkındaki bilginin birleştirmesinin en etkili yöntem olacağı düşünülmektedir.

Diskriminant Analizi'ne ilk olarak modele girecek ayırıcı değişkenlere karar verilerek başlanmalıdır. Daha sonra varsayımların kontrolü ile devam edilmektedir. Ayırıcı değişkenlerin seçiminden sonra diskriminant fonksiyonunun elde edilerek testleri yapılır ve yorumlanır. Son olarak da Diskriminant Analizinin başarısı değerlendirilir.

2.1.1 Ayırıcı Değişkenlerin Seçimi

Diskriminant fonksiyonunun oluşturulmasına geçmeden önce modele girecek uygun bağımsız değişkenlerin seçilmesi gerekmektedir. Bu amaçla ilk önce, birimlerin gruplandırılmasında anlamlı olabileceği düşünülen bağımsız değişkenlerin grup farklarının anlamlı olup olmadığının saptanmasına olanak veren “Grup Ortalamalarının Eşitlik Testi” yapılarak grupları ayırmada etkin

olan deęişkenlerin modele bağımsız deęişken olarak alınmalarına karar verilebilir.¹³³

Genellikle kullanılan bir başka yol da deęişkenlerin grupları ayırmadaki gücüne bakmaksızın araştırma yapılan konu hakkında literatürde önemli olabileceęi düşünölen tüm bağımsız deęişkenler modele alınarak Adım Adım Diskriminant Analizi (Stepwise Discriminant Analysis) ile modele girecek bağımsız deęişkenlere karar verilmesidir¹³⁴. Bu yöntemle bağımsız deęişkenler arasındaki çoklu doğrusal bağlantı problemi de çözülmüş olmaktadır. Yaygın kullanım alanı bulmasına rağmen, deęişkenlerin seçim ve sıralanmasında kullanılan Adım Adım Diskriminant Analizi'nin bazı teorik sakıncalarının da farkında olmak gerekmektedir. Bu yöntemdeki temel problem, seçilen alt kümenin içindeki deęişkenlerin her zaman bağımlı deęişkeni açıklayan en iyi deęişkenler olmamasıdır. Ayrıca bağımsız deęişkenler arasındaki korelasyonun yüksek olduęu durumlarda boyut indirgemek amacıyla Adım Adım yöntemi yerine Faktör Analizi'nin kullanılması daha uygundur.¹³⁵

Diskriminant Analizi sonuçunda elde edilen bağımsız deęişkenlerin grupları ayırmadaki gücünün etkisine genellikle, 1932'de Wilks¹³⁶ tarafından önerilen, modelin Wilks' Lamda istatistięi incelenerek karar verilmektedir. McKay ve Campbell¹³⁷ da 1982'de yaptıkları çalışma ile Wilks' Lamda deęerinin toplam grupları ayırmada endeks olarak kullanılmasının kabul edilebilir olduęunu ifade etmişlerdir. Bununla birlikte Wilks' Lamda kadar önemli olan, doğrusal bileşenlerin (kanonik deęişkenlerin) tümü ile ilgili özdeęerlerin de dikkate

¹³³ Jacques Tacq, **Multivariate Analysis Techniques in Social Science Research**, Sage Publication, 1997, p. 248.

¹³⁴ Joseph F. Hair, v.d., **a.g.e.**, p.293.

¹³⁵ A. Pedro Durate Silva, Antonie Stam, Edited by Laurence G. Grimm, Paul R. Yarnold, **a.g.e.**, p. 289.

¹³⁶ S. S. Wilks, **Certain Generalizations in the Analysis of Variance**, Biometrika, Vol. 24, No. 3/4, 1932, pp. 471–494.

¹³⁷ R. J. McKay, N. A. Campbell, **Variable Selection Techniques in Discriminant Analysis I. Description**, British Journal of Mathematics Statistical Psychology, 35, 1982a, pp. 1–29.

alınması gerekmektedir. Analiz sonucunda elde edilen değişkenler kümesinde yer alan bazı önemli doğrusal bileşenlerin çok küçük özdeğerini dikkate alarak hata yapmaktan kaçınmak gerekmektedir. Grupları ayıran değişkenlerden oluşan doğrusal bileşenlerin önemi kendilerine karşılık gelen özdeğerleri tarafından yansıtılmaktadır. Özdeğeri maksimum yapacak şekilde elde edilen diskriminant fonksiyonunun, grup ayırımında önemli olan doğrusal bileşenleri içinde bulundurduğu anlamına gelmektedir.¹³⁸

Modelde yer alacak değişkenlerin belirlenmesinde kullanılan Adım Adım Diskriminant Analizi'nde üç farklı yaklaşım vardır. Bunlar; İleri Doğru Seçim (Forward Selection), Geriye Doğru Seçim (Backward Selection) ile ileri ve geriye doğru eleme yöntemlerinin birleşiminden oluşan Adım Adım Seçim (Stepwise Selection) yaklaşımıdır. Bu yaklaşımlarda modele girecek veya çıkacak değişkenlere karar verirken genellikle kullanılan seçim kriterleri; Wilks' Λ , Rao's V , Mahalanobis uzaklığı ve gruplararası F oranıdır.¹³⁹

Değişkenlerin diskriminant fonksiyonunda yer almaması önemsiz oldukları anlamına gelmemektedir. Değişkenler tek tek analiz edilip Wilks' Λ değeri bakıldığında en düşük değerle modelde olması beklenen değişkenin bazı durumlarda Adım Adım Diskriminant Analizi yöntemi ile modele girmediği görülmektedir. Bunun nedeni değişkenlerin kendi içlerindeki ilişkileridir. Bunu saptamak için korelasyon matrisi incelenebilir. Bu matrisdeki bazı bağımsız değişkenler arasındaki korelasyon değerlerinin büyük olması bu değerlere ait değişkenlerden ikisinin birden modelde kullanılmaması gerektiğini ifade etmekle birlikte, bazı durumlarda (bağımsız değişkenlerin bağımlı değişkenle olan korelasyon değerlerinin birbirine yakın olması gibi) hangisinin modelde kullanılmasının daha önemli olduğunu söylememektedir.¹⁴⁰

¹³⁸ A. Pedro Durate Silva, Antonie Stam, Edited by Laurence G. Grimm, Paul R. Yarnold, **a.g.e.**, p. 289.

¹³⁹ Subhash Sharma, **a.g.e.**, pp. 265–266.

¹⁴⁰ A.e., p. 272.

Veri setindeki çoklu doğrusal bağlantı için Adım Adım Diskriminant Analizi'nin kullanılıp kullanılmamasına çoklu doğrusal bağlantının kaynağına bağlı olarak karar verilmelidir. Anakütleye dayalı çoklu doğrusal bağlantı, bağımsız değişkenler arasındaki ilişkinin düzenini oluşturulabilecek tüm mümkün örnek seti içinde barındırır. Bu durum için Adım Adım Diskriminant Analizi'nin kullanılması örnekten örneğe değişmeyeceği için uygundur. Diğer taraftan bağımsız değişkenler arasındaki ilişki örnekten örneğe farklılık gösterdiğinde bu yönteminin kullanılması uygun olmayacaktır.¹⁴¹

2.1.2 Diskriminant Analizinin Varsayımları

Diskriminant fonksiyonunun oluşturulmasındaki en önemli varsayımlar çoklu normallik varsayımı ve grupların eşit varyansa sahip olması varsayımlarıdır.

Veri setinin çoklu normallik varsayımını sağlamaması diskriminant fonksiyonunun tahmin edilmesinde soruna neden olabilir. Bu durumda alternatif bir yöntem olan Lojistik Regresyon Analizi'nin kullanılması önerilmektedir. Eşit varyanslılık varsayımının sağlanamaması, başka bir ifade ile eşit olmayan varyans kovaryans matrisi ise sınıflandırma aşamasını olumsuz etkilemektedir. Eğer örnek büyüklüğü küçük ve varyans kovaryans matrisi eşit değilse tahminlerin istatistik anlamlılığını da olumsuz şekilde etkilemektedir. Örnek büyüklüğü yeterli olan gruplar arasında eşit varyanslı olma varsayımının sağlanamaması ise, gözlemlerin varyans kovaryans matris değeri daha büyük olan grupta, olması gerekenden daha fazla yer almasına neden olmaktadır ve bu sıklıkla karşılaşılan bir durumdur.

Bu iki temel varsayımın dışında veri setinin, sonuçları etkileyen bir diğer belirleyici özelliği de bağımsız değişkenler arasında çoklu doğrusal bağlantı problemi olmasıdır.¹⁴² Bu problemi ortadan kaldırmak için genellikle modele

¹⁴¹ Subhash Sharma, a.g.e., pp. 272-273.

¹⁴² Joseph F. Hair, v.d., a.g.e., pp. 290-291.

alınacak değişkenlerin seçiminde “Adım Adım (Stepwise)” veya “En İyi Alt Model Seçimi (Best Subset Selection)” yöntemleri kullanılmaktadır.

Diskriminant Analizi'nin bir diğer varsayımı da bağımlı ve bağımsız değişkenler arasındaki ilişkinin doğrusal olması gerekliliğini vurgulayan, doğrusallık varsayımdır.

Çoklu Normallik Varsayımı

Bağımsız değişkenler vektörünün çoklu normalliği sağlıyor olması aşağıdaki özelliklerin de sağlandığını göstermektedir:¹⁴³

- a- Bağımsız değişkenlerden oluşan doğrusal kombinasyonlar da normal dağılıma uymaktadır.
- b- Bağımsız değişkenlerden oluşan tüm mümkün alt kümeler de çoklu normalliği sağlamaktadır.
- c- Kovaryansın sıfır olması, ele alınan değişkenlerin birbirini etkilemeden (bağımsız olarak) dağıldığını göstermektedir.
- d- Bağımsız değişkenlerin koşullu dağılımı da (çoklu) normaldir.

Çoklu normallik varsayımının test edilmesinin zorlukları nedeniyle, varsayımın testinde birkaç pratik seçenek geliştirilmiştir. Bunlardan biri, her bir değişkenin tek tek Q-Q grafiği kullanılarak normallik testinin yapılmasıdır. Eğer bu çizilen grafiklerden herhangi biri normalliği sağlamıyorsa çoklu normallikle ilgili şüpheler doğmaktadır. Öte yandan tüm değişkenler tek tek normallik varsayımını sağlasa bile bu, çoklu normalliğin göstergesi değildir. Bu nedenle, çoklu normalliğin test edilmesi gerekmektedir.¹⁴⁴

¹⁴³ Richard A. Johnson, Dean W. Wichern, **Applied Multivariate Statistical Analysis**, Third Edition, Prentice Hall, New Jersey, 1992, p. 133.

¹⁴⁴ A. B. Ellem, **Multivariate Analysis** (Stat501), January 10, 2008 (Çevirimiçi) <http://mcs.une.edu.au/~stat501/UnitNotes/unitnotes08.pdf>, 25 Ağustos 2011, p. 28.

Çeşitli çoklu normallik testlerinde ve grafik yöntemde Mahalanobis uzaklığının karesi kullanılmaktadır¹⁴⁵. Mahalanobis uzaklığı, iki nokta arasındaki İstatistik uzaklığa (Standart uzaklık $\left(\frac{x_i - x_j}{s}\right)^2$) değişkenler arasındaki kovaryans ve korelasyon değerlerinde kullanılmasıyla elde edilen uzaklıktır. İki gözlem arasındaki Mahalanobis uzaklığının karesi şu şekilde hesaplanmaktadır;¹⁴⁶

$$MD_{ik}^2 = \frac{1}{1 - r^2} \left[\frac{(x_{i1} - x_{k1})^2}{s_1^2} + \frac{(x_{i2} - x_{k2})^2}{s_2^2} + \frac{2r(x_{i1} - x_{k1})(x_{i2} - x_{k2})}{s_1 s_2} \right]$$

Mahalanobis uzaklığının en önemli özelliği, değişkenler ilişkili değilse Mahalanobis uzaklığının İstatistik uzaklığa dönüşmesidir. Ayrıca değişkenlerin varyansları “1” ise Mahalanobis uzaklığı Öklid uzaklığına indirgenir. Kısaca, Öklid uzaklığı ve İstatistik uzaklık Mahalanobis uzaklığının özel durumudur.¹⁴⁷

p değişkenli durum için, iki gözlem arasındaki Mahalanobis uzaklığı en basit olarak aşağıdaki gibidir;

$$MD_{ik} = (x_i - x_k)' S^{-1} (x_i - x_k)$$

Burada x , $p \times 1$ koordinat vektörü ve S , $p \times p$ kovaryans matrisidir.

Johnson ve Wichern tarafından çoklu normalliğin kontrol edilmesinde, tek değişkenli Q-Q grafiğinin benzeri olan Mahalanobis uzaklığının kullanıldığı χ^2 grafiğinin kullanılması uygundur. Bu durumda uzaklıkların karesi küçükten büyüğe sıralayarak, bu değerlere karşılık gelen beklenen değerleri χ^2 dağılımı kullanarak hesaplanır ve sıralanmış Mahalanobis uzaklık değerleri ile χ^2 dağılımı kullanılarak elde edilen beklenen değerlerinin grafiği çizilir. Bu grafik sonucunda doğrusal çizgiden sapma çoklu normallik varsayımının sağlanmadığı

¹⁴⁵ M. Bilodeau, D. Brenner, **Theory of Multivariate Statistics**, Springer, 1999, p. 170.

¹⁴⁶ Subhash Sharma, **a.g.e.**, p. 44.

¹⁴⁷ Carl J. Huberty, **Applied Discriminant Analysis**, John Wiley & Sons INC., 1994, p. 43.

anlamına gelmektedir. Veri setinde n ve $n - p$ değeri 25 veya 30'dan büyükse Mahalanobis uzaklıklarının karesi χ^2 dağılıma benzer özellikler göstermektedir.

148

Bununla birlikte Bilodeau ve Brenner¹⁴⁹, tek değişkenli durumlar için örnek büyüklüğünün 25'den fazla olması χ^2 yaklaşımının kullanılması için yeterli olsa da değişken sayısı arttığında (örneğin 4 olduğunda) örnek büyüklüğünün 100 olmasının χ^2 yaklaşımının kullanılması için yeterli olmayacağını ifade etmişlerdir. Bunun yerine Beta dağılımının kullanılması önerilmektedir. Bunun dışında, Johnson ve Wichern¹⁵⁰ da küçük örnekler için χ^2 dağılımının daha genel hali olan Gamma dağılımının kullanılmasını önermekle birlikte, uygulamada χ^2 dağılımının kullanılması daha yaygındır.

Lachenbruch, Sneeringer ve Revo 1973'de yaptıkları çalışma ile çoklu normallik varsayımının sağlanmaması ile sınıflandırma oranlarının nasıl etkilendiğini incelemişlerdir. Bu çalışma ile çoklu normallik varsayımının sağlanmaması sonucunda anlamlılık testleri ve doğru sınıflandırma oranlarının güvenilirliğini kaybedeceği saptanmıştır.¹⁵¹ Çoklu normallik varsayımının sağlanmadığı durumda öncelikle uygun dönüşümler yapılarak veri setinin tekrar oluşturulmasına çalışmak gerekmektedir. Dönüşümler denendikten sonra da sorun devam ediyorsa açıklayıcı değişkenler için herhangi bir dağılım varsayımına dayanmayan Lojistik Regresyon Analizi'nin kullanılması daha uygun olacaktır.¹⁵²

¹⁴⁸ Richard A. Johnson, Dean W. Wichern, **a.g.e.**, pp. 158–160.

¹⁴⁹ M. Bilodeau, D. Brenner, **a.g.e.**, p. 186.

¹⁵⁰ Richard A. Johnson, Dean W. Wichern, **a.g.e.**, p. 160.

¹⁵¹ Peter A. Lachenbruch, Cheryl Sneeringer, Lawrence T. Revo, "Robustness of the Linear and Quadratic Discriminant Function to Certain Types of Non-Normality", **Communications in Statistics**, 1973, Vol. 1, pp. 39–56.

¹⁵² Subhash Sharma, **a.g.e.**, p. 8.

Eşit Varyanslılık Varsayımı

Çoklu normallik varsayımı dışında Diskriminant Analizi sonuçlarının güvenilirliği açısından eşit varyanslılık varsayımının da sağlanması gerekmektedir. Tüm grupların varyanslarının eşit olduğunun ileri sürüldüğü sıfır hipotezinin testi için, F dağılımına uyan Box's M değerinden hesaplanan F test istatistiği kullanılmaktadır.¹⁵³

$$M = (n - g) \log |S| - \sum_{i=1}^g (n_i - 1) \log |S_i|$$

$$F = \begin{cases} M/b & \text{eğer } \tau > (1 - \rho)^2 \\ \frac{f_2 M}{f_1(b - M)} & \text{eğer } \tau \leq (1 - \rho)^2 \end{cases}$$

Burada g ; grup sayısı, n_i ; i . gruptaki birim sayısı, S ; ortak kareler toplam matrisi ve S_i ; i . grubun kareler toplam matrisidir. F dağılımına uyan Box's M test istatistiği için f_1 ve f_2 serbestlik dereceleri, modeldeki değişken sayısı p ile gösterilmek üzere sırası ile şu şekilde hesaplanır;¹⁵⁴

$$f_1 = (g - 1)p(p + 1)/2$$

$$f_2 = \frac{f_1 + 2}{|\tau - (1 - \rho)^2|}$$

2.1

İkinci serbestlik derecesindeki τ ve ρ katsayıları da şu şekilde elde edilmektedir;

$$\tau = \frac{(p - 1)(p + 2)}{6(g - 1)} \left(\sum_{i=1}^g \frac{1}{(n_i - 1)^2} - \frac{1}{(n - g)^2} \right)$$

¹⁵³ Kazım Özdamar, Cilt 2, **a.g.e.**, s. 407.

¹⁵⁴ SPSS, (Çevirimiçi)

http://support.spss.com/ProductsExt/SPSS/Documentation/Statistics/algorithms/14.0/app14_boxs_m.pdf, 10 Eylül 2011.

$$\rho = 1 - \frac{2p^2 + 3p - 1}{6(p+1)(g-1)} \left(\sum_{i=1}^g \frac{1}{(n_i - 1)} - \frac{1}{(n - g)} \right)$$

F test istatistiğinin formülünde yer alan b değeri de şu şekilde hesaplanmaktadır,¹⁵⁵

$$b = \begin{cases} \frac{f_1}{\rho - (f_1/f_2)} & \text{eğer } \tau > (1 - \rho)^2 \\ \frac{f_2}{\rho - (2/f_2)} & \text{eğer } \tau \leq (1 - \rho)^2 \end{cases}$$

Eğer $(1 - \rho)^2 - \tau = 0$ ise Box's M değerinden elde edilen F testi yerine *Bartlett* χ^2 test istatistiğinin kullanılması uygundur. χ^2 test istatistiği şu şekilde hesaplanabilir,¹⁵⁶

$$\chi^2 = M\rho$$

χ^2 istatistiği, f_1 (Eşitlik 2.1'de yer alan) serbestlik derecesi ile kabul edilen anlamlılık düzeyindeki χ^2 tablo değeri ile karşılaştırılarak yorumlanır.

Eşit varyanslılık varsayımının sağlanmadığı durumda, doğrusal diskriminant fonksiyonunun yerine karesel diskriminant fonksiyonunun oluşturulması daha uygundur¹⁵⁷.

Karesel diskriminant fonksiyonun nasıl oluşturulacağına “Diskriminant Fonksiyonunun Elde Edilmesi” başlığında (Başlık 2.1.3) değinilmiştir.

Çoklu Doğrusal Bağlantı Problemi Olmaması Varsayımı

Birçok istatistik yöntemin uygulanabilir olması için sağlanması gereken varsayımlardan biri de bağımsız değişkenler arasında ilişki olmamasıdır.

¹⁵⁵ Kazım Özdamar, Cilt 2, **a.g.e.**, s. 407.

¹⁵⁶ A.e., p. 408.

¹⁵⁷ Subhash Sharma, **a.g.e.**, p. 264.

Bağımsız değişkenler arasındaki ilişki en basit şekli ile iki değişken arasındaki korelasyon katsayısının gücü ile belirlenebilmektedir. Çoklu doğrusal bağlantının olması, bir bağımsız değişken diğer bir bağımsız değişken veya değişkenler tarafından tam olarak tahmin edilebiliyorsa bu değişkenlerden biri veya bazılarının gereksiz olacağı anlamına gelmektedir. Bu durumda söz konusu değişken veya değişkenlerin modelde yer almaması daha uygundur.

Çoklu doğrusal bağlantı problemini saptamak için çeşitli yöntemler kullanılmaktadır.¹⁵⁸

- Bu yöntemlerden biri, daha öncede değinildiği gibi, basit korelasyon matrisinin incelenmesidir. İki bağımsız değişken arasındaki basit korelasyon katsayısı anlamlı olması çoklu doğrusal bağlantı probleminin ortaya çıkabileceğini göstermektedir.
- Çoklu doğrusal bağlantının saptanmasında kullanılan bir diğer yöntem de modele bağımsız değişkenler ilave edildikçe belirlilik katsayısındaki değişimlerin incelenmesidir.
- Bir başka yöntem de kısmi korelasyon katsayılarının incelenmesidir. İki değişken arasındaki basit korelasyon katsayıları anlamlı olduğu halde kısmi korelasyon katsayılarının anlamsız çıkması çoklu doğrusal bağlantı problemi için bir gösterge olabilmektedir.¹⁵⁹
- Çoklu doğrusal bağlantı problemi Varyans Artış Faktörleri (Variance Inflation Factors-VIF) incelenerek de saptanabilmektedir.¹⁶⁰

$$VIF = \frac{1}{(1 - R^2)}$$

- Tolerans değeri $T = 1 - R_i^2$ şeklinde hesaplanmaktadır, bağımsız değişkenler arasındaki çoklu bağımlılığın miktarının ölçüsüdür.¹⁶¹

¹⁵⁸ Damodar N. Gujarati, **a.g.e.**, ss. 335–339.

¹⁵⁹ Neyran Orhunbilge, **Uygulamalı Regresyon ve Korelasyon Analizi**, İstanbul, 2002, s. 241.

¹⁶⁰ Damodar N. Gujarati, **a.g.e.**, s. 339.

¹⁶¹ Joseph F. Hair, v.d., **a.g.e.**, p.273.

Eşitlikten de anlaşılacağı gibi küçük Tolerans değeri daha büyük VIF değeri demektir.

- Çoklu doğrusal bağlantı probleminin saptanmasında kullanılan, bir diğer yöntem de Koşul Endeks (Condition Index-CI) değerlerinin hesaplanmasıdır. CI değerleri aşağıdaki gibi hesaplanmaktadır:¹⁶²

$$CI = \sqrt{\frac{V_{max}}{V_{x_i}}}$$

Burada maksimum açıklanan varyans V_{max} , ile i . değişken tarafından açıklanan toplam varyans da V_{x_i} ile gösterilmektedir. CI değerinin 10 ile 30 arasında değer alması çoklu doğrusal bağlantı probleminin orta düzeyde, 30'dan büyük olması da çok güçlü olduğu anlamına gelmektedir.

Çoklu doğrusal bağlantı probleminin saptanmasından sonra, soruna neden olan değişkenlerden hangisinin modelde kalıp hangisinin dışarıda bırakılacağına karar vermek kolay değildir. Bu aşamada çoklu doğrusal bağlantı problemini çözmek* için Adım Adım Regresyon Analizi'ni kullanmak uygun olacaktır. Bunun dışında bazı durumlarda çoklu doğrusal bağlantı problemine neden olan değişkenlere dönüşüm uygulanarak modele alınması daha uygun olabileceği gibi her zaman mümkün olmamakla birlikte örnek büyütülerek çoklu doğrusal bağlantı problemi çözülmeye çalışılabilir. Bunların dışında çoklu doğrusal bağlantı problemini çözmek için veri setine çok değişkenli istatistik yöntemlerden olan Faktör Analizi de uygulanabilir.¹⁶³

Doğrusallık Varsayımı

Diskriminant Analizi'nde de korelasyon katsayılarına dayanan çok değişkenli tekniklerdeki gizli varsayım olan doğrusallık varsayımının sağlanması

¹⁶² Damodar N. Gujarati, **a.g.e.**, s. 338.

* Çoklu doğrusal bağlantı problemi ortaya çıktığında çözüm yöntemleri ile ilgili ayrıntılı bilgi için Neyran Orhunbilge'nin "Uygulamalı Regresyon ve Korelasyon Analizi" kitabı sayfa 242'ye bakınız.

¹⁶³ Damodar N. Gujarati, **a.g.e.**, ss. 340–344.

önemlidir. Bağımlı değişken ile bağımsız değişkenler arasındaki ilişkinin doğrusal olmaması, doğrusal korelasyon katsayılarının gerçek ilişkiyi daha düşük göstermesine neden olmaktadır.¹⁶⁴

2.1.3 Diskriminant Fonksiyonunun Elde Edilmesi

İki Kategorili Diskriminant Analizi'ndeki amaç, grupların ayırımını en iyi şekilde yapacağı düşünülen bağımsız değişkenler setinin iki grubu ayırmada yeterli olup olmadığının analizinin yapılmasıdır. Bu nedenle iki grubun en üst düzeyde ayırımını yapabilecek bağımsız değişkenlerin doğrusal kombinasyonları araştırılmaktadır. Bu doğrusal kombinasyona “diskriminant fonksiyonu” adı verilmekte ve

$$T - \bar{T} = k_1(X_1 - \bar{X}_1) + k_2(X_2 - \bar{X}_2) + \dots + k_p(X_p - \bar{X}_p)$$

$$t = k_1x_1 + k_2x_2 + \dots + k_px_p \quad 2.2$$

şekilinde gösterilmektedir.¹⁶⁵ Eşitlikte küçük harfle gösterilen t ve x_i , ortalamadan sapmaları ifade etmektedir. k_i katsayıları, diskriminant ağırlıkları (kanonik değerler) olarak adlandırılmaktadır. İki grup olduğunda sadece bir diskriminant fonksiyonu vardır.

Diskriminant fonksiyonundaki ağırlıklar, gruplararası varyansın gruplarıçi varyansa oranını veren aşağıdaki eşitlikle hesaplanan λ (eigenvalue-özdeğer) değeri¹⁶⁶

$$\lambda = \frac{\text{Gruplararası Varyans}}{\text{Gruplarıçi Varyans}}$$

¹⁶⁴ Ali Sait Albayrak, **Uygulamalı Çok Değişkenli İstatistik Teknikler**, Asil Yayın Dağıtım Ltd. Şti., 2006, s. 65.

¹⁶⁵ Jacques Tacq, **a.g.e.**, p. 233.

¹⁶⁶ Subhash Sharma, **a.g.e.**, p. 241.

şeklinde yazılmaklabirlikte, yukarıdaki eşitlik Fisher tarafından önerilen ve k katsayılarının elde edilmesinde kullanılmak üzere şu şekilde ifade edilmektedir.¹⁶⁷

$$\lambda = \frac{k'Bk}{k'Wk}$$

Buradan, B ; gruplararası varyans, W ; gruplarıçi varyans matrisi olmak üzere λ değerini maksimum yapacak şekilde k katsayıları elde edilmektedir.*

Bu şekilde elde edilen eşitliğe *Fisher'in Doğrusal Diskriminant Fonksiyonu* adı verilmektedir¹⁶⁸.

Katsayıların hesaplanmasında kullanılan eşitlikler doğrusal olarak bağımsız olmadığında, elde edilen katsayıların oranları birbirine eşit olmak üzere sonsuz sayıda çözüm elde edilecektir. Çözüm kümesi sonsuz olduğunda, standartlaştırma amacıyla diskriminant fonksiyonunun katsayıları normalleştirilerek kullanılmalıdır. Normalleştirme, katsayıların k uzunluğuna* bölünerek elde edilmesidir.¹⁶⁹

$$k_i^* = \frac{k_i}{\sqrt{k'k}} \quad i = 1, 2, \dots, p$$

k ; $p \times 1$ boyutunda p değişkene ait diskriminant ağırlıklar vektörüdür.

Her grup için birer diskriminant fonksiyonu yazılabileceği gibi grupların normal dağılıma uyduğu varsayımına dayanan Kanonik Diskriminant Analizi'nde

¹⁶⁷ Hüseyin Tatlıdil, **a.g.e.**, s. 258.

* λ 'in k 'ya göre kısmi türevleri alınarak sifıra eşitlendiğinde elde edilen $(W^{-1}B - \lambda I)k = 0$ denklem sisteminde (W gruplarıçi ve B gruplararası varyasyonu ifade etmektedir) önce λ değerleri (özdeğerler) hesaplanarak denklem sisteminde yerine konulup diskriminant ağırlıkları hesaplanmaktadır.

¹⁶⁸ Subhash Sharma, **a.g.e.**, p. 245.

* k uzunluğu $\|k\| = \sqrt{k'k}$ şeklinde hesaplanmaktadır.

¹⁶⁹ Jacques Tacq, **a.g.e.**, p. 245.

uygulandığı şekilde, diskriminant ağırlıkları, ortalama fark vektörleri kullanılarak sabit değerin de yer aldığı ortak bir diskriminant fonksiyonu olarak da ifade edilebilmektedir. Matris gösterimi ile diskriminant ağırlıkları¹⁷⁰

$$k = S^{-1}(\bar{X}_1 - \bar{X}_2) \quad 2.3$$

ile bulunurken, fonksiyondaki sabit değer,¹⁷¹

$$k_0 = \left(-\frac{1}{2}\right)\bar{X}'_1 S^{-1}\bar{X}_1 + \left(-\frac{1}{2}\right)\bar{X}'_2 S^{-1}\bar{X}_2 \quad 2.4$$

eşitliği ile hesaplanmaktadır.

Eşitlik 2.3 ve 2.4'deki, S ; $p \times p$ boyutlu ortak varyans kovaryans matrisi olup, iki grubun örnek büyüklüğü ve varyans kovaryans matrisleri kullanılarak şu şekilde elde edilmektedir;

$$S = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{n_1 + n_2 - 2} \quad 2.5$$

$(\bar{X}_1 - \bar{X}_2)$; $p \times p$ boyutunda p tane değişkenin grup ortalamalarının fark vektörüdür.

Dikkat edilmesi gereken önemli bir nokta, diskriminant ağırlıklarının işaretlerinin diskriminant ekseninin yönünü gösteriyor olduğudur. Bağımsız değişkenlerin birbirleriyle ilişkilerinin yönüne göre işaret değişmektedir. Örneğin değişkenler arasındaki ilişki pozitif ise diskriminant katsayıları diskriminant ekseninin yönünü beklenenin aksi şekilde etkilemektedir yani değişkenlerin işaretleri terstir.¹⁷² Sabit terim bağımsız değişkenlerin ortalama farkları kullanılmadığında fonksiyonda yer alır. Diskriminant ağırlıklarının hesaplanmasında her bir değişkenin ortalama farkları

¹⁷⁰ Richard A. Johnson, Dean W. Wichern, **a.g.e.**, p. 505.

¹⁷¹ Kazım Özdamar, Cilt 2, **a.g.e.**, s. 361.

¹⁷² Jacques Tacq, **a.g.e.**, p. 238.

kullanıldığında ise sabit terim yoktur. Ayrıca, değişkenin standartlaştırılmış değerleri (z_i) yer aldığında diskriminant fonksiyonu yine sabit terim içermez. Standartlaştırılmış değerlerin kullanılması grupların belirlenmesinde hangi bağımsız değişkenin daha güçlü olduğunun karşılaştırılmasına imkan verdiği için tercih edilmektedir.

Diskriminant fonksiyonu veya fonksiyonlarının ayırma güçlerinin istatistik anlamlılığı *Wilks'Λ*, *Hotelling's Trace* ve *Pillai's Kriteriya* değerleri ile değerlendirilebilmektedir. *Roy'un En Büyük Karakteristik Kökü*, sadece ilk diskriminant fonksiyonunun ayırma gücünü değerlendirmede kullanılmaktadır.¹⁷³ Öneren istatistikçilerin adı ile anılan yukarıda ki dört yöntem test istatistikleri değerlerini elde ederken λ değerlerini kullanmaktadırlar.*

Burada unutulmaması gereken nokta, diskriminant fonksiyonunun oluşturulmasında, Bölüm 2.1.2'de incelenen varsayımlardan olan çoklu normallik ve eşit varyanslılık varsayımlarının sağlanıyor olması koşuludur.

Verilerin normal dağılıma uyduğu ve grupların eşit varyanslılık varsayımının sağlandığı durumda diskriminant fonksiyonunu oluştururken, fonksiyonun katsayılarının (doğrusal bileşenlerinin) hesaplanmasında Eşitlik 2.5'de verilen ortak varyans kovaryans matrisi kullanılırken, verilerin normal dağılıma uyduğu ancak gruplara ait eşit varyanslılık varsayımının sağlanmadığı durumda ise, grupların varyans kovaryans matrislerinin farkları ($S_1 - S_2$) alınarak karesel diskriminant fonksiyonunun doğrusal bileşenlerinin elde edilmesi daha doğru olacaktır.¹⁷⁴

¹⁷³ Joseph F. Hair, v.d., **a.g.e.**, p. 294.

* *Wilks'Λ*, *Hotelling's Trace*, *Pillai's Kriteriya* ve *Roy'un En Büyük Karakteristik Kökü* ile ilgili ayrıntılı bilgi için bakınız: Neyran Orhunbilge, **Çok Değişkenli İstatistik Yöntemler**, İstanbul, 2010, ss. 342–344.

¹⁷⁴ C. A. B. Smith, "Some Examples of Discrimination", **Annals of Eugenics**, 1947, 13, pp. 272–282.

Bununla birlikte Marks ve Dunn 1974'te sundukları çalışmada varyans kovaryans matrisinin eşit olmadığı durumda diskriminant fonksiyonunu farklı λ değerleri ve örnek büyüklükleri için karesel ve doğrusal fonksiyon denklemlerini test ederek hangisinin daha iyi sonuç verdiğini saptamışlardır. Bu çalışmanın sonucunda büyük örnek ve λ oranının büyük değeri için karesel fonksiyonun doğrusal fonksiyondan daha iyi sonuç verdiği elde edilmiştir.¹⁷⁵ 1969'da Gilbert¹⁷⁶ tarafından yapılan bir diğer çalışmada da çoklu normallik varsayımı sağlanırken varyans kovaryans matrisinin eşit olmadığı durum için doğrusal ve karesel sınıflandırma kuralına göre farklı bağımsız değişken sayısı ve örnek büyüklüğü için grupların sınıflandırılmasında farklı sonuçlar elde edildiği gösterilmiştir.

2.1.4 Diskriminant Fonksiyonunun Test Edilmesi ve Yorumlanması

İki grup için bağımsız değişkenlerin anlamlı şekilde farklı olup olmadığının testleri tek tek yapılabileceği gibi değişkenlerin birlikte anlamlı olup olmadığının da test edilmesi gerekmektedir. Bu bağlamda testler tek ve çok değişkenli testler olmak üzere ikiye ayrılmaktadır.*

- *Tek Değişkenli Testler*

Tek değişkenli testlerde modelde sanki tek değişken varmış gibi düşünülerek testler yapılmaktadır. Bu çalışmada *Student t* testi kullanılmakla birlikte bilgisayar paket program (SPSS) çıktısında Wilks' Lamda ve tek değişkenli *F* test sonuçları da yer almaktadır. *F* istatistik değeri, *Student t* istatistik değerinin

¹⁷⁵ S. Marks, O. J. Dunn, "Discriminant Functions When Covariance Matrices are Unequal", **Journal of the American Statistical Association**, 1974, 69, pp. 555–559.

¹⁷⁶ Ethel S. Gilbert, "The Effect of Unequal Variance-Covariance Matrices on Fisher's Linear Discriminant Function", **Biometrics**, 1969, 25, pp. 505–515.

* Anlamlılık testlerinden tek değişkenli için *Student t* testi ve çok değişkenli için de *Hotelling T²* testi burada ele alınmıştır. Bunun dışındaki anlamlılık testleri (*Wilks'λ* ve *Rao V*) için Neyran Orhunbilge'nin **Çok Değişkenli İstatistik Yöntemler** kitabı sayfa 224-225'e bakınız.

karesine eşittir ve değişkenlerin tek tek gruplararası ve gruplarıçi varyanslarının birbirlerine oranlanması ile de hesaplanabilmektedir. Wilks' Lamda istatistiği ise her bir değişkenin gruplarıçi sapmasının toplam sapmaya oranından bulunmaktadır. Bağımsız değişkenlerin katsayılarının sıfıra eşit olduğunu ileri süren sıfır hipotezinin test sonuçlarının anlamlı olması beklenmektedir.

- *Çok Değişkenli Testler*

Modeldeki bağımsız değişkenlerin hepsinin birlikte anlamlı olup olmadığının testinin yapılması çok değişkenli testler olarak adlandırılmaktadır. *Student t* testinin karesine benzer olan ve çok değişkenli testlerde kullanılan istatistik *Hotelling T²* istatistiğidir. Bu test sadece bir değişkeni değil model için tanımlanan tüm değişkenleri içermektedir ve birden fazla değişken test edildiği için grupların ortalamaları yerine grup merkezlerinin farklarını incelemektedir. Ayrıca tek değişken incelendiğinde formülde $1/s_w^2$ olarak yer alan değer yerine gruplararası kovaryans matrisini ifade eden (C_w) değerin tersi alınmaktadır.¹⁷⁷

$$t^2 = \frac{(\bar{X}_1 - \bar{X}_2)^2}{\frac{s_w^2}{n_1} + \frac{s_w^2}{n_2}} = \frac{n_1 n_2}{n_1 + n_2} (\bar{X}_1 - \bar{X}_2) \frac{1}{s_w^2} (\bar{X}_1 - \bar{X}_2)$$

Hotelling T² test istatistiği

$$\left[\frac{(n - p - 1)}{p(n - 2)} \right] T^2$$

formülü ile hesaplanmakla birlikte p ve $n - p - 1$ serbestlik derecesi ile F dağılımına uymaktadır.

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} d' C_w^{-1} d$$

¹⁷⁷ Jacques Tacq, **a.g.e.**, pp. 250–251.

Formülde T^2 eşitliğindeki $d' C_w^{-1} d$ kısmı Mahalanobis uzaklığını vermektedir.* *Hotelling* T^2 testi dışında bağımlı değişkenin iki kategoriden fazla olduğu durumlar için geliştirilmiş en genel çok değişkenli test olan Wilks' Lamda (bkz. sayfa 107) bağımlı değişkenin iki kategorili olduğu durumlar için de kullanılabilir.¹⁷⁸

2.1.5 Gözlemlerin Gruplarının Belirlenmesi

Veri setindeki birimlerin veya yeni yapılacak gözlemlerin hangi gruba ait olduğunu belirlemek için temelde iki farklı yol izlenmektedir.

1. Gözlemlerin her bir grup ortalama vektörüne olan uzaklıkları belirlenip, bu uzaklık içinden değer olarak en küçük olan grup hangisi ise gözlem o gruba atanır.¹⁷⁹
2. Gözlemin hangi sınıfta yer alacağını belirlemeye yarayan bir diğer yöntem de kriter değerinin kullanılmasıdır. En genel hali ile kriter değeri şu şekilde ifade edilmektedir:¹⁸⁰

$$(\bar{X}_1 - \bar{X}_2)' S^{-1} x_0 \geq \frac{1}{2} [(\bar{X}_1 - \bar{X}_2)' S^{-1} (\bar{X}_1 - \bar{X}_2)] + \ln \left[\frac{C(1|2)}{C(2|1)} \left(\frac{p_2}{p_1} \right) \right]$$

Yukarıdaki eşitsizliğin sağlanması durumunda x_0 gözlemi (incelenen gözleme ait değişkenlerin değerlerinin yer aldığı vektör) 1. gruba atanmaktadır. Eşitsizliğin sağlanmadığı durumda ise gözlem 2. gruba atanmaktadır. Buradaki p_1 ve p_2 sırası ile 1. ve 2. gruba ait ön olasılık değerleridir. $C(1|2)$ ve $C(2|1)$ de sırası ile gözlemin gerçekte 2. grupta

* *Hotelling* T^2 , *Mahalanobis* D^2 uzaklığını içermektedir: $D^2 = d' C_w^{-1} d$

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} d' C_w^{-1} d = \frac{n_1 n_2}{n_1 + n_2} D^2$$

Buradan F testi de D^2 'nin beklenen değeri tarafından şu şekilde ifade edilebilmektedir:

$$F = \frac{n - p - 1}{p(n - 2)} \frac{n_1 n_2}{n_1 + n_2} D^2$$

¹⁷⁸ Jacques Tacq, **a.g.e.**, p. 348.

¹⁷⁹ Kazım Özdamar, Cilt 2, **a.g.e.**, s. 381.

¹⁸⁰ Richard A. Johnson, Dean W. Wichern, **a.g.e.**, p. 499,505.

iken 1. gruba atanması ile gerçekte 1. grupta iken 2. gruba atanmasının maliyetini (cost) göstermektedir.

Gözlemlerin gruplara ait olma ön olasılıklarının ve yanlış atama maliyetlerinin eşit olduğu durumlar için ise yukarıdaki eşitsizliğin sağ tarafındaki ifade

$$\frac{1}{2}[(\bar{X}_1 - \bar{X}_2)'S^{-1}(\bar{X}_1 - \bar{X}_2)]$$

olarak değişmektedir.

2.1.6 Diskriminant Analizinin Ayırıcı Gücünün Değerlendirilmesi

Yeni gözlemlerin iki gruptan hangisine ait olduğunun belirlenmesi Diskriminant Analizi'nin hedeflerinden bir diğeridir. Sınıflandırma genelde Diskriminant Analizi'nden bağımsız bir süreç olarak düşünülmesine karşın, birçok kaynak ve istatistik paket programında bu süreç Diskriminant Analizi'nin bir parçası olarak ele alınmaktadır ve fonksiyondan elde edilen diskriminant değerleri kullanılarak uygulanmaktadır. Diskriminant değerlerinden oluşan bu uzay yanlış sınıflandırma hatasını veya maliyetini minimize edecek şekilde kritik değer veya ayırıcı değer olarak adlandırılan diskriminant değeriyle ikiye bölünmektedir. Bu şekilde, gözlemin hangi gruba ait olduğuna karar verilmektedir.¹⁸¹

Genel uygunluğun değerlendirilmesi için sınıflandırma matrisi (classification matrix) oluşturulmaktadır. Bu matris oluşturulmadan önce her birimin bağımsız değişkenlere ait değerleri diskriminant fonksiyonunda yerine yazılarak hangi gruba ait olduğuna karar verilmektedir. Bu işlemler eğer veri seti yeterince büyükse fonksiyonu oluşturmak ve sınıfları belirlemek için kullanılmak üzere iki parçaya ayrılır ve bir grup gözlem ile fonksiyon yazıldıktan sonra kullanılmayan diğer gözlemlerle de sınıflandırma matrisi elde edilmektedir. Bu aşamadan sonra sınıflandırma matrisinin ayırma gücünün istatistik testi için

¹⁸¹ Ali Sait Albayrak, **a.g.e.**, ss. 319–320.

Press's Q istatistiği kullanılmaktadır. Bu istatistik değeri, doğru sınıflandırma sayısını toplam örnek büyüklüğü ve grup sayısı ile karşılaştıran basit bir ölçümdür. Bulunan değer 1 serbestlik dereceli χ^2 dağılımından elde edilen kritik değerle karşılaştırılır. Eğer hesaplanan değer kritik değerden büyükse, sınıflandırma matrisinin istatistik olarak anlamlı ve yorumlanabilir olduğu anlamına gelmektedir. *Q* istatistiği şu şekilde hesaplanmaktadır:¹⁸²

$$Press's Q = \frac{[n - (n_d g)]^2}{n(g - 1)}$$

Burada n ; toplam örnek büyüklüğü, n_d ; doğru sınıflandırılan birim sayısı, g ; grup sayısını göstermektedir.

Sınıflandırma işlemi tamamlandıktan sonra yapılması gereken bir diğer işlem de yanlış sınıflandırılan birimlerin saptanması ve yanlış sınıflandırmaya neden olan durumun incelenmesidir.¹⁸³

2.2 Diskriminant Fonksiyonunun Dışsal Geçerliliği

Diskriminant Analizi sonucunda elde edilen diskriminant fonksiyonu, gözlemlerin sınıflandırılması için kullanılacaksa dışsal geçerliliğinin* (external validation) incelenmesi gerekmektedir. Diskriminant fonksiyonunun doğruluğunu referans alan dışsal geçerlilik, diğer örneklerdeki gözlemleri sınıflandırabilmelidir. Diskriminant fonksiyonunun tahmin edilmesinde kullanılan gözlemlerin aynı zamanda sınıflandırılmasından elde edilen hata oranı yanlış olmaktadır. Bu nedenle, fonksiyonun yazılmasında kullanılan gözlemler diskriminant fonksiyonunun geçerliliğini ölçmede kullanılmamalıdır.

¹⁸² Joseph F. Hair, v.d., **a.g.e.**, p. 303.

¹⁸³ **A.e.**, p. 304.

* Dışsal geçerlilik (external validation), bir örneklem üzerinde elde edilen araştırma sonuçlarının daha büyük gruplara ve evrene genellenme derecesidir.

Diskriminant fonksiyonunun geçerliliğini ölçmek için sıklıkla kullanılan üç teknik vardır. Bu teknikler şu şekilde sıralanabilir:

- *Gizleme yöntemi (Holdout Method)*

Diskriminant Analizi'nin son adımında sonuçların geçerliliği için çapraz geçerlilik yöntemi de denilen Gizleme Yöntemi uygulanabilir. Bu yöntem en basit şekli ile üç adımdan oluşmaktadır. İlk olarak veri seti rastgele iki parçaya ayrılır. Gözlemlerin bir parçası kullanılarak diskriminant fonksiyonu elde edildikten sonra, fonksiyonun oluşturulmasında kullanılmayan gözlemlere ait değerler bu fonksiyonda yerine konularak doğru sınıflandırma oranları elde edilmektedir. Bazı araştırmacılar bu işlemi birkaç kez tekrarlayıp sınıflandırma işlemi sonucunda bulunan çıktıların ortalamalarını alarak doğru sınıflandırma oranının bulunmasının daha uygun olduğunu düşünmektedirler¹⁸⁴. Ayrıca bu yöntemi kullanabilmek için veri setinin yeterince büyük olması gerekmektedir.

- *U-Yöntemi (U-Method)*

Bu yöntem Lachenbruch¹⁸⁵ tarafından 1967 yılında önerilmiştir. Her seferinde bir gözlem dışarıda tutarak $n - 1$ gözlemlerle diskriminant fonksiyonu tahmin edilip, fonksiyonu oluşturmak için kullanılmayan gözlemin değerleri elde edilen fonksiyonda yerine koyularak gözlemin sınıfı belirlenir ve bu işlem n defa tekrarlanarak elde tutulan n birim için sınıflama yapılır. Bu yöntem, sınıflama oranı açısından neredeyse yansız tahmin yapmakla birlikte varyans ve ortalama hata karelerinin küçük olmaması nedeniyle eleştirilmektedir.

¹⁸⁴ Joseph F. Hair, v.d., **a.g.e.**, p. 311.

¹⁸⁵ P. A. Lachenbruch, "An Almost Unbiased Method of Obtaining Confidence Intervals for the Probability of Misclassification in Discriminant Analysis", **Biometrics**, Vol: 23, No. 4, 1967, pp. 639–645.

- *Bootstrap Yöntemi*

İstatistikte çoğunlukla parametrelerin normal dağılım varsayımına dayanılarak tahmin edilmesi söz konusu iken Bootstrap Yöntemi'nde veri setinden rastgele iadeli örnekler alınarak parametrelerin, dağılım varsayımından bağımsız olarak tahmini yapılabilmektedir. Bootstrap, küçük veri setleri için de kullanılabilen bir yöntemdir.¹⁸⁶

Bootstrap Yöntemi literatürde birkaç farklı şekilde adlandırılmaktadır. İlk olarak dağılım varsayımı olup olmadığına göre "Parametrik Bootstrap" ya da "Parametrik Olmayan Bootstrap" olmak üzere ikiye ayrılmıştır. İkinci olarak ise örneğin alınış şekline ve/veya parametrelerin hesaplanış şekline göre Basit Bootstrap, Çift Bootstrap, Ağırlıklı Bootstrap, Tekrarlamalı Bootstrap gibi farklı isimler almaktadır.¹⁸⁷

En genel şekli ile veri kümesinde yerine koyma yöntemi ile örnekler seçilerek öğrenme kümesi oluşturulur. n örnekten oluşan veri kümesinden yerine konularak yine n örnek seçilir. Bu küme öğrenme kümesi olarak ele alınmaktadır. Öğrenme kümesinde yer almayan örnekler sınama kümesi olarak kullanılır. n örnekten oluşan veri kümesinde bir örneğin seçilmeme olasılığı $1 - \frac{1}{n}$ ve sınama kümesinde yer alma olasılığı $\lim_{n \rightarrow \infty} \left(1 - \frac{1}{n}\right)^n = e^{-1} = 0,368$ 'dir. Diğer bir ifadeyle, öğrenme kümesi veri kümesindeki örneklerin %63,2'sinden oluşmaktadır. Buradan modelin başarısını sadece sınama kümesi kullanarak belirlemek kötümser bir yaklaşımdır. Model başarısı hem öğrenme kümesinin hem de sınama kümesinin başarısı ile değerlendirilmelidir. Bunun için $hata = 0,632hata_{sınama} + 0,368hata_{öğrenme}$ değeri hesaplanır ve işlem birkaç kez

¹⁸⁶ R. Wehrens, H. Putter, L. M. C. Buydens, "The Bootstrap: A Tutorial", **Chemometrics and Intelligent Laboratory Systems**, Vol: 54, 2000, pp. 35–52.

¹⁸⁷ M. H. Aslan, F. Yıldırım, "Endüstri Mühendisliği Yaşam Süresi Analizlerinde Yeniden Örnekleme Yöntemlerinin Kullanılması", **2. Mühendislik ve Teknoloji Sempozyumu**, 30 Nisan- 1 Mayıs 2009, Çankaya Üniversitesi, s. 169.

tekrarlanıp hataların ortalaması bulunarak diskriminant fonksiyonunun dışsal geçerliliği test edilmektedir.¹⁸⁸

Diskriminant fonksiyonunun dışsal geçerliliğini ölçmek için kullanılan Gizleme, U ve Bootstrap yöntemleri arasından son dönemde, bilgisayar paket programların gelişimiyle de birlikte en yaygın kullanılan Bootstrap yöntemidir.

¹⁸⁸ Ş. G. Öğüdücü, “Veri Madenciliği Farklı Sınıflandırma Yöntemleri”, (Çevirimiçi) <http://web.itu.edu.tr/~sgunduz/courses/verimaden>, 15 Ağustos 2011.

BÖLÜM 3

İŞLETME PERFORMANSINI ETKİLEYEBİLECEK FAKTÖRLERİN ARAŞTIRILMASI

Genel bir tanımlama ile işletme performansı, faaliyetlerinin kalitesi olarak ifade edilebilir. İşletmelerin performansı, yöneticilerin maliyet kar dengesini en iyi şekilde yönetmesiyle ilgilidir¹⁸⁹. İşletmelerin performanslarının değerlendirilmesinde çok farklı kriterler kullanılmaktadır. Performansı etkileyen faktörlerin en iyi şekilde saptanması, işletmenin belirlediği finansal ve faaliyetlerine ilişkin hedeflerine ulaşmak için kaynaklarını en etkin şekilde kullanması açısından önemlidir.

Bu çalışmada işletmelerin performanslarını etkileyebilecek değişkenler çeşitli performans göstergeleri kombinasyonlarıyla analiz edilerek işletmelerin belirlediği hedefler doğrultusunda hangi değişkenlerin önemli olduğu saptanmaya çalışılmıştır.

3.1 İşletme Performans Kriterleri

Sosyal ve yardım amaçlı kuruluşlar dışında işletmelerin amacı mali dönemlerinin sonunda elde ettikleri karı en yüksek düzeye taşımaktır (maksimize etmektir).¹⁹⁰ İşletmelerin “karlılık oranları”nın yüksek olması, işletmelerin ulaşması gereken önemli hedeflerden biri olmakla birlikte

¹⁸⁹ Anil M. Pandya, Narender V. Rao, “Diversification and Firm Performance an Empirical Evaluation”, **Journal of Financial and Strategic Decisions**, V.II. N. 2, 1998, p. 67.

¹⁹⁰ N. Bayyurt, “İşletmelerde Performans Değerlendirmenin Önemi ve Performans Göstergeleri Arasındaki İlişkiler”, **Sosyal Siyaset Konferansları Dergisi**, Sayı 53, 2007/2 Prof. Dr. Haşmet Başar’a Armağan Özel Sayı, s. 580.

işletmelerin sürekliliğini sağlamak için yeterli değildir. Karın maksimize edilmesi, işletmelerin faaliyetleri ya da finansal kararları ile bağlantılı olarak gerçekleşebilir. İşletmelerin karlarını kısa vadeli yatırımlarla (hisse senedi, hazine bonusu vb.) yükseltebilseler bile bu durum işletmelerin, faaliyetlerinden elde edecekleri nakit girişlerini olumsuz yönde etkileyerek taşıdıkları riski arttırabilir. Riskin artması ise diğer bir performans göstergesi olan “büyüme”yi etkileyecektir. Sadece büyüme odaklı bir hedef belirlemek ise işletmelerin kontrolsüz borçlanmasına neden olabilir.

Sanayi işletmelerinde, “verimlilik” de yine bir performans göstergesi olarak ele alınmaktadır. Kaynakların verimli bir şekilde kullanılması sonucunda işletmelerin maliyetlerinde önemli düşüşler meydana gelerek kapasite kullanım oranında ve toplam karında artışlar gerçekleşmektedir. Verimlilik artışı ise kapasite kullanım oranlarını yükseltmekte ve birim maliyetlerde önemli düşüşler elde edilmesine neden olabilmektedir. Sanayi işletmelerinin verimli üretim yapmaları ile ekonomide; sürekli, sağlıklı ve dengeli gelişme gerçekleşebilecek, enflasyonla mücadele, ihracatı arttırma ve refahı yükseltme de mümkün olabilecektir.¹⁹¹ İşletmelerin verimlilik hedefi ekonomiye ve işletmelere birçok fayda sağlamasına rağmen tek başına yeterli bir başarı göstergesi değildir.

İşletmeler için farklı karlılık, büyüme, verimlilik vb. hedefler mümkünse de işletmelerin tek bir hedefe yönelmeleri işletmeler için yetersiz ve hatta riskli olmaktadır. Bu nedenle işletmeler performans göstergesi olarak kabul edilen bu hedeflerin hepsini birlikte düzeltmeyi amaçlamalıdır.¹⁹²

3.1.1 Karlılık Oranları

İşletmenin faaliyetleri ile sonucunda elde etmiş olduğu karların karşılaştırıldığı bu oranlar faaliyetlerin ne kadar etkin olduğunu ortaya koymaktadır. Karlılık

¹⁹¹ M. Tekin, “Sanayi İşletmelerinde Verimlilik ve Önemi”, **Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi**, Sayı: 1, 1992, s. 170.

¹⁹² N. Bayyurt, **a.g.e.**, s. 581.

oranları işletme sahipleri açısından, yabancı kaynakların sahiplerine ödenen anapara ve faizlerin işletmenin faaliyetleri sonucunda elde ettiği kardan mı yoksa düşük karlılık oranı nedeniyle öz sermayeden mi karşılandığını ortaya koyması sebebiyle önemlidir. Diğer taraftan bu oranlar işletmenin piyasadaki başarısı açısından yöneticiler, personel, kamu kuruluşları, kredi verenler, potansiyel ortaklıklar tarafından da dikkatle izlenilmelidirler.¹⁹³

Bilançonun farklı kalemlerinin kullanılmasıyla elde edilen çeşitli karlılık oranı hesaplamaları mevcuttur. Literatürde sık kullanılan karlılık oranlarına aşağıda değinilmiştir.

Öz Sermaye Karlılığı:

Öz sermaye karlılığı farklı kar kalemleri ile hesaplanabilmektedir.

$$\text{Öz Sermaye Karlılığı} = \frac{\text{Net Dönem Karı}}{\text{Öz Sermaye}}$$

oranın sonucu, işletme sahipleri açısından öz sermayenin alternatif kullanım alanlarının değerlendirilmesi nedeniyle önem taşımaktadır. Alternatif kullanım alanlarının kar getirisi bu orandan daha yüksek ise öz sermaye sahipleri alternatif alanlara yöneleceklerdir.¹⁹⁴

Ekonomik Karlılık Oranı:

$$\text{Ekonomik Karlılık} = \frac{(\text{Vergi Öncesi Kar} + \text{Finansal Giderler})}{\text{Toplam Varlıklar}}$$

¹⁹³ M. Tükenmez, vd., **Finansal Yönetim**, Vizyon Eğitim ve Danışmanlık Ltd. Şti., İzmir 1999, ss. 413–414.

¹⁹⁴ A.e., s. 418.

olarak hesaplanan bu oran toplam kaynakların ne ölçüde karlı kullanıldığını göstermekte¹⁹⁵ ve işletmeye yatırılan fonların getirisini ölçmekte kullanılmaktadır.

Aktif Karlılık:

$$\text{Aktif Karlılık} = \frac{\text{Net Dönem Karı}}{\text{Toplam Varlıklar}}$$

olarak ifade edilen bu oran, varlıkların ne ölçüde verimli kullanıldığını göstermektedir ve işletmelerin performans göstergelerinden olan çeşitli “karlılık oranı” hesaplamaları içinde önemli bir yer almaktadır.

Faaliyet Kar Marjı:

$$\text{Faaliyet Kar Marjı} = \frac{\text{Faaliyet Karı}}{\text{Satış Gelirleri}}$$

şeklinde hesaplanan faaliyet kar marjının yüksek olması, işletmenin ana faaliyet konusunda başarılı olduğunu ifade etmektedir. Ancak oranın, faaliyet dışı gider ve zararlar ile vergiyi karşıladıktan sonra yeterli ölçüde dönem net karı bırakacak büyüklükte olması önem taşımaktadır.¹⁹⁶

Brüt Kar Marjı:

İşletmeler açısından

$$\text{Brüt Kar Marjı} = \frac{\text{Brüt Kar}}{\text{Satış Gelirleri}}$$

şeklinde hesaplanan bu oranın yeterliliği, aynı sektördeki işletmelerin oranları ile karşılaştırılarak ve işletmenin geçmiş yıllardaki oranları dikkate alınarak

¹⁹⁵ M. Tükenmez, vd., a.g.e., s. 419.

¹⁹⁶ N. Kurnaz, “Finansal (Mali) Tablolar Analizi”, (Çevrimiçi) <http://www.niyazikurnaz.net>, 12 Temmuz 2011.

belirlenebilmektedir¹⁹⁷. İşletmelerin brüt kar marjının değişmesinde hangi etkenlerin geçerli olduğu saptanmalıdır¹⁹⁸. Satışlar ve satılan ürün maliyeti bu oranın yüksek çıkmasında önemli bir nedendir.

Net Kar Marjı:

Aşağıda gösterildiği gibi hesaplanan

$$\text{Net Kar Marjı} = \frac{\text{Net Dönem Karı}}{\text{Satış Gelirleri}}$$

'nın yüksek çıkması işletmenin satışlarını yüksek karla gerçekleştirdiğini göstermektedir. Oranın yeterliliği için belirli bir standart yoktur ancak aynı sektördeki diğer işletmelere ve geçmiş yıllardaki oranlara göre değerlendirilebilmektedir.

Vergi Öncesi Kar Marjı:

$$\text{Vergi Öncesi Kar Marjı} = \frac{\text{Vergi Öncesi Kar}}{\text{Satış Gelirleri}}$$

olarak hesaplanan bu gösterge de net kar marjı ile aynı şekilde yorumlanmaktadır. İki oranın da yüksek çıkması beklenmektedir.

Faiz Karşılama Oranı:

İşletmenin faaliyetleri sonucunda elde ettiği fonlar aracılığı ile ödemek zorunda olduğu faizleri kaç defa kazandığını ölçmek amacıyla kullanılmaktadır¹⁹⁹.

¹⁹⁷ M. Tükenmez, vd., **a.g.e.**, s. 415.

¹⁹⁸ Ö. Akgüç, **Finansal Yönetim**, Avcıol Basım, Yenilenmiş 7. Baskı, İstanbul 1998, s. 71.

¹⁹⁹ M. Tükenmez, vd., **a.g.e.**, s. 425.

İşletmenin faiz giderlerini karşılayamaması, işletmeyi mali açıdan zor durumda bırakarak iflasına neden olabilmektedir.²⁰⁰

$$\text{Faiz Karşılama Oranı} = \frac{\text{Faaliyet Karı}}{\text{Finansal Giderler}}$$

3.1.2 Verimlilik

Verimlilik, üretilen miktarla bu üretimde kullanılan herhangi bir kaynak arasındaki oranı göstermektedir. Üretim faktörlerinin her biri için verimlilik, üretim faktörlerinin sağladığı artışla ifade edilebilmektedir. İşletme performans göstergelerinden olan “verimlilik” için de karlılık oranının da olduğu gibi farklı yaklaşımlar kullanılmaktadır. Bu yaklaşımlardan emek (işgücü) verimliliği belirli bir dönemde mevcut işgücünün üretim miktarında sağladığı artışın dönem başındaki toplam üretime oranlanmasıyla hesaplanmaktadır. Makine verimliliği ise işletmede belirli bir dönemde mevcut makinelerle yapılan üretimdeki artışın dönem başındaki toplam üretime oranlanmasıyla elde edilmektedir. Diğer bir yaklaşımda sermaye verimliliği yani rantabilite (karlılık), mevcut sermayede belirli bir dönemdeki artışın dönembaşındaki sermaye miktarına oranlanması sonucu elde edilmektedir.²⁰¹

Verimliliğin artırılması, özellikle sanayi kesiminde ülkenin gelişimini sürdürebilmesi açısından önem taşımaktadır. Verimlilikteki düşüş, enflasyon seviyesi ve istihdamı olumsuz yönde etkilediği gibi şirket düzeyinde de karlılığı azaltmaktadır. Bu sebeple özellikle son yıllarda makro ve mikro düzeyde verimlilik artışını sürekli kılmak önemli olmuştur.²⁰²

²⁰⁰ N. Akdoğan, N. Tenker, **Finansal Tablolar ve Mali Analiz Teknikleri**, 4. Baskı İstanbul, 1994, s. 452.

²⁰¹ M. Tekin, **a.g.e.**, s. 170.

²⁰² A. Tezeren, **İmalat Sanayinde Verimliliği Etkileyen Faktörler**, 319 Nolu Milli Prodüktivite Merkezi Yayınları, Ankara, 1985.

Kaynakların verimli bir şekilde kullanılması sonucunda işletmelerin kapasite kullanım oranında ve toplam karında önemli artışlar sağlanabilmektedir.²⁰³

Emek verimliliği, sermaye verimliliği, hammadde verimliliği gibi farklı girdilerin verimliliği gibi kısmi verimlilikleri hesaplamak mümkündür. Bunun dışında verimlilik ölçüsü olarak çok faktörlü verimlilik ve toplam faktörlü verimlilik oranları da kullanılabilir.* Ancak hesaplamadaki zorluklar nedeniyle en çok kullanılan kısmi verimlilik oranları olup bu oranlar içinde de emek verimliliğidir ve aksi belirtilmedikçe verimlilik ifadesiyle emek verimliliğinden bahsedilmektedir. Bu çalışmada işletme performans göstergelerinden olan verimlilik için emek kullanılmış ve oran,

$$\text{Emek Verimliliği} = \frac{\text{Katma Değer}}{\text{Çalışan Sayısı}}$$

olarak hesaplanabilmektedir.

3.1.3 Büyüme Oranları

İşletmelerin büyüme oranları; satışlardaki artış, vergi öncesi kar (dönem karı), öz sermaye varlık (aktif) toplamı ve hisse başına kar olarak hesaplanabilmektedir.

Satışlardaki Artış:

Satışlardaki artış cari fiyatlarla hesaplanmaktadır.²⁰⁴

$$\Delta S = \frac{S_t - S_{t-1}}{S_{t-1}} * 100$$

²⁰³ M. Tekin, a.e., ss. 171–172.

* Ayrıntılı bilgi için bkz. Bilge Acar Bolat, **Firma Performansını Etkileyen Faktörlerin Yapısal Eşitlik Modeli ile İncelenmesi**, Basılmış Doktora Tezi, İstanbul, 2010, ss. 32–39.

²⁰⁴ Ö. Akgüç, **Finansal Yönetim**, a.g.e., ss. 80–81.

ΔS ; net satışlardaki % artış, S_t ; t . dönemdeki cari fiyatlarla net satış tutarı, S_{t-1} ; bir önceki dönemdeki net satış tutarını belirtmektedir.

Fiyat artışlarının hızlı olduğu yani fiyat hareketlerinin istikrarlı olmadığı ekonomilerde, cari fiyatlara göre hesaplanan artış hızlarının, fiyat artış hızı dikkate alınarak düzeltilmesi gerekmektedir. Bu amaçla cari fiyatlar yerine reel fiyatlarla hesaplanan artışlar kullanılmaktadır.

$$\Delta S_r = \frac{\%100 + \% \Delta S}{\%100 + \% \Delta F} - 1$$

ΔS_r ; bir önceki döneme göre satışlardaki reel olarak artış oranını, ΔF ; bir önceki yıla göre fiyatlardaki ortalama artış yüzdesini ifade etmektedir.

Dönem Karında Artış:

Cari fiyatlarla dönem karındaki veya vergi öncesi kardaki artış oranı aşağıdaki gibi hesaplanmaktadır.²⁰⁵

$$\Delta k = \frac{K_t - K_{t-1}}{K_{t-1}} * 100$$

Δk ; cari fiyatlarla dönem karındaki % artış, K_t ; cari fiyatlarla dönem karını, K_{t-1} ; bir önceki dönem karını ifade etmektedir.

Enflasyonun etkisinden arındırılmak amacıyla reel kar artış hızı (Δk_r) hesaplanmaktadır.

$$\Delta k_r = \frac{\%100 + \% \Delta k}{\%100 + \% \Delta F} - 1$$

²⁰⁵ Ö. Akgüç, **Finansal Yönetim**, a.g.e., ss. 81–82.

Öz Sermaye Artışı:

$$\Delta\ddot{o} = \frac{\ddot{o}_t - \ddot{o}_{t-1}}{\ddot{o}_{t-1}} * 100$$

$\Delta\ddot{o}$; öz sermaye artışı, \ddot{o}_t ; cari dönem öz sermaye tutarı, \ddot{o}_{t-1} ; bir önceki dönem öz sermaye değeridir.

Bir işletmenin öz sermayesi, bir akışı değil belirli bir tarih itibariyle stoku, başka bir ifadeyle birikmeli bir tutarı gösterdiğinden sabit fiyatlarla veya reel olarak artışı hesaplamak için tüm bilanço kalemlerinin enflasyona göre düzeltilmesi gerekmektedir.²⁰⁶

Varlık (Aktif) Artışı:

Varlık toplamındaki artış cari fiyatlarla

$$\Delta v = \frac{v_t - v_{t-1}}{v_{t-1}} * 100$$

olarak hesaplanmaktadır.

v_t ; cari dönem varlık tutarı, v_{t-1} ise bir önceki dönem varlık tutarını ifade etmektedir. “*Varlıklarda reel artışı hesaplayabilmek için her bir varlık kaleminin hangi tarihte işletmenin varlık tutarına dahil edildiğinin bilinmesi ve o tarih esas alınarak fiyat artışlarına göre düzeltme yapılması gerekmektedir.*”²⁰⁷

İşletme bazlı incelendiğinde büyüme kavramının ekonomik büyüme ve buna bağlı olarak istihdam, verimlilik ve gelişen teknoloji ile ilişkili olduğu görülmektedir. İşletmelerin büyümesini etkileyen sadece kendi iç dinamikleri ve faaliyet alanları değil aynı zamanda faaliyet gösterdiği ülkenin belirlediği

²⁰⁶ Ö. Akgüç, **Finansal Yönetim**, a.g.e., s. 83.

²⁰⁷ A.e., s. 83.

ekonomi stratejileridir. Bu bağlamda ülkenin büyüme stratejileri işletmelerin büyümesini etkileyebilmesine rağmen, işletmelerin büyüme konusu özellikle iktisat teorisinde hakettiği şekilde incelenmemektedir. Mikro ekonomik analizlerde işletmelerin faaliyet gösterdikleri endüstriyel kolların dinamiklerine ilişkin bilgiler yetersiz kalmaktadır.²⁰⁸

İşletmelerin büyümesini etkileyen faktörlerin belirlenmesi ve daha iyi yorumlanması işletme politikalarının oluşturulmasında önemli etkenlerden biridir. Ülkelerin belirledikleri ekonomik stratejiler işletmelerin bu ekonomik faaliyetleri dikkate almasını ve bu stratejiler yönünde hareket etmelerini gerektirmektedir. Türkiye gibi ihracata dayalı olarak büyüyen ülkelerde,²⁰⁹ ihracatın ithalatı karşılama oranı yıllar itibariyle aşağıdaki Tablo 2.1.'de görülebilir. Yıllar itibariyle bu oranın yükselmesi Türkiye'nin ihracat yoluyla elde edilen gelirin ithalatı karşıladığı anlamına gelmektedir. İhracatın ithalatı finanse etmesinin derecesi ihracatın ithalatı karşılama oranıyla açıklanır.²¹⁰

²⁰⁸ M. Karagöz, H. Demirgil, "Firma Büyüme Performansını Etkileyen Faktörler", **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi**, 2009, C. 14, S.2, s. 48.

²⁰⁹ S. Çiftçioğlu, C. Karaaslan, H. Demir, "Saving, Investment, Exports and Economic Growth in Turkey", **Review of Social, Economic&Business**, Vol: 3-4, ss. 20-30.

C. Bilgin, A. Şahbaz, "Türkiye'de Büyüme ve İhracat Arasındaki Nedensellik İlişkisi", **Gaziantep Üniversitesi Sosyal Bilimler Dergisi**, 2009, (8).1, ss. 177-198.

²¹⁰ (Çevrimiçi) <http://www.havvatunc.com/ticaret-acigi-buyuyor-ihracatin-ithalati-karsima-orani-kuculuyor-1107.html>, 10 Eylül 2011.

Tablo 2.1 Yıllara göre İhracatın İthalatı Karşılama ve İhracat Değişim Oranları

Dönem	Yıllara göre İhracatın İthalatı Karşılama Oranı (%)	Yıllara göre İhracat (Değişim Oranı) (%)	Dönem	Yıllara göre İhracatın İthalatı Karşılama Oranı (%)	Yıllara göre İhracat (Değişim Oranı) (%)
2010	61.4	11.5	2003	68.1	31.0
2009	61.4	11.5	2002	69.9	15.1
2008	65.4	23.1	2001	75.7	12.8
2007	63.1	25.4	2000	50.9	4.5
2006	61.3	16.4	1999	65.4	-1.4
2005	62.9	16.3	1998	58.7	2.7
2004	64.8	33.7			

Kaynak: <http://www.tuik.gov.tr/Gosterge.do?metod=IlgiliGosterge&id=3489>

İhracat yolu ile artan gelirler ve Türkiye’de işletmeler arasındaki rekabet nedeniyle satış hacimlerinde ihracatın payının önemini arttırmıştır. İşletmelerin ihracat hacimlerindeki artış büyüme stratejilerine önem vermelerini zorunlu kılmaktadır. Bu artan dış gelirlerle beraber işletmelerin döviz kuru rejimleri, enflasyon gibi ekonomik politikalarından etkilenme riskini arttırabilmektedir.

İşletmelerin performanslarını arttıran, yenilikçi faaliyet ve ürünün geliştirilmesine yönelik düzenli Ar-Ge faaliyetleridir. Bir işletmenin hızlı büyümesi yenilikçi faaliyetlerinin sağladığı olumlu katkıların bir sonucu olarak değerlendirilebilmektedir. Piyasadaki mevcut işletmelerin pazar paylarını alabilen hızlı büyüyen yeni işletmeler ürün ve hizmetlerin daha etkin yürütülmesi için yeni yöntemlerin kullanılmasında daha başarılı olabilmektedirler.²¹¹

²¹¹ M. Karagöz, H. Demirgil, **a.g.e.**, ss. 61–65.

3.1.4 Piyasa Performans Oranları

Hisse Başına Düşen Kar:

Bu oran işletme sahip ve ortaklarının yaptıkları yatırımlar karşılığında yeterli gelir elde edip etmediklerini ifade etmektedir ve vergi sonrası net kardan, imtiyaz veren kurucu hisseler, yönetim kurulu ve çalışanlara ödenen temettü çıkarıldıktan sonra, kalan net karın ortalama hisse senedi sayısına bölünmesi ile bulunmaktadır.²¹²

$$\text{Hisse Başına Düşen Kar} = \frac{\text{Net Kar}}{\text{Hisse Senedi Sayısı}}$$

Fiyat-Kazanç Oranı:

Bu oran hisse senedi fiyatının hisse başına kar değerine oranıdır.²¹³

$$\text{Fiyat - Kazanç} = \frac{\text{Piyasa Değeri}}{\text{Net Kar}}$$

Fiyat-Kazanç oranı yatırımcılar arasında en çok kullanılan değerlendirme yöntemidir. Bu oranın fazla kullanılmasının temel sebepleri şunlardır:²¹⁴

- Bu Fiyat-Kazanç oranı kullanıldığında, geleneksel yöntemlerde* varsayımı ya da hesabı yapılan risk, büyüme oranı ve temettü ödeme oranı gibi bir takım verileri kullanmaya gerek kalmamaktadır.
- Yatırımcılar açısından hesaplanması kolaydır.
- Risk ve büyüme gibi verileri yansıtmaktadır.

²¹² Sermaye Piyasası ve Borsa Temel Bilgiler Kılavuzu, İstanbul Menkul Kıymetler Borsası Yayınları, 17. Basım, Temmuz 2002, s.141.

²¹³ A.e., s. 141.

²¹⁴ M. B. Karan, Yatırım Analizi ve Portföy Yönetimi, Gazi Kitapevi, Ankara, 2001, s. 356.

* Geleneksel yöntemler: Şirketlerin mali tabloları kullanılarak elde edilen oranla işletme performansının değerlendirilmesi yöntemidir.

Piyasa Değeri/Defter Değeri Oranı (PD/DD):

Bu oran bir işletmenin hisse senedi fiyatının hisse başına öz sermaye değerine bölünmesiyle elde edilmektedir.²¹⁵

$$PD/DD \text{ Oranı} = \frac{\text{Piyasa Değeri}}{\text{Öz Sermaye}}$$

Düşük öz sermaye karlılığı ve yüksek PD/DD oranına sahip senetler pahalı olarak kabul edilirken, yüksek öz sermaye karlılığına ve düşük PD/DD oranına sahip senetler ucuz olarak kabul edilmektedir.²¹⁶

3.2 İşletme Performansını Etkileyebilecek Faktörler

3.2.1 Finansal Oranlar

Bir işletmenin mali durumu değerlendirilirken işletmenin mali tablolarında görünen değerlerden ziyade bilanço ve gelir tablosunda yer alan kalemler arasındaki ilişkiler daha anlamlıdır. Bu nedenle finansal analizlerde oranlardan geniş ölçüde yararlanılmaktadır. Oranlar incelenirken piyasadaki aynı sektörde yer alan diğer işletmelerle ve geçmiş yıllarla karşılaştırmalı olarak yorumlanmalıdır.

Finansal oranlar, işletmelerin yalnız geçmiş ve cari mali durumunu değerlendirmek açısından değil, planlama ve denetim işlevini yerine getirmede ve şirket ortaklarının, pay senetlerine yatırım yapmak isteyen birikim sahiplerinin ve finansman kurumlarının, işletmenin mali durumunu ve gücünü nasıl gördüklerini, değerlendirdiklerini anlamak için de finans yöneticisine yarar sağlamaktadır.²¹⁷

²¹⁵ M. B. Karan, **a.g.e.**, ss. 359–361.

²¹⁶ **A.e.**, s. 361.

²¹⁷ Ö. Akgüç, **Finansal Yönetim, a.g.e.**, s. 20.

Farklı şekillerde gruplandırılabilen finansal oranlar bu çalışmada işletme faaliyetlerinin değerlendirilmesindeki kullanım biçimine göre ele alınmıştır ve Likidite Oranları, Finansal Yapı Oranları, Faaliyet Oranları ve Karlılık Oranlarını kapsamaktadır.²¹⁸ Karlılık Oranları işletme performans göstergeleri başlığı altında ele alınmıştır.

3.2.1.1 Likidite Oranları

Likidite oranları, işletmenin kısa süreli borçlarını ödeme gücünü ölçmek, işletme sermayesinin yeterli olup olmadığını saptayabilmek için kullanılmaktadır. Kısa sürede borçlarını ödeyebilecek yeterli likiditeye sahip olmanın, özellikle ekonomik bunalım dönemlerinde şirketin faaliyetini sürdürmesi açısından büyük önemi vardır.²¹⁹ Literatürde en çok kullanılan ve tartışılan likidite oranlarının kullanıldığı bu çalışmadaki oranlar şu şekildedir;

Cari Oran:

Bu oran işletmenin kısa vadeli borç ödeme gücü hakkında bilgi vermektedir.²²⁰

$$Cari\ Oran = \frac{Dönen\ Varlıklar}{Kısa\ vadeli\ Yükümlülükler}$$

Cari oran hesaplanmasındaki amaç, işletmenin kısa vadeli borçlarını ödeme gücünü ölçmek ve net işletme sermayesinin yeterli olup olmadığını saptayabilmektir. Net işletme sermayesi, dönen varlıklar ile kısa vadeli yükümlülükler arasındaki mutlak farktır; başka bir ifadeyle işletmenin tüm kısa vadeli borçlarını ödedikten sonra kalan işletme sermayesi tutarıdır.²²¹

²¹⁸ M. Tükenmez, vd., **a.g.e.**, 384.

²¹⁹ Ö. Akgüç, **Finansal Yönetim**, **a.g.e.**, s. 23.

²²⁰ **Sermaye Piyasası ve Borsa Temel Bilgiler Kılavuzu**, s.139.

²²¹ Ö. Akgüç, **Finansal Yönetim**, **a.g.e.**, s. 24.

Asit-test Oranı (Likidite Oranı):

$$\text{Asit - test Oranı} = \frac{(\text{Dönen Varlıklar} - \text{Stoklar})}{\text{Kısa vadeli Yükümlülükler}}$$

eşitliği ile hesaplanan bu oran, kısa vadede nakde dönüşebilecek dönen varlıkların, kısa vadeli borçları karşılama gücünü göstermektedir.²²² Likidite oranının payında paraya çevrilmesi zaman alabilecek olan değerlere yer verilmediğinden, bu oran cari orana kıyasla daha duyarlı bir ölçü olarak kabul edilmektedir.²²³

Batılı finans kurumlarında bu oranın 2 olması genel kabul gördüğü halde, yüksek enflasyonlu ve kit fon kaynaklarına sahip ülkelerde endüstri ve sektörlerin değişik özellikleri göz önünde bulundurulduğunda cari oranın 1,5 olması genellikle yeterli olarak kabul edilmektedir.²²⁴

Çalışmada kullanılan veriler incelendiğinde veri setini oluşturan işletmelerden yüksek likidite oranına sahip sektörlerin taş, toprak ve gıda sektörü olduğu gözlenmiştir. Bu oranın yüksek olması işletmelerin faaliyet alanları ile doğrudan ilişkilidir.

Nakit Oranı:

Nakit oran, işletmenin satışlarının durması ve alacaklarını tahsil edememesi durumunda, kısa süreli borçlarını karşılayabilme gücünü göstermektedir. Bu oran hazır değer oranı olarak da bilinmektedir.²²⁵

$$\text{Nakit Oranı} = \frac{(\text{Nakit ve nakit benzerleri} + \text{Finansal yatırımlar})}{\text{Kısa vadeli Yükümlülükler}}$$

²²² **Sermaye Piyasası ve Borsa Temel Bilgiler Kılavuzu**, s. 139.

²²³ Ö. Akgüç, **Finansal Yönetim, a.g.e.**, s. 28.

²²⁴ (Çevirimiçi) <http://www.bilgaz.net/dosyalar/OranAnalizi.pdf>, 10 Ağustos 2011.

²²⁵ Ö. Akgüç, **Finansal Yönetim, a.g.e.**, s. 30.

Stokların Toplam Varlıklara Oranı:

$$\text{Stokların/Toplam Varlıklara Oranı} = \frac{\text{Stoklar}}{\text{Toplam Varlıklar}}$$

Toplam varlıkların içinde stok oranını gösteren bu oran, stokların varlık yapısındaki durumunu saptamak için kullanılmaktadır. Bu oran işletmelerin faaliyetlerine ilişkin bilgi sağlamaktadır. Bu oranın yüksek olması işletmelerin yüksek stok hacmi ile çalıştığı ve stoklarını eritme konusunda sıkıntı yaşadığı şeklinde yorumlanabilmektedir.

Kısa vadeli (K.Vd.) Ticari Alacakların Toplam varlıklara Oranı:

$$\begin{aligned} & \text{K.Vd.Tic. Alacakların/Toplam Varlıklara Oranı} \\ & = \frac{\text{K.Vd.Ticari Alacaklar}}{\text{Toplam Varlıklar}} \end{aligned}$$

Bu oranın yüksek olması işletmelerin alacak politikalarına ilişkin risk taşıdıklarını göstermektedir.

3.2.1.2 Finansal Yapı Analizinde Kullanılan Oranlar

İşletmenin varlıklarını hangi kaynaklardan elde ettiğini, kaynakların yapısı ile dağılımını ve uzun vadedeki borçlarını ödeme gücünü analiz eden oranlar bu başlık altında toplanmaktadır. İşletmenin dönen ve duran varlıklarının yabancı kaynak veya özkaynaklar aracılığı ile finanse edilmesindeki denge; dolayısıyla özkaynak, yabancı kaynak dengesinin sağlanması, yabancı kaynak finansmanındaki anapara ve faiz ödemelerinde işletmenin güçlü bulunması anlamına gelmektedir.²²⁶

²²⁶ M. Tükenmez, vd., **a.g.e.**, s. 393.

Finansal Kaldıraç Oranı:

$$\text{Finansal Kaldıraç Oranı} = \frac{(\text{Kısa vadeli Yük.} + \text{Uzun vadeli Yük.})}{\text{Toplam Varlıklar}}$$

İşletmenin kaynakları arasında işletmeye sabit yük getiren kaynakların varlığı, finansal kaldıraçın nedenini oluşturmaktadır.²²⁷ Finansal kaldıraç oranının yüksek çıkması öz sermayenin yetersiz olduğu anlamına geldiği için kredi verenlerin alacaklarının tahsili riskli duruma gelebilmektedir. Yani işletmenin borçlarını ödeyememe riskinin yüksek olduğu ortaya çıkmaktadır.²²⁸

Kısa Vadeli Yabancı Kaynakların (K.Vd. Yük.) Toplam varlıklara Oranı:

$$\text{K.Vd. Yük./Toplam Varlıklara Oranı} = \frac{\text{Kısa vadeli Yük.}}{\text{Toplam Varlıklar}}$$

Eşitliği ile hesaplanan bu oran, işletmenin iktisadi varlıklarının yüzde kaçının kısa süreli yabancı kaynaklarla finanse edildiğini ortaya koymaktadır. Bu oranın yüksek olması, iktisadi varlıkların büyük bir bölümünün kısa süreli borçlarla finanse edildiğini göstermektedir. Bir işletme gelecekte ödeme güçlükleri ile karşılaşmak istemiyorsa, kısa vadeli borçlarla duran varlıklarını sürekli olarak finanse etmekten kaçınmalıdır.²²⁹

Uzun Vadeli Yabancı Kaynakların (U.Vd. Yük.) Devamlı Sermayeye Oranı:

$$\begin{aligned} &\text{U.Vd. Yük./Devamlı Sermayeye Oranı} \\ &= \frac{\text{Uzun Vadeli Yük.}}{\text{Uzun Vadeli Yük.} + \text{Öz Sermaye}} \end{aligned}$$

Bu oranın yüksek (%25'den fazla) olması yabancı kaynakların oranının yüksek olduğuna ya da işletmelerde büyük oranda uzun vadeli yabancı kaynak

²²⁷ Ö. Akgüç, **Finansal Yönetim, a.g.e.**, s. 32.

²²⁸ M. Tükenmez, vd., **a.g.e.**, s. 394.

²²⁹ Ö. Akgüç, **Mali Tablolar Analizi**, Avcıol Basım-Yayın, İstanbul, Genişletilmiş 9. Bası, 1995, ss. 363–364.

kullanıldığına işaret etmektedir.²³⁰ Uzun süreli yabancı kaynaklarla varlıklarını finanse eden işletmeler, özellikle durgunluk dönemlerinde borç taksitlerini ödemedede büyük güçlüklerle karşılaşabilirler.²³¹

Duran Varlıkların Öz Sermayeye Oranı:

$$\text{Duran Varlıkların/Öz Sermayeye Oranı} = \frac{\text{Duran Varlıklar}}{\text{Öz Sermaye}}$$

Duran varlıkların öz sermayeye bölünmesiyle hesaplanan bu oranın düşük olması durumunda maddi duran varlıkların öz kaynaklarla karşılanabildiği kabul edilmektedir.²³²

Dönen Varlıkların Toplam Varlıklara Oranı:

$$\text{Dönen Varlıkların/Toplam Varlıklara Oranı} = \frac{\text{Dönen Varlıklar}}{\text{Toplam Varlıklar}}$$

Dönen varlıkların toplam varlıklara bölünmesiyle hesaplanan bu oranın yüksek olmaması (%50'nin altında), kaynakların çoğunluğunun duran varlıklara bağlandığı anlamına gelmektedir. Bu durumda işletmenin kaynak sıkıntısına düşmesi muhtemeldir.²³³

3.2.1.3 Faaliyet Oranları

Bu oranlar, işletmenin faaliyetlerinde kullandığı varlıklarını etkin bir şekilde kullanıp kullanmadığını ölçmektedir. Bu çalışmada literatürde incelenen ve hesaplanması mümkün bütün faaliyet oranları ele alınmıştır. Çalışmada kullanılan faaliyet oranları şu şekildedir;

²³⁰ N. Kurnaz, (Çevirimiçi) www.niyazikurnaz.net, 12 Temmuz 2011.

²³¹ Ö. Akgüç, **Mali Tablolar Analizi**, s. 364.

²³² N. Kurnaz, **a.g.e.**, 12 Temmuz 2011.

²³³ **A.e.**, N. Kurnaz.

Hazır Değerler Devir Hızı:

$$\text{Hazır Değerler Devir Hızı} = \frac{\text{Satış Gelirleri}}{\text{Hazır Değerler}}$$

Bu oranda Hazır Değer, (Nakit ve Nakit Ben. + Finansal Yatırımlar) toplamına eşittir ve hazır değerlerin kaç katı satış yapıldığını göstermektedir.

Hazır Değer Devir Hızı işletmede bulundurulması gerekli nakit miktarını belirlemektedir, ancak bu konuda belirli bir standart yoktur. Bu nedenle aynı sektördeki işletmelerle ve geçmiş yıllarla karşılaştırılarak bir yoruma ve sonuca gidilmeye çalışılması gerekmektedir.²³⁴

Stok Devir Hızı:

$$\text{Stok Devir Hızı} = \frac{\text{Satışların Maliyeti}}{\text{Ortalama Stok}}$$

eşitliği ile ifade edilen bu orandaki Ortalama Stok;

$$\text{Ortalama Stok} = (\text{bir dönem önceki Stoklar} + \text{o döneme ait Stoklar})/2$$

olarak hesaplanmaktadır. İşletmeler düşük fiyat seviyesinde satış hacimlerini arttırarak stok devir hızını yüksek tutabilmektedirler. Bu durum rekabet avantajını arttırmaktadır.²³⁵ Yüksek devir hızı işletmelerin iyi bir stok kontrol sistemi ve etkin bir pazarlama stratejisi olduğunun göstergesi olarak değerlendirilebilir.

İşletmelerin stok devir hızlarının yüksek olması, stokların daha optimal düzeyde tutulduğunu ve kullanıldığını gösterir. Böyle bir durumda işletmeler daha az işletme sermayesi ile daha fazla kar elde etme olanağına sahiptirler; fakat

²³⁴ (Çevirimiçi) <http://www.bacegrup.com/3-4-oran-analizi.ntml>, 10 Temmuz 2011.

²³⁵ N. Kurnaz, a.g.e., 12 Temmuz 2011.

işletmelerin stok devir hızlarının yüksek olması, yetersiz stoklarla çalışmasından da ileri geliyor olabilir.²³⁶

Stok devir hızı stokların verimli kullanımı ve satışlar yoluyla stokların alacaklara dönüşüm hızı hakkında bilgi vermektedir. Stok devir hızının düşmesi nakit akışlarının yavaşlamasına neden olduğundan işletme sermayesi ihtiyacını arttırmaktadır. Stok devir hızının artması nakit akışını arttırdığından ve stok maliyetini azalttığından işletme karlılığı ve verimliliğini arttırmaktadır.

Bir işletme üretimini (faaliyetini) sürdürebilmek amacıyla gerekli üretim girdilerini (hammadde), üretim sürecinde işlemleri tamamlanmamış mamulleri (yarı mamul), üretimini tamamlayıp satış için hazır tuttuğu ürünleri (mamul) stok olarak tutar. Stok devir hızı analizinde amaç stok olarak tutulan bu varlıkların işletme tarafından ne kadar hızla üretim içinde tüketildiği ve satışa hazır hale getirildiği görebilmektir. Bu şekilde stokların belli bir dönem içinde kaç kere yenilendiği ortaya çıkar.²³⁷

Stok devir hızının düşmesi nakit akışlarının yavaşlamasına neden olduğundan işletme sermayesi ihtiyacını arttırmaktadır. Stok devir hızının artması nakit akışını arttırdığından ve stok maliyetini azalttığından işletme karlılığı ve verimliliğini arttırmaktadır. Stok devir hızının düşük seviyede olması işletmenin satışlarında sorunlar yaşadığı ya da uygun stok politikaları kullanmadığı anlamına gelmektedir.

Alacak Devir Hızı:

Alacakların devir hızı bir işletmenin alacaklarını tahsil edebilirliğini yani likiditesini gösteren iyi bir ölçektir.²³⁸

$$\text{Alacak Devir Hızı} = \frac{\text{Satış Gelirleri}}{\text{Ortalama Ticari Alacaklar}}$$

²³⁶ (Çevirimichi) <http://www.bilgaz.net/dosyalar/OranAnalizi.pdf>, 10 Ağustos 2011.

²³⁷ A.e., 10 Ağustos 2011.

²³⁸ Ö. Akgüç, **Mali Tablolara Analizi**, s. 380.

Yukarıdaki eşitlikte paydada yer alan Ortalama Ticari Alacaklar

$$= (\text{bir dönem önceki Ticari Alacaklar} + \text{o döneme ait Ticari Alacaklar})/2$$

olarak hesaplanmaktadır. Alacak devir hızının artması, işletmenin alacaklarını vadesinde tahsil edebildiğini ve vadelerinin kısaldığını ifade etmektedir.²³⁹ Bu oranının yüksek olması alacakların vadeli alacaklar için tahsilat süresinin kısa olduğu anlamına gelmektedir. Yıllar itibariyle azalan alacak devir hızı işletmelerin nakit seviyesini de yükselmesine neden olmaktadır. Alacak devir hızının yüksek olduğu en önemli sektörlerden biri perakende sektörüdür.

Bir işletme alacaklarını hızlı bir şekilde (devir hızının yüksek olması durumu) tahsil edebiliyorsa, likiditesi yüksek kabul edilebilir ve işletme bu sayede hem nakit sıkıntısı içine girmez, hem de alacaklarının değeri fazla erimeden bunları daha iktisadi alanlarda kullanabilir.²⁴⁰

Dönen Varlık Devir Hızı:

$$\text{Dönen Varlık Devir Hızı} = \frac{\text{Satış Gelirleri}}{\text{Dönen Varlıklar}}$$

eşitliğiyle hesaplanan bu oran işletme sermayesi olarak da adlandırılmaktadır. Kısa vadeli borçlarla öz sermaye arasında denge olup olmadığı bu oranla ölçülmektedir. Kısa süreli borçlar uzun süreli borçlardan daha büyük boyutlarda dalgalanma gösterdiği gibi daha kısa aralıklarla yenilenmesi veya ödenmesi gerekmektedir. Bu nedenle bu oranın yüksekliği fonlama açısından daha büyük risk göstergesidir.²⁴¹

²³⁹ Ö. Akgüç, **Finansal Yönetim**, s.44.

²⁴⁰ (Çevirimiçi) <http://www.bilgaz.net/dosyalar/OranAnalizi.pdf>

²⁴¹ Ö. Akgüç, **Mali Tablolara Analizi**, s. 386.

Duran Varlık Devir Hızı:

$$\text{Duran Varlık Devir Hızı} = \frac{\text{Satış Gelirleri}}{\text{Duran Varlıklar}}$$

Bu oranın düşük çıkması duran varlıkların fazlalığına işaret ederken, kaynakların verimsiz kullanıldığının belirtisi olarak da görülmektedir. Oranın yüksek çıkması ise duran değerlere aşırı bir yatırım yapılmadığının ve kaynakların verimli kullanılmadığının göstergesi olarak kabul edilmektedir.²⁴²

Öz Sermaye Devir Hızı:

Öz sermayenin verimliliğini ölçen oran

$$\text{Öz Sermaye Devir Hızı} = \frac{\text{Satış Gelirleri}}{\text{Öz Sermaye}}$$

olarak hesaplanmaktadır. Oranın yüksek çıkması, öz sermayenin verimli olarak kullanıldığı, düşük çıkması da atıl öz sermaye bulunduğu anlamına gelmektedir. Oranın normalin çok üstünde çıkması ise net satışların, öz sermaye yanında yabancı kaynak kullanımı ile sağlandığının bir göstergesi olarak kabul edilmektedir. Oranın belli bir standardı bulunmamaktadır.²⁴³

Toplam Varlık Devir Hızı:

$$\text{Toplam Varlık Devir Hızı} = \frac{\text{Satış Gelirleri}}{\text{Toplam Varlıklar}}$$

Bu oran bir işletmede teknoloji kullanımının bir göstergesi, varlık kullanımında bir etkinlik ölçüsü ya da işletmenin risk derecesinin göstergesi olarak yorumlanabilmektedir. Varlık devir hızının yüksek olduğu endüstri kollarında risk göreceli olarak daha düşük olmaktadır.²⁴⁴ Bunun yanı sıra düşük oran, işletmenin aktiflerinin atıl kapasiteyle çalıştığının bir göstergesi olarak kabul

²⁴² N. Kurnaz, (Çevirimiçi) www.niyazikurnaz.net, 12 Temmuz 2011.

²⁴³ M. Tükenmez, vd., **a.g.e.**, s. 413.

²⁴⁴ Ö. Akgüç, **Finansal Yönetim**, ss.57–58.

edilmektedir. Oranın düşük olması kredi verenler açısından işletme varlıkları içinde duran varlıkların büyük bir ağırlığa sahip olduğu anlamını taşımaktadır. Bununla birlikte duran varlıklarda sürekli gelir elde edilmesi oranın yükselmesine neden olacaktır.²⁴⁵

Varlık devir hızı bir işletmenin karlılığını belirleyen önemli etmenlerden biridir. Bir işletme diğer koşullar aynı kalmak üzere varlık devir hızını arttırmak yoluyla öz sermaye karlılığını yükseltebilmektedir.²⁴⁶

3.2.2 Finansal Olmayan Değişkenler

Sermaye Yoğunluğu:

Sermaye yoğunluğu, emek üretkenliğini belirleyen önemli bir göstergedir. Sermaye yoğunluğu değerlendirilirken en çok kullanılan oranlar çalışan sayısına göre duran varlık, toplam varlık ve makine teçhizatla elde edilen değerlerdir.²⁴⁷ İşletmelerin makine-teçhizatla ilgili verilerine ulaşımadaki zorluklar nedeniyle bu çalışmada aşağıdaki iki oran kullanılmıştır.

$$\text{Sermaye Yoğunluğu}_1 = \frac{\text{Duran Varlık}}{\text{Çalışan Sayısı}}$$

$$\text{Sermaye Yoğunluğu}_2 = \frac{\text{Toplam Varlık}}{\text{Çalışan Sayısı}}$$

İktisatta sermayenin emeğe oranı sermaye yoğunluğu olarak ifade edilmektedir. Ekonomik büyümenin temel belirleyicileri; sermaye birikimi, teknolojik gelişme ve istihdam artışıdır. İmalat sanayinde 1980’li yıllardan sonra üretimin sürecinde gerileme yaşanmıştır. Sermaye birikimi oranındaki ciddi yavaşlama bu oluşumun esas sebebinin oluşturmaktadır. 2000–2003 yıllarında sermaye stoku

²⁴⁵ Ö. Akgüç, Finansal Yönetim, s.57.

²⁴⁶ A.e., s. 58.

²⁴⁷ Z. Akal, **İmalatçı Kamu Kuruluşlarında İşletmelerarası Toplam Performans, Verimlilik, Karlılık ve Maliyet Karşılaştırması**, Ankara, MPM: 538, 1994, s. 53.

yeniden ilerleme kaydetmiş ve 1985’den bu yana düşmekte olan performans yeniden pozitif bir konuma geçmiştir. Son yıllarda yaşanan emek verimliliğinin artışında, özellikle özel sektör yatırımlarında kendini gösteren sermaye birikimi gelişiminin payı olabileceği öne sürülmektedir. Ancak, bunun ne ölçüde “sürdürülebilir” olduğu tartışılmaktadır. Bu noktada büyüme stratejilerinin özellikle teknoloji ve verimlilik bağlantılı olması önerilmektedir.²⁴⁸

Yabancı Sermaye Oranı:

İşletmelerin öz sermayesinin yabancı oranını gösteren bu değişken işletmelere doğrudan yapılmış yabancı firmaların ortaklığı ile ilgili bilgi vermektedir.

Hızlı bir büyüme politikası benimseyen Türkiye’nin tasarruf, dış ticaret açığı gibi kalkınmasını engelleyen problemlerden kendisini koruyabilmesi için yabancı sermaye politikası önemli bir ekonomik stratejidir.²⁴⁹

İşletmelerin büyümek ve yeni teknolojileri en hızlı şekilde takip edebilmeleri için ihtiyaç duydukları gerekli sermaye nedeniyle yabancı sermaye yatırımlarının önemi artmıştır. Gelişen ekonomisiyle Türkiye, 2002 yılından itibaren uluslararası sermaye ve yatırımcıların ilgi odağı olmaktadır.

Tablo 2.2 incelendiğinde 2007 yılında toplam 22 milyar dolar değerinde uluslararası sermayenin ülkeye girdiği görülmektedir. Küresel ekonomik krizin etkili olduğu 2009 yılında ise ülkeye giren uluslararası sermaye çok ciddi anlamda düşüş göstermiştir. Türkiye’nin iç kaynaklarının yetersiz ve dış borçlanma maliyetlerinin çok yüksek olması sebebi ile doğrudan yabancı yatırımlar teşvik edilmektedir. Doğrudan yabancı yatırımların getireceği

²⁴⁸ H. Suiçmez, Türkiye’de Ekonomik Büyüme ve Verimlilik Artış Performansı İşığında Nasıl bir Kalkınma Politikası Benimsenmeli, **Türkiye İşveren Sendikaları Konfederasyonu**, Ocak, 2007. (Çevirimiçi) http://www.tisk.org.tr/isveren_sayfa.asp?yazi_id=1618&id=84, 11 Haziran 2011

²⁴⁹ Açık Gazete, Yabancı Sermaye Nedir? Hazırlayan Faruk Eskioglu, (Çevirimiçi) <http://www.acikgazete.com/ozel-dosyalar/2009/03/24/yabanci-sermaye-nedir.htm?aid=27995>, 01 Temmuz 2011

teknoloji, know-how, modern işletme ve pazarlama yöntemleri işletmelerin büyüme stratejilerinde önemli hale gelmektedir.

Tablo 2.2 Yabancı Sermaye ve Ekonomik Büyüme

Yıllar	Doğrudan Yabancı Sermaye (Fiili giriş) (milyon \$)	Ekonomik Büyüme (%)
2002	1042	6.2
2003	1693	5.3
2004	2779	9.4
2005	10010	8.4
2006	20223	6.9
2007	22023	4.7
2008	18148	0.7
2009	7610	-4.7

Kaynak. Hazine Müsteşarlığı

Yenilik:

Araştırma Geliştirme (Ar-Ge) harcamalarının işletmelerin satış gelirleri içindeki payını gösteren yenilikle ilgili oran işletmelerin yenilik düzeyinin göstergesidir.²⁵⁰

$$Yenilik = \frac{Ar - Ge \text{ Harcamaları}}{Satış Gelirleri}$$

Son yıllarda işletmeler rekabet avantajı sağlayabilmek için ürün maliyetlerini düşürmeye çalışmaktadırlar. Bu durum işletmelerin bilgi ve iletişim teknolojilerini kullanmalarını ve takip etmelerini zorunlu kılmaktadır. Uzun dönemde sürekli büyüme sağlayabilmek için işletmeler hatta ülkeler, teknolojik yenilikleri takip ederek verimlilik düzeylerini arttırmak zorundadırlar.

²⁵⁰ Z. Griliches, "Productivity, R&D and Basic Research at the Firm Level in the 1970's", *The American Economic Review*, 1986, pp. 141–153.

Ar-Ge destek programlarının uygulanmasının en önemli nedenlerinden biri, teknolojik yenilik sonucu üretkenliğin ve rekabet gücünün artırılmasıdır. Yapılan bir anket çalışması sonucunda Türkiye’deki işletmelerin sadece yaklaşık %23’ü Ar-Ge yatırımı yapmaktadır.²⁵¹

Üretilen yeni teknolojilerle yeni ürünler ve yeni üretim yöntemleri oluşturulmaktadır. Ürün yeniliği, yeni ürünler için piyasalar geliştirmek veya iyileştirilmiş ürünlere olan talebi artırarak üretim düzeyinde bir artışa yol açabilir. Bu nedenle işletme ve sektör düzeyinde, ürün yeniliklerinin işgücü talebini ve istihdamı artırma eğiliminde olması, buna paralel olarak da kişilerin ve ülkenin refah düzeyinin sürekli artış göstermesi beklenmektedir.²⁵² Bu nedenle uzun dönemde büyüme ve refah için anahtar kavram teknolojik değişim ve yeniliklerdir.²⁵³

İhracat:

İhracat, işletmelerin yurt dışına yapmış oldukları satışları ifade etmektedir. İhracatın işletmelere sağladığı faydalar; alternatif pazar yaratarak pazar payının büyümesine olanak sağlaması ve böylece pazar payının büyümesi ile kapasite kullanım oranını arttırabilmesi, işletmenin iç pazarda rekabet gücünü arttırması ve son olarak da işletmenin iç pazarda oluşabilecek ekonomik durgunluklardan daha az etkilenmesini sağlaması olarak sıralanabilir. İşletmeler ihracat yaparak bir takım riskler de taşımaktadırlar bu risklerden bazıları; işletmenin ihracat yaptığı ülkenin yaşadığı istirsazlıkların ihracat gelirlerinde problemlere neden olabilmesi, döviz kuru dalgalanmalarının olumsuz etkisi ve herhangi bir sorun

²⁵¹ “Yatırım Ortamı Değerlendirmesi Krizden Özel Sektör Öncülüğünde Büyümeye”. (Çevirimiçi) <http://www.hazine.gov.tr>, 07 Temmuz 2011.

²⁵² Bölüm-10, “Teknolojik Yenilik ve Ekonomik Performans”, (Çevirimiçi) <http://www.inovasyon.org/pdf/blm10.pdf>,_01 Temmuz 2011.

²⁵³ H. Suiçmez, **a.g.e.**

karşısında ticari ve hukuki anlaşmazlıkların çözümünün zaman kaybına neden olması şeklinde ifade edilebilir.²⁵⁴

“1966’da Vernon tarafından ileri sürülen büyüme dayalı ihracat hipotezine göre, ülkelerin sahip oldukları büyüme oranlarının, ihracatları üzerinde pozitif yönlü bir etkiye yol açtığı ve dolayısıyla ülkenin ihracatında önemli ölçüde bir genişlemeye neden olduğu savunulmaktadır.”²⁵⁵

Ekonomik büyüme ve ihracat arasındaki bu karşılıklı ilişki nedeniyle Türkiye’de özellikle büyüme önem veren işletmeler ihracata yönelmişlerdir.

Büyüyen işletmelerin Ar-Ge harcamalarını özellikle son yıllarda arttırarak daha yeni teknolojiye dayalı ve kaliteli ürünler ürettikleri de görülmektedir.²⁵⁶

Pazar Payı:

Pazar payı, sektördeki toplam satışlar içerisindeki işletme satışlarının payını ifade etmektedir. Bu gösterge işletmenin faaliyet sektörü içindeki yerini ve rekabet durumunu belirlemede, başka bir ifadeyle işletme faaliyetinin başarısını ortaya çıkartmaktadır. Ayrıca bu oran işletmenin ürün çeşitliğinin rakibine kıyasla genişliğini ölçmektedir.²⁵⁷

$$\text{Pazar payı} = \frac{\text{İşletme Satışları}}{\text{Sektördeki Satışlar}}$$

PIMS (Profit Impact of Market Strategies_Pazarlama Stratejilerinin Karlılığa Etkisi) proje araştırması* sonuçları aktif karlılığa etki eden 37 faktörü analiz

²⁵⁴ Bölüm-1, “Niçin İhracat”, **100 Soruda Dış Ticaret, İhracat Geliştirme Etüd Merkezi**, Ekim, 2006, (Çevirimiçi)

http://www.dkib.org.tr/dosya/pratik_100soru.pdf, 10 Temmuz 2011.

²⁵⁵ C. Bilgin, A. Şahbaz, **a.g.e.**, ss. 177–198.

²⁵⁶ “Yatırım Ortamı Değerlendirmesi Krizden Özel Sektör Öncülüğünde Büyüme”, **a.g.e.**, 07 Temmuz 2011.

²⁵⁷ Ö. Dinçer, **Stratejik Yönetim ve İşletme Politikası**, İstanbul, 1992, s. 116.

* Stratejik planlama ile işletmenin karlılığı arasındaki ilişkiyi araştırmak üzere 620 değişik konuda faaliyet gösteren 57 ana işletmede yapılan bir araştırmadır. Projenin

etmiştir. Bu faktörler içerisinde yer alan pazar payına ilişkin PIMS'in elde ettiği sonuçlar kısaca şu şekilde özetlenebilir;²⁵⁸

- Pazar payının karlılık üzerinde gerçekten büyük etkisi vardır. Pazar payı artarken karlılık süratle yükselmektedir.
- Büyük pazar payına sahip işletmelerin ortalama sermaye devir hızları yüksek ve pazarlama masraflarının satışlara oranı düşüktür. Bu nedenle pazarlama masraflarının satılan ürün birimi başına payı düşmekte ve bu durum karlılığı arttırmaktadır.
- Yüksek pazar payına sahip işletmelerin ürünlerinin kalitesi de yüksektir.
- Pazar payı zayıf olan işletmelerde Ar-Ge masraflarının satışlara oranı arttığı takdirde karlılık %50'den fazla azalmaktadır. Pazar payı yüksek olan işletmelerde ise karlılık oranı %18 artmaktadır.

Büyüklik:

İşletmenin büyüklüğünü belirlemede kantitatif ve kalitatif olmak üzere çeşitli kriterlerden yararlanılmaktadır. Kantitatif büyüklük ölçüleri olarak; yıllık satışlar, yıllık karlar, varlıklar, öz sermaye, yatırım toplamı, yaratılan katma değer, çalışanların sayısı, harcanan enerji miktarı ve Ar-Ge için harcanan para miktarı kullanılabilir. Kalitatif büyüklük ölçüleri ise; sermayedarların sayısı, yönetim biçim, bölgeye yönelik olup- olmama, endüstri dalındaki nisbi durum ve hukuki şekil olarak sıralanabilir.²⁵⁹

araştırma sorumluluğunu Harvard İşletme Fakültesi Profesörlerinden Robert D. Buzzel yapmıştır.

²⁵⁸ (Çevirimiçi) <http://notoku.com/pims-analizi/> , 12 Temmuz 2011.

²⁵⁹ İ. Mucuk, **Modern İşletmecilik**, 7. Baskı, Türkmen Kitapevi, İstanbul, 1997, ss. 98–99.

DÖRDÜNCÜ BÖLÜM

İŞLETME PERFORMANSININ TAHMİNİNE YÖNELİK İMALAT SANAYİ SEKTÖRÜ UYGULAMASI

4.1 Araştırmanın Amacı ve Konusu

İşletmelerin performanslarının sürekliliğinin sağlanmasının önemi özellikle yatırımcılar, ortaklar ve işletme yöneticileri açısından her geçen gün artmaktadır. İşletmelerin performanslarını birden çok ve farklı kriter ile ölçmeleri ve değerlendirmeleri mümkündür. Hem işletmelerin yatırımcıları ve ortaklarının kararlarında, hem de ülke ekonomilerinin verimliliğinin ve etkinliğinin belirlenmesinde işletmelerin performanslarının değerlendirilmesi için işletme performanslarını etkileyen faktörlerin en iyi şekilde belirlenmesinin önemi günümüze kadar birçok araştırmacının dikkatini çekmektedir.

Bu çalışmada İstanbul Menkul Kıymetler Borsası (İMKB) imalat ve sanayi endeksinde bulunan işletmelerin, finansal ve finansal olmayan açıklayıcı değişkenlerinin işletmelerin geleneksel performans göstergeleri (karlılık, verimlilik, büyüme ve piyasa) üzerinde etkisinin araştırılması hedeflenmiştir. Günümüze kadar yapılan çalışmalar incelendiğinde performans göstergelerinin tek tek ele alındığı ve performansı etkileyen faktörlerin kurulan modeller için ayrı ayrı belirlendiği gözlenmiştir. Bu çalışmada ise farklı olarak, performans göstergelerinin tüm mümkün kombinasyonları denenmiş ve anlamlı olan modeller araştırmanın kapsamında incelenmiştir. İşletme performans göstergelerinin birarada ele alınması hedeflendiği için kategorik olarak tanımlanması gerekmektedir.

4.2 Araştırmanın Veri Seti

Araştırmada kullanılan veri seti 2007 yılında İMKB’de İmalat Sanayi Sektöründe faaliyet gösteren işletmelerden oluşmaktadır. 2008 yılında gerçekleşen ve Türkiye’yi de etkileyen global kriz nedeniyle işletmelerin finansal ve finansal olmayan göstergelerden özellikle “Stok Devir Hızı” ve “Alacak Devir Hızı” ile Tablo 2.2’deki “Yabancı Sermaye ve Ekonomik Büyüme” değerleri 2008 ve 2009 yıllarında, yaşanan krizin etkilerini taşımaktadır. Ayrıca çalışmada kullanılacak bazı finansal olmayan oranlar için İstanbul Sanayi Odası’nın (İSO) her sene Eylül ayında yayımladığı özel sayıdan yararlanılması nedeniyle 2010 yılı verileri incelenememiştir. Bu nedenlerden dolayı çalışmada, işletme performans kriterlerini etkileyebilecek uygun faktörlerin saptanması amacıyla ekonomik istikrarın olduğu 2007 yılı verilerinin kullanılması uygun görülmüştür.

İşletmelere ait finansal değişkenlerin ve piyasa performans oranlarının hesaplanmasında İMKB’nin internet üzerinden yayınladığı yıllık raporlar, mali tablolar ve bu tablolara ait dipnotlardan yararlanılmıştır. Ayrıca finansal olmayan oranların belirlenmesinde İSO’nun 2007 yılında yayımladığı özel sayıdan elde edilen değerler kullanılmıştır.

Çalışmada performans göstergelerinin olası tüm kombinasyonları için incelenen işletme sayıları farklılık göstermektedir. İşletmelerin sayısına ilişkin bilgiler “Verilerin Analize Hazırlanması” başlığında (4.4) ayrıntılı olarak belirtilmiştir. Araştırma verilerinin analizinde SPSS’in son sürümlerinden olan PASW Statistics 18 ve STATA 9.2 paket programları kullanılmıştır.

4.3 Araştırma Değişkenlerinin Seçimi

Çalışmanın amacı olan işletme performansı (bağımlı değişken) belirlenirken karlılık, büyüme, piyasa ve verimlilik, verilerinin her biri birden çok değişkenle ifade edilmiştir. Bu değişkenlerden performans kriterlerini en iyi temsil edenleri saptanarak, kullanılacak oranları belirlemek amacıyla tek tek karlılık, büyüme

ve piyasa performans göstergeleri için oranlar arasındaki ilişkiyi ortaya çıkarmaya yönelik bir ön analiz yapılmıştır. Bu amaçla önce “Veri setindeki değişkenlerin sıraları arasında uygunluk yoktur” sıfır hipotezi ile kurulan Kendall Uygunluk Katsayısı Testi kullanılmış ve test değerlerine göre her üç performans göstergesi değişkenlerinin de sıraları arasında uygunluk olduğu görülmüştür (H_0 red). Sıfır hipotezinin rededilmesinden sonra gerçek değerler yerine değişkenlerin sıralarının kullanıldığı Spearman Korelasyon değerleri ile değişkenler arasındaki ilişkinin gücü ortaya konmuştur.

Karlılık, büyüme ve piyasa kriterleri için uygulanan Kendall Uygunluk Katsayısı Testi ve Spearman Sıra Korelasyon ön analiz sonuçları sırası ile EK.1, EK.2 ve EK 3’de verilmiştir. Ön analiz sonuçlarına göre karlılık kriteri açısından Başlık 3.1.1’de tanımlanan tüm karlılık oranlarının birbirleriyle olan ilişkisi yüksek çıkmıştır. Ancak yapılan çalışmalarda çoğunlukla tercih edilmesi sebebiyle bu çalışmada da “Aktif Karlılık” oranı seçilmiştir.

Büyüme kriteri açısından, Başlık 3.1.3’te değinilen büyüme oranlarının birbirleriyle ilişkilerinin çok güçlü olmadığı görülmüştür. Bu nedenle, kullanılacak değişkenin belirlenmesi için işletmelerin sınıflamada diğer performans kriterleri ile birlikte değişkenlerin hepsi tek tek gruplanmış ve işletme sayısı baz alındığında en iyi sınıflandırmayı veren büyüme performans göstergesinin “Satışlardaki Reel Artış” ve “Varlık Artış” oranları olduğuna karar verilmiştir.

Piyasa performans kriteri açısından Başlık 3.1.4’te tanımlanan değişkenlerin birbirleri ile olan ilişkilerinin yüksek olmadığı görülmüştür. Bu durumda, büyüme kriterinde olduğu gibi tüm mümkün piyasa performans göstergeleri için karşılaştırma yapılmış ve en uygun gruplandırma sonucuna göre bağımsız değişkenin belirlenmesinde “Fiyat-Kazanç” oranının kullanılmasına karar verilmiştir. Yatırımcıların şirket analizinde performans göstergesi olarak en çok başvurduğu değişkenin Fiyat-Kazanç oranı olması da bu değişkenin tercih edilmesinde etkili olmuştur.

Bir diğer performans kriter değeri olan verimlilik için ise; karşılaştırılabilirlik, yorumlanabilirlik ve güvenilirlik hususlarında en az sorun işgücü verilerinde olduğundan, özellikle imalat sanayi sektöründe en yaygın kullanılan göstergenin literatürde “Emek Verimliliği” olduğu gözlemlenmiştir. Bu nedenle çalışmada da emek verimliliği oranı seçilmiştir.

Çalışmada kullanılan bağımsız değişkenler ise finansal oranlar (Tablo 4.1) ve finansal olmayan oranlar (Tablo 4.2.) olmak üzere grupta toplanmıştır.

Tablo 4.1 Bağımsız Değişken Olarak Kullanılan Finansal Oranlar

Oran	Formüller	Modeldeki Karşılığı
Likidite Oranları		
Cari Oran	Dönen Varlık / Kısa vadeli Yükümlülükler	Likidite1
Asit-Test Oranı	(Dönen Varlık-Stoklar) / Kısa vadeli Yükümlülükler	Likidite2
Nakit Oranı	(Nakit ve Nakit Benzerleri+Finansal Yatırımlar) / Kısa Vadeli Yükümlülükler	Likidite3
Stokların Toplam Varlıklara oranı	Stoklar / Toplam Varlıklar	Likidite4
Kısa vadeli Ticari Alacakların Toplam Varlıklara Oranı	Kısa vadeli Ticari Alacaklar / Toplam Varlıklar	Likidite5
Finansal Yapı Oranları		
Finansal Kaldıraç Oranı	(Kısa Vadeli Yük.+ Uzun Vadeli Yük.) / Toplam Varlıklar	FinYap1
Kısa Vadeli Yabancı Kaynakların Toplam Varlıklara Oranı	Kısa Vadeli Yük. / Toplam Varlıklar	FinYap2

Uzun Vadeli Yabancı Kaynakların Devamlı Sermayeye Oranı	$\frac{\text{Uzun Vadeli Yük.}}{\text{Yük.} + \text{Öz Sermaye}}$ (Uzun Vadeli)	FinYap3
Duran Varlıkların Öz Sermayeye Oranı	$\frac{\text{Duran Varlıklar}}{\text{Öz Sermaye}}$	FinYap4
Dönen Varlıkların Toplam Varlıklara Oranı	$\frac{\text{Dönen Varlıklar}}{\text{Toplam Varlıklar}}$	FinYap5
<i>Faaliyet Oranları</i>		
Hazır Değerler Devir Hızı	$\frac{\text{Satış Gelirleri}}{\text{Hazır Değerler}}$	Faaliyet1
Stok Devir Hızı	$\frac{\text{Satışların Maliyeti}}{\text{Ortalama Stok}}$	Faaliyet2
Alacak Devir Hızı	$\frac{\text{Satışların Gelirleri}}{\text{Alacaklar}}$ Ortalama Ticari	Faaliyet3
Dönen Varlık Devir Hızı	$\frac{\text{Satış Gelirleri}}{\text{Dönen Varlıklar}}$	Faaliyet4
Duran Varlık Devir Hızı	$\frac{\text{Satış Gelirleri}}{\text{Duran Varlıklar}}$	Faaliyet5
Öz Sermaye Devir Hızı	$\frac{\text{Satış Gelirleri}}{\text{Öz Sermaye}}$	Faaliyet6
Toplam Varlık Devir Hızı	$\frac{\text{Satış Gelirleri}}{\text{Toplam Varlıklar}}$	Faaliyet7

Kaynak: İMKB

Tablo 4.2 Bağımsız Değişken Olarak Kullanılan Finansal Olmayan Oranlar

Oran	Formülleri	Modeldeki Karşılıkları
Sermaye Yoğunluğu-1	Duran Varlık / Çalışan Sayısı	SermayeYog1
Sermaye Yoğunluğu-2	Toplam Varlık / Çalışan Sayısı	SermayeYog2
Yenilik	Ar-Ge Harcamaları / Satış Gelirleri	Yenilik
Pazar Payı	İşletme Satışları / Sektördeki Satışlar	PazarPay
Yabancı Sermaye Oranı*		YabSermaye
İhracat*		İhracat
Büyükklük*		Buyukluk

Kaynak: İMKB, İSO

4.4 Verilerin Analize Hazırlanması

Araştırmanın veri setinde yer alan işletmelerin her biri belirlenen dört performans kriterlerinin olası tüm dördü, üçlü ve ikili kombinasyonları için küçükten büyüğe doğru sıralanmıştır. Bu işlem sonucunda tüm kriterlere göre küçük değer alan işletmeleri “0” büyük değer alan işletmelere ise “1” grubu atayarak iki ayrı sınıf oluşturulmuştur. Kategorize edilemeyen işletmeler ise analiz dışında bırakılmıştır. Bu şekilde oluşturulan kategorik bağımlı değişkeni açıklayabilecek faktörleri saptamak üzere 24 bağımsız değişken, bahsedilen tüm kombinasyonların modellenmesinde kullanılmıştır. Dördü ve üçlü performans göstergesi kombinasyonlarından Karlılık-Verimlilik-Büyüme kombinasyonu dışındaki göstergeler için işletmeler sınıflandırılmamıştır. Olası tüm ikili performanslar incelendiğinde ise Verimlilik-Piyasa ve Büyüme-Piyasa

* Finansal olmayan oranlardan yabancı sermaye oranı ile ihracat değişkeni ISO dergisinden doğrudan alınmıştır. Büyükklük değişkeni ise yine aynı dergiden alınan şirketlerin katma değerlerine göre yüksek olanlara “1” düşük olanlara “0” değeri atanarak kategorik veriye dönüştürülmüştür.

göstergelelerini açıklayabilecek bağımsız değişkenler elde edilemediği için modelleri kurulamamıştır. Sorun yaşanan kombinasyonlar çıkarıldığında kalan beş kombinasyonla elde edilen modeller çalışmada ayrıntılı olarak incelenmiştir. Oluşturulan modeller ile her bir model için işletme sayıları şu şekildedir:

Model 1: Karlılık-Verimlilik-Büyüme göstergelerine göre (79 işletme için),

Model 2: Karlılık-Verimlilik göstergelerine göre (104 işletme için),

Model 3: Karlılık-Büyüme göstergelerine göre (95 işletme için),

Model 4: Karlılık-Piyasa göstergelerine göre (78 işletme için),

Model 5: Verimlilik-Büyüme göstergelerine göre (88 işletme için) performans göstergelerini etkileyen faktörler belirlenmiştir.

Kurulan modellerde karlılık için aktif karlılığı, verimlilik için emek verimliliği, piyasa için de fiyat-kazanç oranı kullanılmıştır. Büyüme değişkeni olarak Model 1’de varlık artış oranı, Model 3 ve Model 5’te ise satışlardaki reel artış oranı benimsenmiştir.

4.5 Finansal Performans Tahmin Modellerinin Uygulanması

Her bir modelin açıklayıcı değişkenlerinin saptanması için hem Diskriminant Analizi hem de Lojistik Regresyon Analizi uygulanmıştır. Diskriminant Analizi için PASW Statistics 18 paket programı kullanılmıştır. Lojistik Regresyon Analizi için ise “Kesirli Polinomlar”, “Roc Eğrisi” ve modelin oluşturulmasında kullanılan “En İyi Model Seçim” sonuçlarının elde edilebilmesi amacıyla PASW Statistics 18 ile birlikte Stata 9.2 istatistik paket programı da kullanılmıştır.

4.5.1 Model 1: Performans Göstergesi Olarak Karlılık Verimlilik Büyüme

Bu modelde performans göstergesi olarak karlılık, verimlilik ve büyüme için seçilen “aktif karlılığı”, “emek verimliliği” ile “varlık artış” değerlerini birlikte açıklayan faktörlerin saptanması amaçlanmıştır.

Model 1 oluşturulurken, bağımsız değişkenler arasındaki yüksek korelasyon sebebiyle ortaya çıkabilecek çoklu doğrusal bağlantı sorunun önüne geçebilmek amacıyla Adım Adım (Stepwise) Regresyon Yöntemi hem Diskriminant hem de Lojistik Regresyon Analizi’nde uygulanmıştır.

Diskriminant Analizi için çoklu doğrusal bağlantı varsayımı haricinde eşit varyanslılık ve çoklu normallik varsayımlarının testleri yapılarak modellerin grupları tahmin etmedeki başarısı değerlendirilmiştir.

Lojistik Regresyon Analizi için ise Adım Adım Regresyon Yöntemi uygulandıktan sonra bulunan modelden elde edilen lojit değerleri ile modeldeki bağımsız değişkenlerin doğrusallıkları, Kesirli Polinomlar (Fractional Polynomial) yöntemi ile kontrol edildikten sonra, bulunan model ve parametrelerin uygunluk test sonuçları incelenmiş ve klasik yöntem ve ROC eğrisi çizilerek modelin sınıfları tahmin etmedeki başarısı değerlendirilmiştir.

Diskriminant Analizi Sonuçları

Model 1 için öncelikle analizin sağlaması gereken eşit varyanslılık, çoklu normallik ve bağımsız değişkenler arasında ilişki olmaması varsayımlarının sağlanıp sağlanmadığının kontrolleri yapılmıştır.

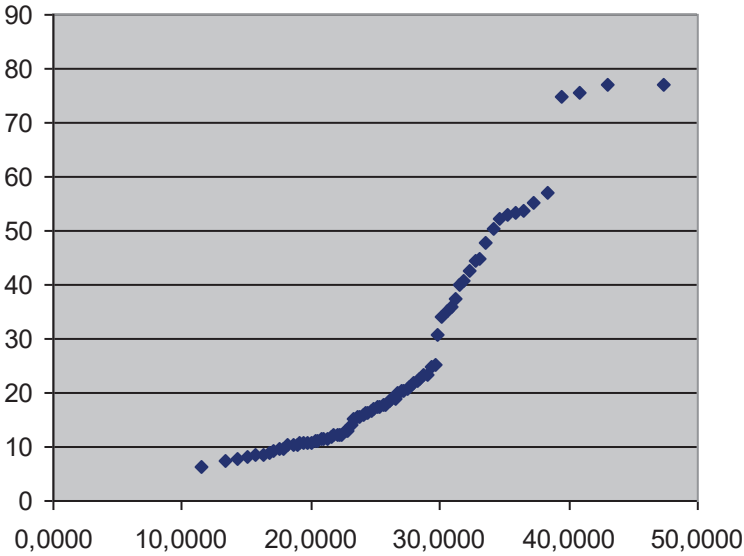
Diskriminant Analizi ile yapılan testler ve sonuçlarının yorumlanmasında kullanılan değerler Tablo 4.3’te özetlenmiştir (Paket program sonuçları için bkz. EK. 4).

Eşit varyanslılık için iki grubun varyans-kovaryans matrislerinin eşit olduğunu ileri süren sıfır hipotezinin Box's M test sonucuna göre reddedildiği görülmektedir (sig. 0,000). Bu durumda eşit varyanslılık varsayımı sağlanmamaktadır.

Tablo 4.3 Model 1 için Diskriminant Analizi Testleri ve Sonuçları

Box's M		Eigenvalues			Wilks' Lambda	
Değeri	Sig.	Değeri	Varyans (%)	Kanonik Korelasyon	Değeri	Sig.
1835,091	0,000	0,874	100	0,683	0,534	0,000

Veri setinden hesaplanan Mahalanobis uzaklığının sıralı değerleri ile bunların χ^2 değerlerinden elde edilen serpilme diyagramı çizildiğinde çoklu normallik varsayımının da sağlanmadığı görülmektedir.



Şekil 4.1 Model 1 için Serpilme Diyagramı

Çoklu doğrusal bağlantı sorunu, analiz uygulanırken Adım Adım Regresyon Yöntemi kullanılarak aşılacak istenmiş ve Tablo 0.4'ten standartlaştırılmış kanonik diskriminant fonksiyonu şu şekilde elde edilmiştir.

$$T = 0,548 \text{ Likidite5} + 0,634 \text{ FinYap2} + 1,768 \text{ SermayeYog1} \\ - 2,000 \text{ SermayeYog2} - 0,376 \text{ YabSermaye} \\ - 0,530 \text{ Buyukluk}$$

Diskriminant Modeli'ne giren değişkenler incelendiğinde sermaye yoğunluğu için hesaplanan iki oranın (SermayeYog1 ve SermayeYog2) birbiriyle yüksek korelasyona sahip olması nedeniyle (bkz. EK. 4) Adım Adım Regresyon Yönteminin çoklu doğrusal bağlantı problemini çözmede etkin olamadığı görülmüştür.

Tablo 4.4 Model 1 için Diskriminant Modeli ve Doğru Sınıflandırma Sonuçları

Standartlaştırılmamış ve Standartlaştırılmış Kanonik Diskriminant Fonksiyon Katsayıları			Sınıflandırma Sonuçları			
Değişkenler	β	β^*	Sınıf	0	1	Doğru Yüzdesi
Likidite5	4,206	0,548	0	23	9	71,9
FinYap2	3,786	0,634	1	6	41	87,2
SermayeYog1	0,000	1,768	Toplam			81,0
SermayeYog2	0,000	-2,000				
YabSermaye	-0,015	-0,376				
Buyukluk	-1,444	-0,530				
(Sabit)	-0,871					

Tablo 4.4 ile verilen Diskriminant modeli sonuçları, bağımsız değişkenlerin sınıfları ayırmadaki önemini ortaya koyan “Grup Ortalamalarının Eşitliği Testi” ile de modele giren değişkenlerin grup ortalamalarının farklı olması nedeniyle desteklenmektedir (bkz. EK. 4).

Diskriminant Analizi sonuçlarına göre model, bağımlı değişkendeki varyansın yaklaşık olarak 0,466'sını ($0,683^2$) açıklayabilmektedir. Wilks' Lamda değeri ile de diskriminant skorlarındaki toplam varyansın gruplar arasındaki farklar tarafından açıklanamayan kısmının 0,534 olduğu görülmektedir (Tablo 4.3). Bu değer ayrıca, fonksiyonun iyi bir ayırım yapıp yapmadığının göstergesi olan özdeğer (eigenvalue) istatistiğinin anlamlılığının testinde kullanılmaktadır. Wilks' Lamda değeri Tablo 4.3'de görüldüğü üzere anlamlıdır (sig. 0.000).

Model 1 için yapılan analiz sonucunda Diskriminant Fonksiyonu'nun grupları doğru tahmin etme oranı %81 olarak bulunmuş olmasına rağmen, varsayımlarının sağlanmaması nedeniyle işletme performans kriterlerini etkileyen faktörlerin saptanması için daha az varsayımın sağlanmasının yeterli olduğu Lojistik Regresyon Analizi'nin uygulanmasına karar verilmiştir.

Lojistik Regresyon Analizi Sonuçları

Modelin ve parametrelerin uygunluk testine geçmeden önce, modelde bulunan bağımsız değişkenlerin modelin lojit değerleriyle doğrusallık kontrolünü yapmak amacıyla Kesikli Polinomlar Yöntemi uygulanmıştır (STATA paket programında kullanılan kodlar ve sonuçlar için bkz. EK. 5). Bu yöntemin sonuçlarına, göre değişkenlerin modelin lojit değerleri ile doğrusal olduğu ve modele değişkenlerin kendi değerlerinin girmesinin 0,01 anlamlılık düzeyinde uygun olduğu görülmüştür.

Modelin uygunluğunun sınanması* için uygulanan Hosmer-Lemeshow Testi sonucuna göre 0,961 anlamlılık düzeyinde sıfır hipotezi ret edilememektedir. Başka bir ifade ile Model 1, veri setini iyi temsil etmektedir.

* H_0 : Teorik model veriyi iyi temsil etmektedir.

H_1 : Teorik model veriyi iyi temsil etmemektedir.

Modeldeki parametrelerin uygunluğunun sınanması için ise parametrelerin birlikte anlamlılığını test eden Omnibus Testi uygulanmıştır. Bu testin sonucuna göre modeldeki parametrelerin anlamlı olduğu sonucuna varılmıştır (sig. 0,000).

Tablo 4.5 Model 1 için Test ve Doğru Sınıflandırma Sonuçları

Test Sonuçları				Sınıflandırma Sonuçları			
				Sınıf	0	1	Doğru Yüzdesi
Testler	χ^2	<i>df</i>	<i>Sig.</i>	0	25	7	78,1
Hosmer-Lemeshow Testi	2,523	8	0,961	1	5	42	89,4
Omnibus Testi	56,842	6	0,000	Toplam			84,8

$exp(\beta)$ değerlerinin yorumlanmasında sağladığı kolaylık nedeniyle standartlaştırılmış bağımsız değişken değerleriyle model tekrar kurulmuş ve parametreler %0,1 ile %1,5 düzeyinde anlamlı bulunmuştur.

Tablo 4.6 Model 1'deki Açıklayıcı Değişkenler ve Katsayıları

Modele Giren Değişkenler				Standartlaştırılmış Değişkenler için Değerler	
Değişkenler	β	<i>Wald</i>	<i>Sig.</i>	β^*	$exp(\beta^*)$
FinYap2	-9,933	8,914	0,003	-1,832	0,160
SermayeYog2	0,000	11,026	0,001	2,697	14,841
Buyukluk	3,737	7,230	0,007	1,436	4,205
YabSermaye	0,042	6,612	0,010	1,078	2,938
Likidite5	-10,854	5,899	0,015	-1,464	0,231
FinYap5	8,508	6,508	0,011	1,603	4,969
(Sabit)	-1,689	7,469	0,006	1,265	3,544

Model 1'i (Karlılık-Verimlilik-Büyüme) etkileyen bağımlı değişkenlerden en önemlisinin sermaye yoğunluğu (SermayeYog2) olduğu Tablo 4.6'dan açıkça görülmektedir. Modeldeki diğer değişkenlerin sabit kaldığı varsayımıyla

sermaye yoğunluğu değişkenindeki bir birimlik değişim; işletmenin, performansı yüksek işletmeler olarak tanımlanan sınıfa (“1” sınıfı) atanması olasılığının, performansı düşük işletmeler olarak tanımlanan sınıfa (“0” sınıfı) atanması olasılığına oranının $exp(\beta)=14,841$ kat etkilendiği şeklinde yorumlanmaktadır. Kukla değişken “büyüklük” dahil, modeldeki diğer değişkenlerin katsayılarının üstel değerleri de aynı şekilde yorumlanabilmektedir.

Modelin grupları doğru sınıflandırma yüzdesi Tablo 4.5’den de görüldüğü gibi yanlış sınıflandırma maliyetlerinin birbirine eşit olduğu varsayımıyla, kesim değeri (cut value) “0,50” olarak belirlenmiştir. Analiz sonucunda düşük performanslı işletmeleri doğru sınıflandırma oranı olan duyarlılık (sensitivity) değeri %78,1, yüksek performanslı işletmeleri doğru sınıflandırma oranı olan özgünlük (specificity) değeri %89,4, toplamda grupları doğru sınıflandırma oranı olan doğruluk (accuracy) değeri ise % 84,8 olarak elde edilmiştir.

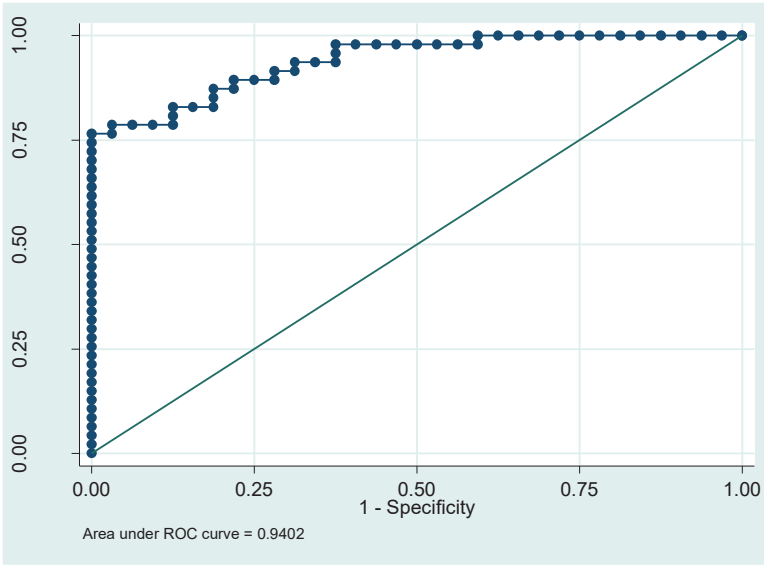
Sınıflandırma modellerinin değerlendirilmesi ve kıyaslanması için önerilen bir diğer yöntem de ROC eğrisidir (Receiver Operating Characteristic Curves).

Farklı kesim değerleri hesaplanarak oluşturulan ROC eğrisi; modelin sınıflandırma gücünün belirlenmesine, çeşitli modellerin etkinliğinin kıyaslanmasına, en uygun kesim değerinin belirlenmesine ve sonuçların kalitesinin izlenmesine olanak sağlamaktadır.²⁶⁰

ROC eğrilerinin oluşturulduğu koordinat sisteminin ordinatında, testin gerçek pozitif oranı (duyarlılık), apsisinde ise yanlış pozitif oranı (1-özgünlük) yer almaktadır. Model ne kadar iyi ise eğrinin o kadar yukarıya (yüksek duyarlılık bölgesine) ve sola (düşük yanlış pozitif oran (1-özgünlük) bölgesine) doğru kayması beklenmektedir.

²⁶⁰ A. Dirican, “Tanı Testi Performanslarının Değerlendirilmesi ve Kıyaslanması”, **Cerrahpaşa Tıp Dergisi**, Cilt 32, Sayı 1, 2001, ss. 25–30.

ROC eğrisi altında kalan alan, aynı örnek üzerinde geliştirilmiş modellerin karşılaştırılmasında önemli bir performans ölçüsü olmanın yanı sıra modelin doğru tahmin olasılığını da göstermektedir²⁶¹ ve 0,5 ile 1,0 arasında değer almaktadır. Genel kural olarak, elde edilen değer 0,5 ise ($ROC = 0,5$) model sınıfları ayırmada başarılı değildir (Para atarak gruplara karar vermekle aynı yöntemdir). $0,7 \leq ROC < 0,8$ aralığında modelin grupları kabul edilebilir düzeyde, $0,8 \leq RC < 0,9$ aralığında mükemmel, $ROC \geq 0,9$ olduğunda ise olağanüstü ayırdığı kabul edilmektedir.



Şekil 4.2 Model 1 için ROC Eğrisi

ROC eğrisi altındaki alan 0,9402 olarak bulunduğundan, modelin grupları olağanüstü ayırdığı yorumu yapılabilmektedir.

Sonuç olarak, işletmelerin Karlılık-Verimlilik-Büyüme'den oluşan performans göstergelerini etkileyen değişkenlerin saptanması amacıyla, kategorik bağımlı

²⁶¹ N. A. Obuchowski, "ROC Analysis", *American Journal of Roentgenology*, 184, 2005, pp. 364–372.

değişken analizlerinden olan Lojistik Regresyon Analizi'nin kullanılmasının uygunluğu açıkça görülmektedir.

4.5.2 Model 2: Performans Göstergesi Olarak Karlılık ve Verimlilik

İşletme performans göstergelerinden olan karlılık ve verimlilik kriterleri için seçilen “aktif karlılığı” ve “emek verimliliği” değerlerini birlikte etkileyen değişkenlerin saptanması amacıyla kurulan Model 2, sırasıyla Diskriminant ve Lojistik Regresyon Analizi kullanılmıştır.

Diskriminant Analizi Sonuçları

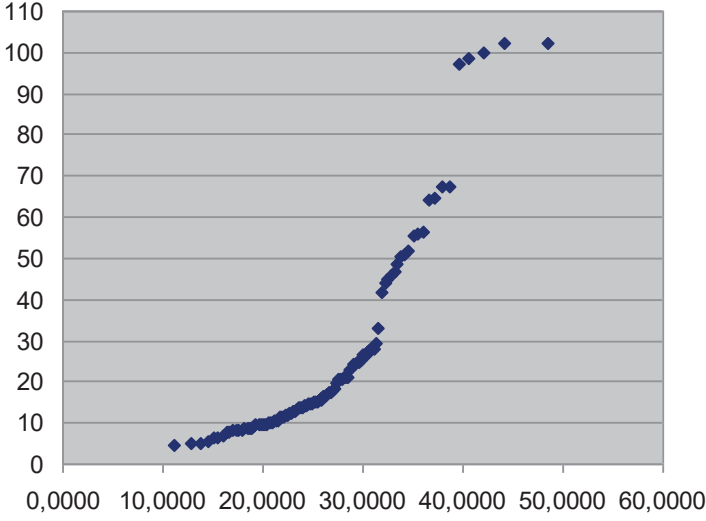
Model 2 oluşturulurken birçok bağımsız değişken arasında güçlü ilişki olduğu korelasyon matrisine bakılarak saptanmıştır. Bunun sonucunda diskriminant fonksiyonuna girecek değişkenlerinin de Adım Adım Regresyon Yöntemi ile seçilmesine karar verilmiştir.

Eşit varyanslılık varsayımı için grupların varyansları birbirine eşittir şeklindeki sıfır hipotezinin, Box's M test sonucuna göre reddedildiği görülmektedir (sig. 0,000). Diğer bir ifadeyle, eşit varyanslılık varsayımı sağlanmamaktadır.

Tablo 4.7 Model 2 için Diskriminant Analizi Testleri ve Sonuçları

Box's M		Eigenvalues			Wilks' Lambda	
Değeri	Sig.	Değeri	Varyans (%)	Kanonik Korelasyon	Değeri	Sig.
2341,607	0,000	0,777	100,00	0,661	0,563	0,000

Çoklu normallik varsayımının da sağlanmadığı Şekil 4.3'ten görülmektedir.



Şekil 4.3 Model 2 için Serpilme Diyagramı

Adım Adım Regresyon Yöntemi sonucunda elde edilen diskriminant fonksiyonu Tablo 4.8’de verilen standartlaştırılmış kanonik değerler kullanılarak aşağıdaki gibi yazılmaktadır.

$$T = -0,809 \text{ FinYap1} + 0,444 \text{ SermayeYog2} + 0,557 \text{ Buyukluk}$$

Model sonuçları, bağımsız değişkenlerin sınıfları ayırmadaki önemini ortaya koyan “Grup Ortalamalarının Eşitliği Testi” ile de, modele giren değişkenlerin grup ortalamalarının farklı olması nedeniyle desteklenmektedir.

Tablo 4.8 Model 2 için Diskriminant Modeli ve Doğru Sınıflandırma Sonuçları

Standartlaştırılmış ve Standartlaştırılmamış Kanonik Diskriminant Fonksiyon Katsayıları			Sınıflandırma Sonuçları			
Değişkenler	β	β^*	Sınıf	0	1	Doğru Yüzdesi
FinYap1	- 4,629	-0,809	0	37	10	78,7
SermayeYog2	0,000	0,444	1	9	48	84,2
Buyukluk (Sabit)	1,592 1,276	0,557	Toplam			81,7

Diskriminant Analizi sonuçlarına göre model, bağımlı değişken performansındaki varyansın yaklaşık olarak 0,437'sini ($0,661^2$) açıklayabilmektedir. Wilks' Lamda değeri ile de diskriminant skorlarındaki toplam varyansın gruplar arasındaki farklar tarafından açıklanamayan kısmının 0,563 olduğu görülmektedir (Tablo 4.7). Bu değer ayrıca, fonksiyonun iyi bir ayırım yapıp yapmadığının göstergesi olan özdeğer (eigenvalues) istatistiğinin anlamlılığında kullanılan Wilks' Lamda değeri Tablo 4.7'den görüldüğü üzere anlamlıdır (sig. 0.000).

Model 2 için yapılan analiz sonucunda diskriminant fonksiyonunun grupları doğru tahmin etme oranı % 81,7 olarak bulunmuş olmasına rağmen, çoklu normallik ve eşit varyanslılık varsayımlarının sağlanmaması nedeniyle işletme performans kriterlerini etkileyen faktörlerin saptanması için Lojistik Regresyon Analizi'nin uygulanmasına karar verilmiştir.

Lojistik Regresyon Analizi Sonuçları

Modelin ve parametrelerin uygunluk testine geçmeden, önce modelde bulunan bağımsız değişkenlerin modelin lojit değerleriyle doğrusallık kontrolünü yapmak amacıyla Kesikli Polinomlar Yöntemi uygulanmıştır. Bu yöntemin sonuçlarına göre, değişkenlerin modelin lojit değerleri ile doğrusal olduğu ve modele değişkenlerin kendi değerlerinin girmesinin 0,01 anlamlılık düzeyinde uygun olduğu görülmüştür.

Modelin uygunluğu için kullanılan Hosmer-Lemeshow testi sonucuna göre modelin verileri iyi temsil ettiğinin iddia edildiği sıfır hipotezi 0.597 anlamlılık düzeyinde kabul edilmektedir.

Modeldeki parametrelerin birlikte anlamlılığının sınanması için Omnibus Testi uygulanmış ve elde edilen sonuçlar incelendiğinde modeldeki parametrelerin anlamlı olduğuna karar verilmiştir (sig. 0,000).

Tablo 4.9 Model 2 için Test ve Doğru Sınıflandırma Sonuçları

Test Sonuçları				Sınıflandırma Sonuçları			
				Sınıf	0	1	Doğru Yüzdesi
Testler	χ^2	<i>df</i>	<i>Sig.</i>	0	39	8	83,0
Hosmer-Lemeshow Testi	6,446	8	0,597	1	6	51	89,5
Omnibus Testi	78,873	4	0,000	Toplam			86,5

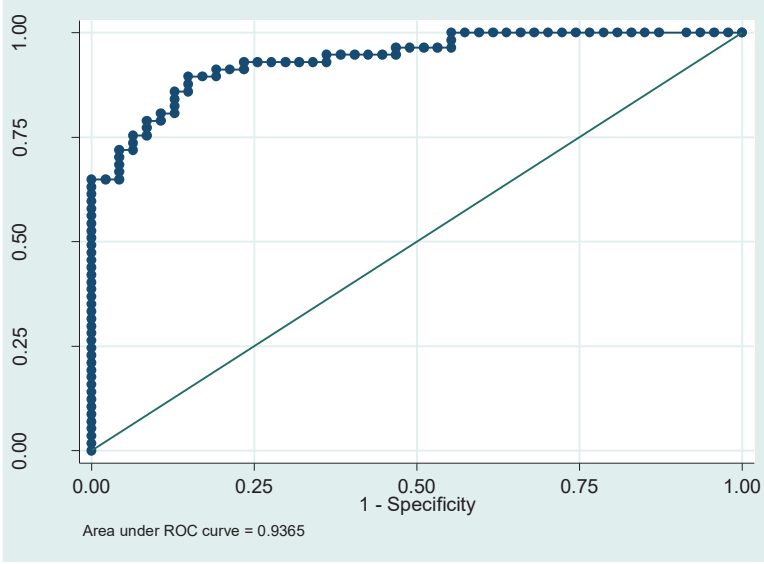
Adım Adım Regresyon Yöntemiyle kurulan Lojistik Regresyon Modeli'ndeki bağımsız değişkenler Tablo 4.10'da görülmektedir. $exp(\beta)$ değerlerinin yorumlanmasında sağladığı kolaylık nedeniyle standartlaştırılmış bağımsız değişken değerleriyle model tekrar kurulmuş ve FinYap1, SermayeYog2 ve Buyukluk değişkenlerinin anlamlı olduğuna karar verilmiştir (sig. 0,000) ayrıca YabSermaye değişkeni de %3,8 düzeyinde anlamlı bulunmuştur.

Tablo 4.10 Model 2'deki Açıklayıcı Değişkenler ve Katsayıları

Modele Giren Değişkenler				Standartlaştırılmış Değişkenler için Değerler	
Değişkenler	β	<i>Wald</i>	<i>Sig.</i>	β^*	$exp(\beta^*)$
FinYap1	-9,273	16,915	0,000	-1,874	0,154
SermayeYog2	0,000	15,117	0,000	3,585	36,043
Buyukluk	5,397	13,436	0,000	2,006	7,430
YabSermaye	0,023	4,306	0,038	0,632	1,881
(Sabit)	1,090	1,678	0,195	1,274	3,574

İşletmeler için performans göstergelerinden olan karlılık-verimlilik kriterleri birlikte ele alındığında Model 2'nin Model 1'deki değişkenlere oldukça benzer olduğu görülmektedir. İki modelin değişkenleri arasındaki fark, Model 1'de likidite oranlarından beşincisinin yer alması ve Model 1'de ikinci ve beşinci finansal yapı oranları yer alırken Model 2'de finansal yapı oranlarından sadece birincisinin modelde bulunmasıdır.

Modelin grupları doğru sınıflandırma oranı Tablo 4.9'dan da görülebileceği üzere %86,5 olarak bulunmuştur. Sınıflandırma modellerinin performansının değerlendirilmesi için önerilen bir yöntem olan ROC eğrisi altındaki alan 0,9365 olarak bulunmuştur. Bu sonuçla modelin grup ayırımını mükemmel yaptığı söylenebilmektedir.



Şekil 4.4 Model 2 için ROC Eğrisi

Sonuç olarak, işletmelerin Karlılık-Verimlilik'ten oluşan performans göstergelerini etkileyen değişkenlerin tespit edilmesi için Lojistik Regresyon Analizinin kullanılmasının uygunluğu açıkça görülmektedir.

4.5.3 Model 3: Performans Göstergesi Olarak Karlılık ve Büyümeye

İşletme performans kriterlerinden karlılık ve büyüme piyasa göstergeleri değerlerini etkileyen faktörlerin saptanması amacıyla kurulan Model 3 için Diskriminant ve Lojistik Regresyon Analiz sonuçları aşağıda verilmiştir. Bu modelde karlılık ve piyasa göstergeleri için seçilen “aktif karlılığı” ve “satışlardaki reel artış oranı” değerlerini birlikte açıklayan değişkenlerin belirlenmesi amaçlanmıştır.

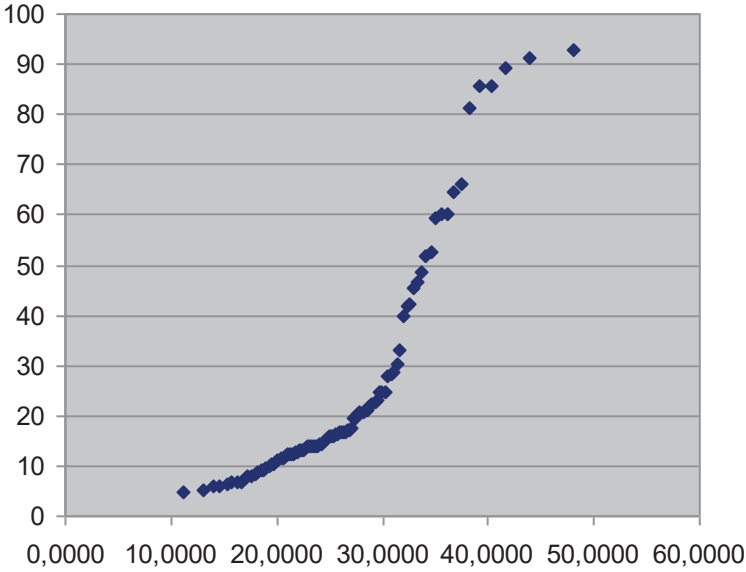
Diskriminant Analizi Sonuçları

Model 3 için uygulanan Diskriminant Analizi sonuçlarına göre, grupların varyanslarının birbirine eşit olduğunu ileri süren sıfır hipotezi Box's M testine göre reddedilmekte (sig. 0,000) ve eşit varyanslılık varsayımı bu modelde de sağlanmamaktadır.

Tablo 4.11 Model 3 için Diskriminant Analizi Testleri ve Sonuçları

Box's M		Eigenvalues			Wilks' Lambda	
Değeri	Sig.	Değeri	Varyans (%)	Kanonik Korelasyon	Değeri	Sig.
1868,295	0,000	0,563	100,00	0,600	0,640	0,000

Ayrıca, çoklu normallik varsayımının da sağlanmadığı çizilen serpilme diyagramından görülmektedir (Şekil 4.5).



Şekil 4.5 Model 3 için Serpilme Diyagramı

Adım Adım Regresyon Yöntemi sonucunda elde edilen standartlaştırılmış diskriminant fonksiyonu Tablo 4.12'de verilen değerler kullanılarak aşağıdaki gibi yazılmaktadır.

$$T = 0,444 \text{ Likidite5} + 0,945 \text{ FinYap2} + 0,528 \text{ Faaliyet1} - 1,060 \text{ Faaliyet7} - 0,627 \text{ YabSermaye}$$

Model sonuçları, bağımsız değişkenlerin sınıfları ayırmadaki anlamlılığına bakan “Grup Ortalamalarının Eşitliği Testi” ile desteklenmektedir.

Tablo 4.12 Model 3 için Diskriminant Modeli ve Doğru Sınıflandırma Sonuçları

Standartlaştırılmış ve Standartlaştırılmamış Kanonik Diskriminant Fonksiyon Katsayıları			Sınıflandırma Sonuçları			
Değişkenler	β	β^*	Sınıf	0	1	Doğru Yüzdesi
Likidite5	3,906	0,444	0	28	13	68,3
FinYap2	5,526	0,945	1	4	50	92,6
Faaliyet1	0,001	0,528	Toplam			81,7
Faaliyet7	-1,072	-1,060				
YabSermaye	-0,023	-0,627				
(Sabit)	-1.030					

Diskriminant Analizi sonuçlarına göre model, bağımlı değişkendeki varyansın yaklaşık olarak 0,36'sini ($0,600^2$) açıklayabilmektedir. Wilks' Lamda değeri ile de diskriminant skorlarındaki toplam varyansın gruplar arasındaki farklar tarafından açıklanamayan kısmının 0,640 olduğu görülmektedir (Tablo 4.11). Aynı zamanda fonksiyonun iyi bir ayırım yapıp yapmadığının göstergesi olan özdeğer (eigenvalues) istatistiğinin anlamlılığında kullanılan Wilks' Lamda değeri Tablo 4.11'den görüldüğü üzere anlamlıdır (sig. 0.000).

Model 3 için yapılan analiz sonucunda diskriminant fonksiyonunun grupları doğru tahmin etme oranı %82,1 olarak bulunmuş olmasına rağmen, varsayımlarının sağlanmaması nedeniyle işletme performans kriterlerini etkileyen faktörlerin belirlenmesi için Lojistik Regresyon Analizi'nin uygulanmasına karar verilmiştir.

Lojistik Regresyon Analizi Sonuçları

Modelin ve parametrelerin uygunluk testinden önce gerçekleştirilen Kesikli Polinomlar Yöntemi sonuçlarına göre, değişkenlerin modelin lojit değerleri ile doğrusal olduğu ve modele değişkenlerin kendi değerlerinin girmesinin 0,01 anlamlılık düzeyinde uygun olduğu görülmüştür.

Modelin uygunluğu için kullanılan Hosmer-Lemeshow test sonucuna göre, 0,064 anlamlılık düzeyinde teorik model veri setini iyi temsil ettiğini ileri süren sıfır hipotezinin anlamlı olduğu kabul edilmektedir.

Model parametrelerinin birlikte anlamlılığını sınavan Omnibus test sonucu incelendiğinde ise, modeldeki bağımsız değişkenlere ait parametrelerin anlamlı olduğu söylenebilmektedir.

Tablo 4.13 Model 3 için Test ve Doğru Sınıflandırma Sonuçları

Test Sonuçları				Sınıflandırma Sonuçları			
				Sınıf	0	1	Doğru Yüzdesi
Testler	χ^2	<i>df</i>	<i>Sig.</i>	0	30	11	73,1
Hosmer-Lemeshow Testi	14,759	8	0,064	1	5	49	90,7
Omnibus Testi	44,385	6	0,000	Toplam			83,2

Çoklu doğrusal bağlantı sorunu ile karşılaşmamak için modele girecek değişkenlerin seçimi yine Adım Adım Regresyon Yöntemi ile yapılmıştır. Modele giren değişkenler Tablo 4.14’te verilmiştir.

Yorumlamadaki kolaylığı nedeniyle, standartlaştırılmış bağımsız değişken değerleri ile model tekrar elde edilmiştir. Tablo 4.14’ten görüldüğü gibi Model 3 için FinYap2 ve Faaliyet7 oranları önemli birer açıklayıcı değişkendir. Değişkenlerin tek tek anlamlılıkları test edildiğinde Faaliyet1’in dışındaki

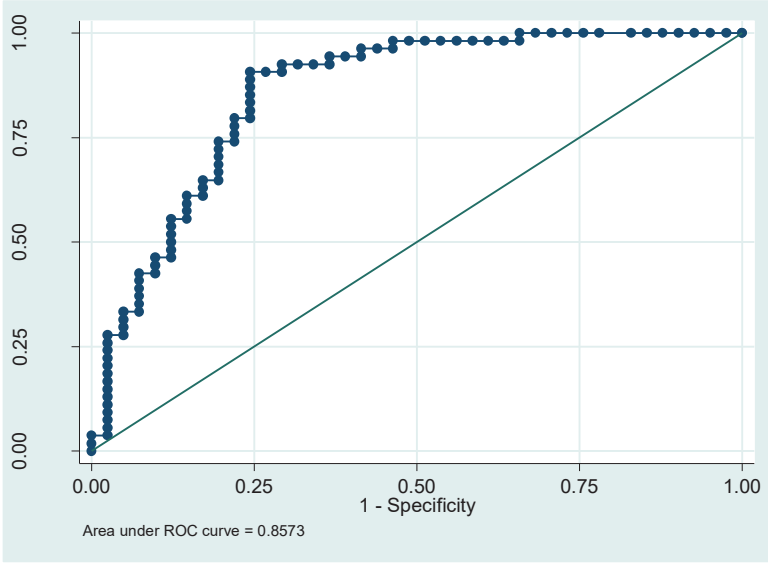
değişkenlerin %0,1 ile %2,8 düzeyinde modelde olması anlamlı bulunmuştur. Faaliyet1'in de anlamlılık düzeyinin 0,057 değeri ile modelde bulunduğu grupları doğru atama oranını arttırdığı ve bu değişkenin işletmelerin nakit politikalarına ilişkin değerlendirme yapılmasına imkan vermesi nedeniyle modelde kalması uygun bulunmuştur.

Tablo 4.14 Model 3'deki Açıklayıcı Değişkenler ve Katsayıları

Modele Giren Değişkenler				Standartlaştırılmış Değişkenler için Değerler	
Değişkenler	β	<i>Wald</i>	<i>Sig.</i>	β^*	$\exp(\beta^*)$
FinYap2	-9,099	15,028	0,000	-1,693	0,184
Faaliyet7	2,272	11,301	0,001	2,255	9,534
YabSermaye	0,039	5,178	0,023	1,033	2,810
Likidite5	-6,762	4,831	0,028	-0,781	0,458
Faaliyet1	-0,003	3,614	0,057	-1,663	0,190
(Sabit)	1,558	0,468	0,036	0,209	1,233

Modelin grupları doğru sınıflandırma yüzdesi Tablo 4.13'ten de görüldüğü gibi yanlış sınıflandırma maliyetlerinin birbirine eşit olduğu varsayımıyla, kesim değeri (cut value) "0,50" olarak belirlenmiştir. Analiz sonucunda düşük performanslı işletmeleri doğru sınıflandırma oranı olan duyarlılık (sensitivity) değeri %73,1, yüksek performanslı işletmeleri doğru sınıflandırma oranı olan özgünlük (specificity) değeri %90,7, toplamda grupları doğru sınıflandırma oranı olan doğruluk (accuracy) değeri ise % 83,2 olarak elde edilmiştir.

ROC eğrisi çizildiğinde de eğrinin altında kalan alanın 0,8573 değeriyle $0,8 \leq ROC < 0,9$ aralığında yer aldığı ve sonuç olarak modelin, grupları mükemmel ayırdığı görülmektedir.



Şekil 4.6 Model 3 için ROC Eğrisi

Sonuç olarak, işletmelerin Karlılık-Büyüme'den oluşan performans göstergelerini etkileyen değişkenlerin tespit edilmesi amacıyla Lojistik Regresyon Analizi kullanılmasının uygunluğu açıkça görülmektedir.

4.5.4 Model 4: Performans Göstergesi Olarak Karlılık ve Piyasa

Bu modelde karlılık ve piyasa performans göstergeleri için seçilen “aktif karlılığı” ile “fiyat-kazanç” değerlerini birlikte açıklayan faktörlerin saptanması amaçlanmıştır.

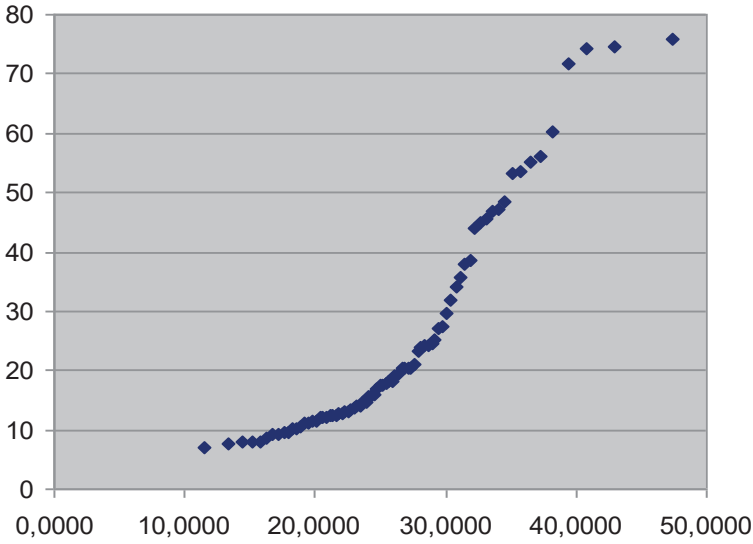
Diskriminant Analizi Sonuçları

Model 4 için uygulanan Diskriminant Analizi sonuçlarına bakıldığında grupların varyansların birbirine eşit olduğunu ileri süren sıfır hipotezi Box's M testine göre reddedilmekte (sig. 0,000) ve yine eşit varyanslılık varsayımı sağlanmamaktadır.

Tablo 4.15 Model 4 için Diskriminant Analizi Testleri ve Sonuçları

Box's M		Eigenvalues			Wilks' Lambda	
Değeri	Sig.	Değeri	Varyans (%)	Kanonik Korelasyon	Değeri	Sig.
1792,673	0,000	0,617	100,00	0,618	0,600	0,000

Ayrıca, çoklu normallik varsayımının da sağlanmadığı çizilen serpilme diyagramından görülmektedir (Şekil 4.7).

**Şekil 4.7 Model 4 için Serpilme Diyagramı**

Adım Adım Regresyon Yöntemi sonucunda elde edilen standartlaştırılmış diskriminant fonksiyonu Tablo 4.16’da verilen değerler kullanılarak aşağıdaki gibi yazılmaktadır.

$$T = -0,913FinYap1 + 0,505PazarPay + 0,448Buyukluk$$

Model sonuçları, bağımsız değişkenlerin sınıfları ayırmadaki önemini ortaya koyan “Grup Ortalamalarının Eşitliği Testi” ile desteklenmektedir.

Tablo 4.16 Model 4 için Diskriminant Modeli ve Doğru Sınıflandırma Sonuçları

Standartlaştırılmış ve Standartlaştırılmamış Kanonik Diskriminant Fonksiyon Katsayıları			Sınıflandırma Sonuçları			
Değişkenler	β	β^*	Sınıf	0	1	Doğru Yüzdesi
FinYap1	-5,240	-0,913	0	21	10	67,6
PazarPay	6,939	0,505	1	6	41	87,2
Buyukluk	1,213	0,448	Toplam			79,5
(Sabit)	1,747					

Diskriminant Analizi sonuçlarına göre model, bağımlı değişkendeki varyansın yaklaşık olarak 0,38'ini (0,618²) açıklayabilmektedir. Wilks' Lamda değeri ile de diskriminant skorlarındaki toplam varyansın gruplar arasındaki farklar tarafından açıklanamayan kısmının 0,600 olduğu görülmektedir (Tablo 0.15) Aynı zamanda fonksiyonun iyi bir ayırım yapıp yapmadığının göstergesi olan ve özdeğer (eigenvalues) istatistiğinin anlamlılığında kullanılan Wilks' Lamda değerinin anlamlı olduğu görülmektedir (sig. 0.000).

Model 4 için yapılan analiz sonucunda diskriminant fonksiyonunun grupları doğru sınıflandırma oranı %79,5 olarak bulunmuş olmasına rağmen, varsayımlarının sağlanmaması nedeniyle işletme performans kriterlerini etkileyen faktörlerin tespiti için Lojistik Regresyon Analizi'nin uygulanmasına karar verilmiştir.

Lojistik Regresyon Analizi Sonuçları

Karlılık-Piyasa kriterlerine göre işletme performansını etkileyen faktörler olarak Adım Adım Regresyon Yöntemi sonucunda modele giren değişkenler; FinYap1, PazarPayı ve Likidite1 oranlarıdır (bkz. Tablo 4.18). Bu değişkenler ile modelden hesaplanan lojit değerlerinin doğrusallık kontrolleri yapıldığında, modele giren üç değişkenin de 0,01 anlamlılık düzeyinde doğrusal olduğu görülmektedir.

Modelin uygunluğu için kullanılan Hosmer-Lemeshow testi sonucuna göre modelin veri setini iyi temsil ettiğini ileri süren sıfır hipotezi 0,530 anlamlılık düzeyinde kabul edilmiştir. Ayrıca Omnibus test sonucuna bakılarak, modeldeki üç parametrenin birlikte modelde bulunmasının anlamlı olduğuna karar verilmiştir.

Tablo 4.17 Model 4 için Test ve Doğru Sınıflandırma Sonuçları

Test Sonuçları				Sınıflandırma Sonuçları			
				Sınıf	0	1	Doğru Yüzdesi
Testler	χ^2	<i>df</i>	<i>Sig.</i>	0	23	8	74,2
Hosmer-Lemeshow Testi	7,061	8	0,530	1	6	41	87,2
Omnibus Testi	44,422	3	0,000	Toplam			82,1

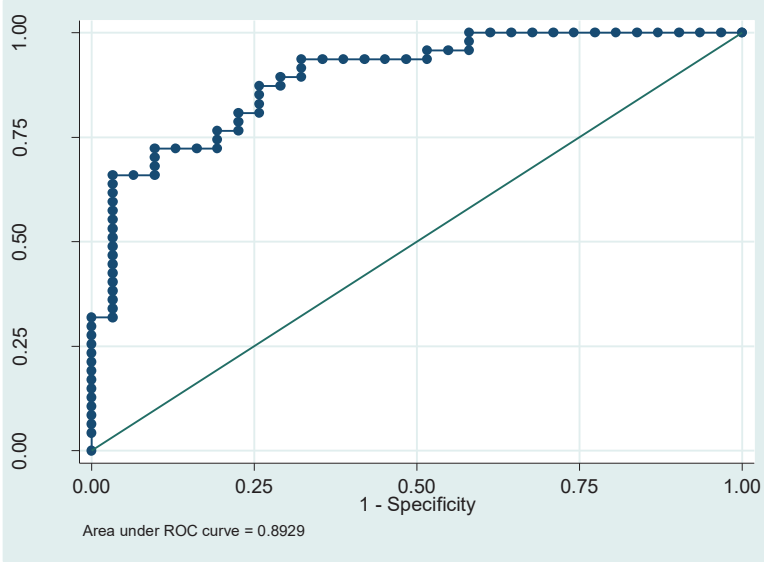
Model kurulurken kullanılan Adım Adım Regresyon Yöntemi'ne göre, modele giren değişkenler ve yorumlamadaki kolaylığı ile kullanılan standartlaştırılmış değişkenler için değerleri Tablo 4.18'de verilmiştir ve FinYap1 değişkenini 0,000 anlamlılık düzeyi, PazarPay değişkenininin %0,5 ve Likidite1 değişkenininin de %4,9 düzeyinde modelde kalması anlamlı bulunmuştur.

Tablo 4.18 Model 4'teki Açıklayıcı Değişkenler ve Katsayıları

Modele Giren Değişkenler				Standartlaştırılmış Değişkenler için Değerler	
Değişkenler	β	<i>Wald</i>	<i>Sig.</i>	β^*	$\exp(\beta^*)$
FinYap1	-12,687	15,356	0,000	-2,504	0,082
PazarPay	65,391	7,948	0,005	4,902	134,604
Likidite1	-0,583	3,888	0,049	-0,931	0,394
(Sabit)	5,536	9,755	0,002	1,478	4,382

Modelin grupları doğru sınıflandırma yüzdesi Tablo 4.17'de görüldüğü gibi % 82,1 olarak elde edilmiştir. ROC eğrisi altında kalan alan ise 0,8929 olarak

bulunmuştur (Şekil 4.8). Bu sonuçla modelin grup ayrımını mükemmel yaptığı söylenebilmektedir.



Şekil 4.8 Model 4 için ROC Eğrisi

Sonuç olarak, işletmelerin Karlılık-Piyasa performans göstergelerini birlikte açıklayan değişkenlerin belirlenmesi için Lojistik Regresyon Analizi sonuçlarının kullanılmasının uygunluğu açıkça görülmektedir.

4.5.5 Model 5: Performans Göstergesi Olarak Verimlilik ve Büyüme

Bu modelde verimlilik ve büyüme performans göstergeleri için seçilen “emek verimliliği” ile “satışlardaki reel artış” değerlerini birlikte açıklayan faktörlerin saptanması amaçlanmıştır. Diskriminant ve Lojistik Regresyon Analiz sonuçları aşağıda verilmiştir.

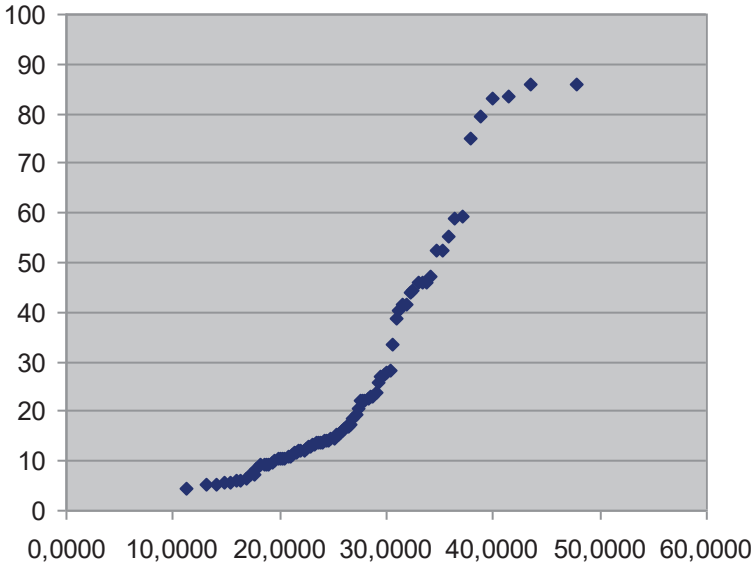
Diskriminant Analizi Sonuçları

Model 5 için uygulanan Diskriminant Analizi sonuçları incelendiğinde, grupların varyansların birbirine eşit olduğunu ileri süren sıfır hipotezi Box’s M testine göre reddedilmekte (sig. 0,000) ve eşit varyanslılık varsayımının bu modelde de sağlanmadığı görülmüştür.

Tablo 4.19 Model 5 için Diskriminant Analizi Testleri ve Sonuçları

Box's M		Eigenvalues			Wilks' Lambda	
Değeri	Sig.	Değeri	Varyans (%)	Kanonik Korelasyon	Değeri	Sig.
1966,907	0,000	0,502	100,00	0,578	0,666	0,000

Ayrıca, çoklu normallik varsayımının da sağlanmadığı Şekil 4.9'da verilen serpilme diyagramından görülmektedir.

**Şekil 4.9 Model 5 için Serpilme Diyagramı**

Adım Adım Regresyon Yöntemi sonucunda elde edilen standartlaştırılmış diskriminant fonksiyonu Tablo 0.20'de verilen değerler kullanılarak aşağıdaki gibi yazılmaktadır.

$$T = -0,700 \text{ FinYap2} + 0,386 \text{ Faaliyet4} + 0,444 \text{ SermayeYg1} \\ + 0,712 \text{ YabSermaye}$$

Model sonuçları, bağımsız değişkenlerin sınıfları ayırmadaki anlamlılığına bakan “Grup Ortalamalarının Eşitliği Testi” ile desteklenmektedir.

Tablo 4.20 Model 5 için Diskriminant Modeli ve Doğru Sınıflandırma Sonuçları

Standartlaştırılmış ve Standartlaştırılmamış Kanonik Diskriminant Fonksiyon Katsayıları			Sınıflandırma Sonuçları			
Değişkenler	β	β^*	Sınıf	0	1	Doğru Yüzdesi
FinYap2	-3,504	-0,700	0	30	11	73,2
Faaliyet4	0,398	0,386	1	8	39	83,0
SermayeYog1	0,000	0,444	Toplam			78,4
YabSermaye	0,028	0,712				
(Sabit)	-0,445					

Diskriminant Analizi sonuçlarına göre model, bağımlı değişkendeki varyansın yaklaşık olarak 0,334'ünü ($0,578^2$) açıklayabilmektedir. Wilks' Lamda değeri ile de diskriminant skorlarındaki toplam varyansın gruplar arasındaki farklar tarafından açıklanamayan kısmının 0,666 olduğu görülmektedir (Tablo 4.19). Aynı zamanda fonksiyonun iyi bir ayırım yapıp yapmadığının göstergesi olan ve özdeğer (eigenvalues) istatistiğinin anlamlılığında da kullanılan Wilks' Lamda değeri anlamlıdır (sig. 0.000).

Model 5 için yapılan analiz ile diskriminant fonksiyonunun grupları doğru tahmin etme oranı %78,4 olarak bulunmuş olmasına rağmen, varsayımlarının sağlanmaması nedeniyle işletme performans kriterlerini etkileyen faktörlerin saptanması için Lojistik Regresyon Analizi'nin uygulanmasına karar verilmiştir.

Lojistik Regresyon Analizi Sonuçları

Verimlilik-Büyüme kriterlerine göre işletme performansı değerlendirilmek istendiğinde, bu performans göstergelerini etkileyen açıklayıcı değişkenler Tablo 4.22'de yer almaktadır.

Mode giren değişkenlerle bu modelden hesaplanan lojit değerleri arasında doğrusallık kontrolü sonucunda değişkenlerin modelde dönüşüm yapılmadan kullanılmasının uygun olduğuna karar verilmiştir.

Modelin uygunluğu için Hosmer-Lemeshow testi sonucu incelendiğinde, modelin veri setini iyi temsil ettiği söylenebilmektedir (sig. 0.444). Ayrıca modeldeki parametrelerin birlikte anlamlılığını test eden Omnibus testi sonucuna bakılarak parametrelerin birlikte modelde olmasının anlamlı olduğuna karar verilmiştir (sig.0.000).

Tablo 4.21 Model 5 için Test ve Doğru Sınıflandırma Sonuçları

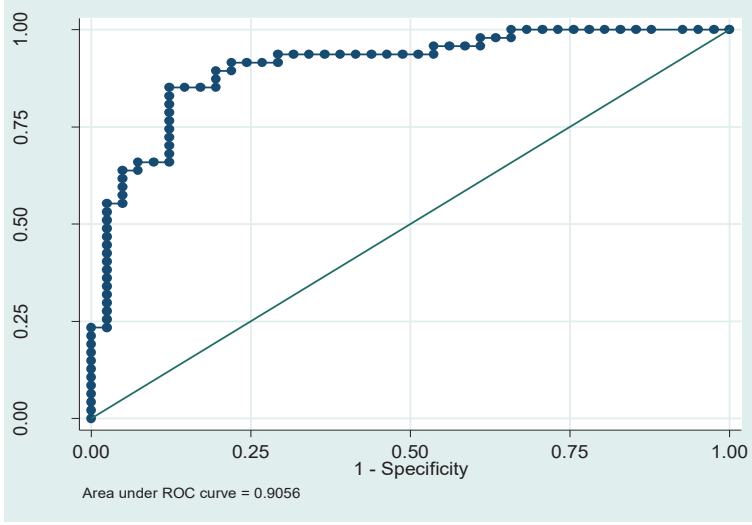
Test Sonuçları				Sınıflandırma Sonuçları			
				Sınıf	0	1	Doğru Yüzdesi
Testler	χ^2	<i>df</i>	<i>Sig.</i>	0	33	8	80,5
Hosmer-Lemeshow Testi	7,890	8	0,444	1	6	41	87,2
Omnibus Testi	53,593	5	0,000	Toplam			84,1

Adım Adım Regresyon Yöntemi'yle kurulan Lojistik Regresyon Modeli'ndeki bağımsız değişkenler Tablo 4.22'de görülmektedir. $exp(\beta)$ değerlerinin yorumlanmasında sağladığı kolaylık nedeniyle, standartlaştırılmış bağımsız değişken değerleriyle model tekrar kurulmuş ve parametrelerin %0,1 ile %2,4 düzeyinde anlamlı oldukları saptanmıştır.

Tablo 4.22 Model 5'deki Açıklayıcı Değişkenler ve Katsayıları

Modele Giren Değişkenler				Standartlaştırılmış Değişkenler için Değerler	
Değişkenler	β	<i>Wald</i>	<i>Sig.</i>	β^*	$exp(\beta^*)$
YabSermaye	0,056	10,737	0,001	1,476	4,376
Faaliyet7	2,729	8,875	0,003	1,257	3,516
FinYap2	-7,337	8,344	0,004	-1,570	0,208
SermayeYog2	0,000	7,838	0,005	3,232	25,338
Likidite5	-7,240	5,066	0,024	-0,868	0,420
(Sabit)	-1,205	1,189	0,275	0,522	1,685

Tablo 4.21’den de görülebileceği üzere, modelin sınıfları doğru tahmin etme oranı %84,1 olarak bulunmuştur. Bu orana alternatif bir yaklaşım olan ROC eğrisi değerine bakıldığında ise eğrinin altında kalan alanın 0,9056 olduğu ve modelin sınıfları olağanüstü ayırdığı sonucuna varılabilmektedir.



Şekil 4.10 Model 5 için ROC Eğrisi

Sonuç olarak, işletmelerin Verimlilik-Büyüme’den oluşan performans göstergelerini etkileyen değişkenlerin belirlenmesi amacıyla, kategorik bağımlı değişken analizlerinden olan Lojistik Regresyon Analizi’nin kullanılmasının uygunluğu açıkça görülmektedir.

4.6 Uygulama Sonuçları ve Yorumları

Çalışmanın sonuç bölümünde işletmenin finansal performansını (karlılık, verimlilik, büyüme ve piyasa) etkileyen faktörler uygulama bölümünde belirlenen kriterler dikkate alınarak elde edilen modeller çerçevesinde değerlendirilmiştir. Sonuçlar yorumlanırken verilerin elde edildiği 2007 yılına ait ekonomik koşullar da dikkate alınmıştır. Modeller incelendiğinde Lojistik Regresyon Analizi sonuçları şu şekilde elde edilmiştir:

- **Model 1’de (Karlılık-Verimlilik-Büyüme)** yer alan değişkenler: SermayeYog2, FinYap2(-), Büyüklük, YabSermaye, FinYap5 ve Likidite5(-).
Grupları doğru sınıflandırma oranı: % 84,8.
- **Model 2’de (Karlılık-Verimlilik)** yer alan değişkenler: FinYap1(-), SermayeYog2, Büyüklük, YabSermaye.
Grupları doğru sınıflandırma oranı: % 86,5.
- **Model 3’de (Karlılık-Büyüme)** yer alan değişkenler: FinYap2(-), Faaliyet7, Likidite5(-) ve Faaliyet1.
Grupları doğru sınıflandırma oranı: % 83,2.
- **Model 4’te (Karlılık-Piyasa)** yer alan değişkenler: FinYap1(-), PazarPay ve Likidite1(-).
Grupları doğru sınıflandırma oranı: % 82,1.
- **Model 5’de (Verimlilik-Büyüme)** yer alan değişkenler: YabSermaye, Faaliyet7, FinYap2(-), SermayeYog2 ve Likidite5(-).
Grupları doğru sınıflandırma oranı: % 84,1.

Çalışmada Model 4 (Karlılık-Piyasa) dışında diğer tüm modeller performansı etkileyen faktör olarak “yabancı sermaye”yi belirlenmiştir. Bu değişkenin 2007 yılında açıklayıcı bir değişken olmasındaki nedenin, Türkiye’ye gelen yabancı sermaye (Bkz. Tablo 2.2) hacmindeki artış olduğu söylenebilir. İşletmelerin performanslarının değerlendirilmesinde faaliyet gösterdikleri ülkelerin ekonomik politikalarının ve yatırım kararlarının etkisini görmek mümkündür. Türkiye’nin 2007 yılında ekonomik istikrar yaşaması ve AB uyum sürecinde doğrudan yabancı sermaye yatırımları* çekmek için yaptığı yasal düzenlemelerin etkisiyle yüksek seviyede gerçekleşen yabancı sermaye akışı,

* “Doğrudan yabancı sermaye yatırımları uluslararası sermaye akışı demektir: ya yabancı bir ülkede bir firma kurulur, ya da yabancı ülkedeki bir işletmeye sermaye yatırılarak ana firma büyür.” Suna Oksay, “Çokuluslu Şirketler Teorileri Çerçevesinde, Yabancı Sermaye Yatırımlarının İncelenerek, Değerlendirilmesi”, Çevirimiçi
<http://www.econturk.org/Turkiyeekonomisi/oksay3.pdf>, 10 Eylül 2011.

performans kriterlerini etkileyen faktörlerde önemli bir açıklayıcı değişken olarak görülmektedir. Bununla birlikte aynı dönem için yabancı sermayenin hem ekonomik gelişmelere hem de işletmelerin performansına olan etkisinin önemi de dikkat edilmesi gereken bir durumdur.

Rekabet koşullarının her geçen gün daha da zorlaşması işletmelerin rakipleri karşısında daha esnek, yenilikçi ve farklılık yaratmaları gerekliliğini ortaya çıkartmaktadır. Böylece işletmeler maliyetlerini düşürerek kaliteli ürün üretme yarışına girmişlerdir. Sermaye yoğunluğu işletme verimliliğini doğrudan etkileyen bir değişken olarak analiz sonuçlarında yer almıştır. Model 1 (Karlılık-Verimlilik-Büyüme), Model 2 (Karlılık-Verimlilik) ve Model 5 (Verimlilik-Büyüme) incelendiğinde “sermaye yoğunluğu” değişkeninin işletmelerin performansı değerlendirilirken önemli bir kriter olduğu sonucuna varılmıştır. Bu değişkenin modellerin içinde yer alması, performans göstergelerinden özellikle “verimlilik” değişkeninin açıklanmasında önemli bir faktör olmasından kaynaklanabilir. İşletme verimliliği ve sermaye yoğunluğu arasındaki doğrudan ve yüksek ilişkinin işletmelerin 2007 yılında finansal yapılarının güçlenmesiyle (yabancı sermaye yatırımlarıyla birlikte) fazla fonları Ar-Ge faaliyetleri için değerlendirmelerinden kaynaklandığı sonucuna varmak mümkündür.

2007 yılının en önemli ekonomik faaliyetlerinden olan ihracattaki büyüme (Bkz. Tablo 0.1), işletmelerin performanslarını özellikle karlılık ve verimlilik açısından etkilediği görülmüştür. İşletmelerin artan ihracatları satış hacimlerini yükseltmiş ve bu durum pazar payında olumlu etki yaratmıştır. İhracat verimliliğinin artmasına yol açabildiği gibi, verimliliğin artması da ihracatın gelişmesinde etkili olabilmektedir. Dış piyasa ile olan rekabet işletmelerin düşük fiyat avantajı elde edebilmek için maliyetlerini azaltma konusunda baskı oluşturmakta ve uluslararası pazarlara olan açılımla da işletmeler yeni

teknolojileri kullanmaya zorlamaktadır. Bu durum da doğrudan verimliliği arttırıcı bir faktör olarak ortaya çıkmaktadır.²⁶²

İhracat genişlemesi faktör verimliliğini yükseltmekte, teknolojik yeniliklerin adaptasyonu ve kaynakların daha etkin kullanımını sağlamaktadır. Ayrıca, yabancı rekabetin getireceği avantajlar ve uluslararası piyasalara açılımın doğurduğu ölçek etkisinin getirisi gibi unsurlar dolayısı ile ekonomik büyümede artış gözlemlenmektedir.²⁶³

Çalışmada yer alan “büyüklük” değişkeni de (kukla değişken), Model 1 (Karlılık-Verimlilik-Büyüme) ve Model 2 (Karlılık-Verimlilik) için işletmeleri performansa göre sınıflamada modele girmiştir. İşletmelerin karlılık, verimlilik ve büyüme gibi performans kriterleri incelenirken yeni teknoloji yatırımlarının ve bu yatırımlar için işletmelerin katlandığı Ar-Ge harcamalarına hak ettiği önemin verilmesi gerekmektedir. Yeni teknolojilere yapılan yatırımlar işletmelerin üretim maliyetlerini düşürdüğünden karlılığın artmasında önemli bir faktör haline gelmektedir. Araştırmalar satışlardaki artış veya karlılık gibi performans göstergeleri ile Ar-Ge harcamaları arasında güçlü ilişkiler olduğunu ortaya koymaktadır²⁶⁴. Yapılan çalışmalarda pazar payları yüksek olan işletmelerin karlılık düzeyinin de yüksek olduğu sonucuna varılmıştır²⁶⁵.

²⁶² Bayfuss, 1987, p.175.

²⁶³ Demetrios Moschos, “Export Expansion, Growth and the Level of Economic Development: An Empirical Analysis”, **Journal of Development Economics**, Vol: 30 Issue 1, 1989, p. 93.

Epaminondas Panas, George Wamvoukas, “Further Evidence on The Export-Led Growth Hypothesis”, **Applied Economics Letters**, 9, 2002, p.731.

²⁶⁴ Graham K. Morbey, Robert M. Reithner, “How Research and Development Affects Sales Growth, Productivity and Profitability”, **Research-Technology Management**, Vol: 33, 3, 1990, pp. 11–15.

Henry G. Grabowski, Dennis C. Mueller, “Industrial Research and Development, Intangible Capital Stocks, and Firm Profit Rates”, **The Bell Journal of Economics**, Vol: 9, No. 2, 1978, pp. 328–343.

²⁶⁵ David M. Szymanski, Sundar G. Bharadwaj, P. Rajan Varadarjan, “An Analysis of the Market Share-Profitability Relationship”, **The Journal of Marketing**, Vol: 57, No. 3, 1993, pp. 1–18.

Büyükölük deęişkeni işletmelerin ekonomiyeye yaptıęı katkı olarak ele alındığında performanslarını etkileyen hatta arttırıcı bir deęişken olarak öne çıkmaktadır.

İşletmelerin finansal performanslarını etkileyen faktörlerin araştırılmasına yönelik geniş bir literatür bulunmaktadır. Yapılan araştırmalar incelendikçe birçok çalışmada birbiri ile tutarlı sonuçlar elde edilirken bazılarının ise tam tersi sonuçlara ulaştığı görülmektedir. Bu çalışmalar özellikle işletmelerin sermaye yapısına (borç ve öz sermaye) ilişkin yapılan araştırmalardan oluşmaktadır. İşletmelerin borçlanma stratejilerinin performanslarını ne kadar etkilediğini ve borçlanma yöntemlerinden hangilerinin performansı daha çok arttırdığı yönünde yapılan araştırmalar finans literatüründe geniş yer bulan sermaye yapısı teorileriyle* de desteklenmektedir. Araştırmacıların mikro düzeyde işletmelerin performanslarını arttırıcı yönde neler yapabilecekleri üzerinde bu kadar çok durmalarının en önemli nedeni, işletmelerin performanslarının iyi olmasının ülkelerin ekonomik etkinliğinin ölçülmesinde önemli göstergelerden biri olmasından kaynaklanmaktadır. Kayo ve Kimura (2011)²⁶⁶ yaptıkları çalışmada ülkelerin ekonomik özelliklerinin işletmelerin borçlanma yapılarını etkilediğini belirtmişlerdir. Bunun yanı sıra işletmelerin borçlanma yapılarının ağırlıklı olarak kendi içsel özelliklerinden kaynaklandığının da altını çizmişlerdir.

İşletmelerin performanslarını ölçmek için farklı finansal oranlardan yararlanılmaktadır. Çalışmalarda genel olarak finansal performans ölçütü olarak aktif karlılığı, net karlılık, öz sermaye karlılığı gibi oranların kullanıldığı gözlenmiştir. Bu çalışmada işletmelerin performansının belirlenmesinde hem finansal hem de finansal olmayan ölçütlerden yararlanılarak çeşitli kombinasyonlarla modeller oluşturulmuş ve deęişkenler incelenmiştir.

* Sermaye yapısı teorileri hisse senetleri fiyatlarının ve sermaye maliyetlerinin sermaye yapısı tarafından nasıl etkilendiğini ortaya koymak için geliştirilmiştir. Finans literatüründe kullanılan ve optimal sermaye yapısını açıklamaya yönelik temel teoriler dört kısımda incelenmektedir. Dengeleme, Hiyerarşi, Vekalet ve Açıklama Teorileri. (<http://eidergisi.istanbul.edu.tr/sayil1/ueis11m3.pdf>)

²⁶⁶ Eduardo K. Kayo, Herbert Kimura, “Hierarchical Determinants of Capital Structure”, **Journal of Banking & Finance**, 35, 2011, pp. 358–371.

Yapılan çalışmalar incelendiğinde elde edilen sonuçların ülkelerin ekonomik gelişmişlik düzeyi, araştırılan dönem, işletmelerin faaliyet kolları (özellikle hizmet, üretim) gibi farklılıklar göstermektedir. Türkiye gibi gelişmekte olan ülkelerde faaliyet gösteren işletmelerin performansı değerlendirilirken de benzer kategoride olan ülkeler üzerine yapılan çalışmalar ele alınmıştır. Ekonomik gelişmişlik düzey farklılığı nedeniyle gelişmiş ülkeler ve gelişmekte olan ülkelerle yapılan araştırmaların sonuçlarında farklılıklar olduğu belirlenmiştir.

Bütün modeller ayrı ayrı incelendiğinde finansal yapı oranlarından Kısa Vadeli Yük./Toplam Varlıklar ve Toplam Yük./Toplam Varlık oranlarının anlamlı ve negatif yönlü bir ilişkiye sahip olduğu görülmektedir. Bu durum işletmelerin uzun vadeli ve öz sermayeyle finansman yolunu tercih ettiği anlamına gelmektedir. Finansal yapısı güçlü olan işletmelerin iyi performans göstermesi beklenmektedir. 2007 yılı itibariyle de işletmelerin, özellikle öz sermayenin varlık finansmanında payının arttığı görülmektedir.²⁶⁷ Yapılan düzenlemelerle Türkiye’de yatırım fırsatlarını değerlendirmek isteyen yabancı yatırımcıların sağladığı fonların işletmelerin finansal yapılarının güçlenmesine katkı da bulunduğu sonucuna varılabilir.

İşlemelerin borç oranı, karlılık ya da işletme riski yüksek olduğu zaman düşük ya da negatif yönlüdür. İşletmelerin maddi olmayan duran varlıkları ile düşük borç oranı birbirleri ile ilişkilidir. Büyüme fırsatlarına sahip şirketlerin düşük borç oranına sahip olduğu görülmektedir²⁶⁸. Çalışmadaki veri seti incelendiğinde de büyüme oranının yüksek olduğu işletmelerin düşük borç oranına sahip olduğu gözlenmiştir. Aşağıda özetlenmiş olan, finansal yapı ile işletme performansı arasındaki ilişkiyi inceleyen çalışmalarda birbirinden farklı sonuçlar elde edilse de genel olarak ilişkinin negatif yönlü olduğu sonucuna varılmıştır.

²⁶⁷ **İstanbul Sanayi Odası Dergisi**, “Türkiye’nin 500 Büyük Sanayi Kuruluşu 2007, Özel Sayı, Sayı 509, Ağustos 2008, s. 165.

²⁶⁸ Stewart C. Myers, “Capital Structure”, **Journal of Economic Perspectives**, Vol:15, Number 2, Spring 2001, pp. 81–102.

Ebaid, Ibrahim El-Sayed (2009)²⁶⁹ Mısır için yaptıkları çalışmada işletme performansını öz sermaye karlılığı, aktif karlılığı ve brüt kar gibi değişkenlerle ölçerken çoklu regresyon yöntemi kullanarak sermaye yapısı ile işletme performansı arasında herhangi bir ilişki belirleyememişlerdir. Booth ve diğerleri (2001)²⁷⁰ ve Wiwattanakantang (1999)²⁷¹ gelişmekte olan ülkeler için yaptıkları çalışmalarda finansal yapı (kaldıraç) oranları ile karlılık arasında negatif yönlü bir ilişki olduğu sonucuna varmışlardır.

Friend ve Lang (1988)²⁷², Titman ve Wessel (1988)²⁷³, Kester (1986)²⁷⁴ ve Myers (2001)²⁷⁵ gelişmiş ülkeler için finansal yapı oranları ile karlılık arasında negatif yönlü bir ilişki olduğu sonucuna varmışlardır. Long ve Malitz (1985)²⁷⁶ ile Wald (1999)²⁷⁷ ise bu ilişkinin pozitif yönlü olduğunu saptayarak literatürde yapılan çalışmaların farklı sonuçlar elde ettiğini göstermişlerdir.

²⁶⁹ Ibrahim El-Sayed Ebaid, “The Impact of Capital-Structure Choice on Firm Performance: Empirical Evidence from Egypt”, **Journal of Risk Finance**, Vol: 10, Issue 5, 2009, pp. 477–487.

²⁷⁰ Laurence Booth, Varouj Aivazian, Aslı Demircuc-Kunt, Vojislav Maksimovic, “Capital Structures in Developing Countries”, **The Journal of Finance**, The Journal of Finance, Vol: LVI, No. 1, February 2001, pp. 87–130.

²⁷¹ Yupana Wiwattanakantang, “An Empirical Study on the Determinants of the Capital Structure of Thai Firms”, **Pacific-Basin Finance Journal**, Vol: 7, 1999, pp. 371–403.

²⁷² Irwin Friend, Larry H. P. Lang, “An Empirical Test of the Impact of Managerial Self-Interest on Corporate Capital Structure”, **The Journal of Finance**, Vol: WLIII, No. 2, June 1988, pp. 271–281.

²⁷³ Sheridan Titman, Roberto Wessels, “The Determinants of Capital Structure Choice”, **The Journal of Finance**, Vol: XLIII, No. 1, March 1988, pp. 1–19.

²⁷⁴ W. Carl Kester, “Capital and Ownership Structure: A Comparison of United States and Japanese Manufacturing Corporations”, **Financial Management in Japan**, Spring 1986, pp. 5–16.

²⁷⁵ Stewart C. Myers, **a.g.e.**, p. 83.

²⁷⁶ Michael S. Long, Ilen B. Malitz, “Investment Patterns and Financial Leverage”, **National Bureau of Economic Research**, Ed. Benjamin M. Friedman, Chapter 9, 1985, pp. 325–352.

²⁷⁷ John K. Wald, “How Firm Characteristics Affect Capital Structure: An International Comparison”, **The Journal of Financial Research**, Vol: XXII, No. 2, Summer 1999, pp. 161–187.

Adekunle ve Sunday (2010)²⁷⁸ Nijerya için 30 firma ile yaptıkları çalışmada borç oranının karlılıkla negatif yönlü ilişkisi olduğunu belirlemiştir. Chhibber ve Majumdar (1997)²⁷⁹ ile Rao, M-Yahyaee ve Syed (2007)²⁸⁰ ise yaptıkları çalışmalarda finansal kaldıraç ile işletme performansı arasında negatif bir ilişki elde etmişlerdir. Bu çalışmalarda ayrıca finansal performansın likidite üzerindeki etkisinin de negatif olduğu vurgulanmıştır.

Finansal yapı oranlarının performansa etkisi incelendikten sonra performansı etkileyen bir diğer değişken olan likidite oranları da çalışmada Model 3 (Karlılık-Büyüme), Model 4 (Karlılık-Piyasa) ve Model 5'e (Verimlilik-Büyüme) ters işaretli olarak girmiştir. Eriotis (2007)²⁸¹ yapmış olduğu çalışmada işletmenin borçlanma oranı ile büyüme oranı, likidite oranı ve faiz karşılama oranları arasında negatif bir ilişki olduğunu belirtirken işletmelerin büyüklüğünün borçlanma oranı ile pozitif ilişkisi olduğunu ortaya koymuştur.

İşletme yöneticilerinin amaçlarında birisi varlık devir hızını (toplam varlık devir hızını) arttırmaktır. Bu değer artırılmasında toplam varlıklara göre satış gelirlerinin maksimizasyonu önemlidir. Varlık devir hızı (finansal verimlilik) arttırılmaksızın elde edilen kar artışı olumlu bir gelişme olarak değerlendirilemez. Bunun nedeni karın hesaplanması sırasında stok değerlendirme yöntemleri, finansal tabloların hazırlanması sırasında genel kabul görmüş muhasebe ilkelerine tam olarak uyulup uyulmadığı gibi konulardaki belirsizlikler işletmeler için elde edilen kar değerini etkilemektedir.²⁸²

²⁷⁸ Adekunle A. Onaolapo, Sunday O. Kajola, "Capital Structure and Firm Performance: Evidence from Nigeria", **European Journal of Economics, Finance and Administrative Sciences**, Issue 25, 2010, pp. 70–82.

²⁷⁹ Pradeep K. Chhibber, Sumit K. Majumdar, "Foreign Ownership and Profitability: Property Rights, Strategic Control and Corporate Performance in Indian Industry", **The William Davidson Institute**, Working Paper Number 64, April 1997, pp. 1–39.

²⁸⁰ Narendar V. Rao, Khamis Hamed Mohamed Al-Yahyaee, Lateef A. M. Syed, "Capital Structure and Financial Performance: Evidence from Oman", **Indian Journal of Economics and Business**, June 2007.

²⁸¹ Nikolaos Eriotis, "How Firm Characteristics Affect Capital Structure: An Empirical Study", **Managerial Finance**, Vol: 33 No. 5, 2007, pp. 321–331.

²⁸² Hüseyin Yılmaz, "Finansal Verimlilik Artışının Finans Maliyetlerine Etkisini Belirlemeye Yönelik Bir İnceleme", **V. Ulusal Üretim Araştırmaları Sempozyumu**, İstanbul Ticaret Üniversitesi, 25–27 Kasım 2005, s. 484.

Bir ekonomide imalat ve hizmet sektöründe varlık devir hızlarını arttırmaksızın sorunlara çözüm üretmek zorlaşmaktadır. Bu bağlamda işletmelerin gerçek anlamda başarılı olması finansman gereksinimlerini ve giderlerini azaltmak için kar marjları negatif olmaması koşuluyla devir hızlarını arttırmalarına bağlı olmaktadır.²⁸³ Çalışmada da Varlık Devir Hızı oranı hem Model 5 (Verimlilik-Büyüme) hem de Model 3'te (Karlılık-Büyüme) yer almaktadır. Ayrıca Model 3'te Hazır Değerler Devir Hızı oranı negatif işaretli olarak bulunmaktadır. İşletmelerin sağladığı rekabet avantajı satış hacimlerini yükseltmiş ve Satışlar/Toplam Varlık oranını Model 3 (Karlılık-Büyüme) ve Model 5'te (Verimlilik-Büyüme) önemli bir bağımsız değişken haline getirmiştir. İşletmelerin 2007 yılında artan satışlarının ve ekonomik istikrarın, ticari alacakların vadelerini etkilediği ve satış politikalarının esnek bir yapı kazanmasında, müşterilerinin taleplerine göre vadeli satışlarının vadelerinin uzamasında önemli bir etken olduğu düşünülmektedir. Varlık devir hızı bir işletmenin karlılığını belirleyen önemli faktörlerden biridir. Bir işletme diğer koşullar aynı kalmak üzere, varlık devir hızını arttırmak yoluyla öz sermaye karlılığını yükseltebilir. Faaliyet etkinliği göstergesi olarak kullanılan varlık devir hızı Model 3 ve Model 5 için anlamlı bir değişken olarak analize girmiştir. Varlık devir hızı ile performans göstergeleri arasındaki pozitif ilişki işletmelerin yatırımlarının 2007 yılında artması nedeniyle ortaya çıkmıştır.

Son olarak, işletme performansını değerlendirmek amacıyla yapılan temel analizlerde* kullanılan “fiyat-kazanç” değerinin “aktif karlılığı” ile birlikte modellenmeye çalışıldığı Model 4'te, pazar payı değişkeninin de pozitif bir etki yarattığı gözlenmiştir. Demir ve Kocabıyık (2008)²⁸⁴ yaptıkları çalışmada bireysel yatırımcıların hisse senedi yatırımlarını yaparken dikkat ettikleri

²⁸³ Hüseyin Yılmaz, a.g.e., s. 484.

* “Temel Analiz: Halka açıklanan bilgilerden yararlanarak bir firmanın değerinin araştırılmasıdır.” Bkz. Mehmet Baha Karan, a.g.e., s. 460.

²⁸⁴ Yusuf Demir, Turan Kocabıyık, “Türkiye’de Kurumsal Yatırımcılar ve Kurumsal Yatırımcıların Portföy Oluşturulurken Gösterdikleri Davranış Şekillerine İlişkin Bir Araştırma”, **Marmara Üniversitesi İ.İ.B.F. Dergisi**, Cilt. XXIV, Sayı. 1, 2008, ss. 181–197.

faktörleri anket yolu ile analiz etmişlerdir. Yatırımcıların işletmelere yatırım yaparken işletmelerin üretim teknolojisi, pazar payı, esas faaliyet karı, dönem net karı, tanınmışlık, kar payı, işlem hacmi gibi faktörlerden etkilendikleri ortaya koymuşlardır. Çalışmada performans göstergelerinden biri olarak ele alınan fiyat-kazanç oranının varlığının, pazar payının açıklayıcı değişken olarak Model 4'te yer almasına neden olduğunu düşündürmektedir.

Çalışmanın sonuçlarında elde edilen Model 4 incelendiğinde pazar payının açıklayıcı değişken olarak devreye girmesi, piyasa göstergelerinden fiyat-kazanç oranında yani sadece Model 4'te bulunmasını açıklayabilir.

Özellikle Türkiye gibi gelişmekte olan ülkelerde yatırımcılar, ortaklar ve işletme yöneticileri açısından işletmelerin performanslarının sürekliliğinin sağlanmasının önemi artmakta ve yapılan araştırmalarla da desteklenmektedir. Hem işletmeler açısından yatırımcı ve ortakların kararlarında, hem de ülke ekonomilerinin verimliliğinin ve etkinliğinin belirlenmesinde işletme performanslarını etkileyen faktörlerin en iyi şekilde belirlenmesi her geçen yıl daha da önem kazanmaktadır.

Bu çalışmada 2007 yılına ait imalat sanayi sektöründe yer alan işletmelerin Lojistik Regresyon ve Diskriminant Analizi kullanılarak performanslarını etkileyen faktörler belirlenmeye çalışılmıştır. Sonuçların benzer çalışmalar ve teori ile tutarlı olduğu, işletmelerin performanslarının finansal yapı, faaliyet, sermaye yoğunluğu ve likidite oranları ile daha iyi açıklandığı gözlenmiştir. Bunun yanı sıra ülkelerin ekonomik politikalarıyla beraber işletmelerin performansı değerlendirildiğinde yabancı sermaye ve pazar payı değişkenlerinin model de yer aldığı görülmektedir. Çalışma sadece 2007 yılıyla sınırlı kalmayarak başka yıllar içinde yapılarak sonuçların karşılaştırılması veya daha uzun bir dönemi içerecek şekilde Panel Lojistik Regresyon Analizi kullanılarak işletme performansını etkileyen faktörlerin zaman içerisindeki değişimlerini de incelemek üzere genişletilebilir.

KAYNAKÇA

- Agresti, Alan: **An Introduction to Categorical Data Analysis**, John Wiley & Sons, 1996, p. 72.
- Akaike, H.: “Information Theory and an Extension of the Maximum Likelihood Principle”, **Perspective in Statistics Selected Paper of Hirotugu Akaike**, Springer, 1998, p. 209.
- Akal, Z. : **İmalatçı Kamu Kuruluşlarında İşletmelerarası Toplam Performans, Verimlilik, Karlılık ve Maliyet Karşılaştırması**, Ankara, MPM: 538, 1994, s. 53.
- Akdoğan, N., Tenker, N. : **Finansal Tablolar ve Mali Analiz Teknikleri**, 4. Baskı İstanbul, 1994, s. 452.
- Akgüç, Ö. : **Mali Tablolar Analizi**, Avcıol Basım-Yayın, İstanbul, Genişletilmiş 9. Bası, 1995, ss. 363–364.
- Akgüç, Ö. : **Finansal Yönetim**, Avcıol Basım, Yenilenmiş 7. Baskı, İstanbul 1998, s. 71.
- Alam, M. Khorshed, Rao, M. Bhaskara, Cheng, Fu-Chih : “Sample Size Determination in Logistic Regression”, **The Indian Journal of Statistics**, Vol: 72-B, 2010, pp. 58–75.
- Albayrak, A. S. : **Uygulamalı Çok Değişkenli İstatistik Teknikler**, Asil Yayın Dağıtım Ltd. Şti., 2006, s. 65.
- Aldrich, John H., Nelson, Forrest D. : **Linear Probability Logit and Probit Models**, USA, Sage Publication, 1984, p. 41.
- Andersen, Erling B. : **Introduction to the Statistical Analysis of Categorical Data**, Springer, 1997, p. 30, 177.
- Aslan, M.H., Yıldırım, F. : “Endüstri Mühendisliği Yaşam Süresi Analizlerinde Yeniden Örnekleme Yöntemlerinin Kullanılması”, **2. Mühendislik ve Teknoloji Sempozyumu, 30 Nisan- 1 Mayıs 2009**, Çankaya Üniversitesi, s. 169.
- Bayyurt, N. : “İşletmelerde Performans Değerlendirmenin Önemi ve Performans Göstergeleri Arasındaki İlişkiler”, **Sosyal Siyaset Konferansları Dergisi**, Sayı 53, 2007/2 Prof. Dr. Haşmet Başar’a Armağan Özel Sayı, s. 580.
- Bilgin, C., Şahbaz, A. : “Türkiye’de Büyüme ve İhracat Arasındaki Nedensellik İlişkisi”, **Gaziantep Üniversitesi Sosyal Bilimler Dergisi**, 2009, (8).1, ss. 177–198.
- Bilodeau, M., Brenner, D. : **Theory of Multivariate Statistics**, Springer, 1999, p. 170.
- Buse, A. : “The Likelihood Ratio, Wald and Lagrange Multiplier Tests: An Expository Note”, **The American Statistician**, Vol: 36, No: 3, 1982, pp. 153–157.
- Bolat, Bilge A. : **Firma Performansını Etkileyen Faktörlerin Yapısal Eşitlik Modeli ile İncelenmesi**, Basılmış Doktora Tezi, İstanbul, 2010, ss. 32–39.
- Booth, Laurence, Aivazian, Varouj, Demirguc-Kunt, A., Maksimovic, V. : “Capital Structures in Developing Countries”, **The Journal of Finance**, The Journal of Finance, Vol: LVI, No. 1, February 2001, pp. 87–130.
- Chatterjee, Sampri, Hadi, Ali S., Price, Bertram : **Regression Analysis by Example**, Third Edition John Wiley & Sons, 2000, p.36.

- Chhibber, Pradeep K., Majumdar, Sumit K. : “Foreign Ownership and Profitability: Property Rights, Strategic Control and Corporate Performance in Indian Industry”, **The William Davidson Institute**, Working Paper Number 64, April 1997, pp. 1–39.
- Christensen, Ronald : **Log-Linear Models and Logistic Regression**, 2nd Edition, Springer, 1997, p. 297.
- Çiftçiöğlü, S., Karaaslan, C., Demir, H. : “Saving, Investment, Exports and Economic Growth in Turkey”, **Review of Social, Economic & Business**, Vol. 3–4, pp. 20–30.
- Copas J. B. : “Plotting p Against”, **Applied Statistics**, Vol: 32, 1983, pp. 25–31.
- Cox, D. R., Snell, E. J. : **Analysis of Binary Data**, Second Edition, Chapman & Hall, 1989, p. 21.
- Cramer, J. : “Predictive Performance of the Binary Logit Model in Unbalanced Samples”, **Journal of the Royal Statistical Society**, Series D, Vol: 88, 1990, pp. 85–94.
- Demir, Yusuf, Kocabıyık, Turan : “Türkiye’de Kurumsal Yatırımcılar ve Kurumsal Yatırımcıların Portföy Oluşturulurken Gösterdikleri Davranış Şekillerine İlişkin Bir Araştırma”, **Marmara Üniversitesi İ.İ.B.F. Dergisi**, Cilt. XXIV, Sayı. 1, 2008, ss. 181–197.
- Dinçer, Ö. : **Stratejik Yönetim ve İşletme Politikası**, İstanbul, 1992, s. 116.
- Dirican, A. : “Tanı Testi Performanslarının Değerlendirilmesi ve Kıyaslanması”, **Cerrahpaşa Tıp Dergisi**, Cilt 32, Sayı 1, 2001, ss. 25–30.
- Ebaid, I. El-Sayed : “The Impact of Capital-Structure Choice on Firm Performance: Empirical Evidence from Egypt”, **Journal of Risk Finance**, Vol. 10, Issue 5, 2009, pp. 477–487.
- Efron, Bradley : “Regression and ANOVA with Zero-One Data: Measures of Residual Variation”, **Journal of the American Statistical Association**, Vol: 73, 1978, pp. 113–121.
- Ellem, A. B. : **Multivariate Analysis (Stat501)**, January 10, 2008 (Çevirimiçi) <http://mcs.une.edu.au/~stat501/UnitNotes/unitnotes08.pdf>, 25 Ağustos 2011, p. 28.
- Eriotis, Nikolaos : “How Firm Characteristics Affect Capital Structure: An Empirical Study”, **Managerial Finance**, Vol: 33 No. 5, 2007, pp. 321–331.
- Fisher, R. A. : “The Use of Multiple Measurements in Taxonomic Problems”, **Annals of Eugenics**, 7(2), 1936, pp. 466–475.
- Freund, John E. : **Mathematical Statistics**, Fifth Edition, Prentice Hall International Editions, p. 224
- Friend, Irwin, Lang, Larry H. P. : “An Empirical Test of the Impact of Managerial Self-Interest on Corporate Capital Structure”, **The Journal of Finance**, Vol: WLIII, No. 2, June 1988, pp. 271–281.
- Gilbert, Ethel S. : “The Effect of Unequal Variance-Covariance Matrices on Fisher’s Linear Discriminant Function”, **Biometrics**, 1969, 25, pp. 505–515.
- Gill, Jeff : **Generalized Linear Models**, USA Sage University Paper, 2001, p. 2.
- Grabowski, Henry G., Mueller, Dennis C. : “Industrial Research and Development, Intangible Capital Stocks, and Firm Profit Rates”, **The Bell Journal of Economics**, Vol: 9, No. 2, 1978, pp. 328–343.
- Griliches, Z. : “Productivity, R&D and Basic Research at the Firm Level in the 1970’s”, **The American Economic Review**, 1986, pp. 141–153.
- Gujarati, Damodar N. : **Temel Ekonometri**, Literatür Yayıncılık, 1999.

- Hair, Joseph F., Black, William C., Babin, Barry J., Anderson, Rolph E., Tatham, Ronald L. : **Multivariate Data Analysis**, Sixth Edition, Pearson, 2006, p. 411.
- Hauck, W. W., Donner, A. : “Wald’s Test as Applied to Hypotheses in Logit Analysis”, **American Statistical Association**, Vol: 72, No: 360, 1977, pp. 851–853.
- Hausman, Jerry A., David A. : “A Conditional Probit Model for Qualitative Wise Choice: Discrete Decision recognizing Interdependence and Heterogeneous Preferences”, **Econometrica**, Vol: 46, No: 2, 1978, pp. 403–426.
- Hoel, P. G., Peterson, R. P. : “A Solution to the Problem of Optimum Classification”, **The Annals of Mathematical Statistics**, 1949, Vol: 20, No. 3 (Sep.), p. 433.
- Hosmer, David W., Jovanovic, B., Lemeshow, S. : “Best Subsets Logistic Regression”, **Biometrics**, Vol: 45, 1989, pp. 1265–1270.
- Hosmer, David W., Lemeshow, Stanley : **Applied Logistic Regression**, 2nd Edition, USA, John Wiley & Sons, 2000, p.1.
- Hsieh, F.Y., Bloch, Daniel A., Larsen, Michael D. : “A Simple Method of Sample Size Calculation for Linear and Logistic Regression”, **Statistics in Medicine**, Vol: 17, 1998, pp. 1623–1634.
- Huberty, Carl J. : **Applied Discriminant Analysis**, John Wiley & Sons INC., 1994, p. 43.
- Johnson, Richard A., Wichern Dean W. : **Applied Multivariate Statistical Analysis**, Third Edition, Prentice Hall, New Jersey, 1992, p. 133.
- Karagöz, M., Demirgil, H. : “Firma Büyüme Performansını Etkileyen Faktörler”, **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi**, 2009, C. 14, S.2, s. 48.
- Karan, M. B. : **Yatırım Analizi ve Portföy Yönetimi**, Gazi Kitabevi, Ankara, 2001, s. 356.
- Kay, R., Little, S. : “Transformation of the Explanatory Variables in the Logistic Regression Model for Binary Data”, **Biometrica**, Vol: 74, 1987, pp. 495–501.
- Kayo, Eduardo K., Kimura, Herbert : “Hierarchical Determinants of Capital Structure”, **Journal of Banking & Finance**, 35, 2011, pp. 358–371.
- Kester, W. Carl : “Capital and Ownership Structure: A Comparison of United States and Japanese Manufacturing Corporations”, **Financial Management in Japan**, Spring 1986, pp. 5–16.
- Kurnaz, N. : “Finansal (Mali) Tablolar Analizi”, (Çevrimiçi) <http://www.niyazikurnaz.net>, 12 Temmuz 2011.
- Lachenbruch, P. A. : “An Almost Unbiased Method of Obtaining Confidence Intervals for the Probability of Misclassification in Discriminant Analysis”, **Biometrics**, Vol. 23, No. 4, 1967, pp. 639–645.
- Lachenbruch, Peter A., Sneeringer, C., Revo, L. T. : “Robustness of the Linear and Quadratic Discriminant Function to Certain Types of Non-Normality”, **Communications in Statistics**, 1973, Vol: 1, pp. 39–56.
- Liao, Tim Futing : **Interpreting Probability Models Logit Probit and Other Generalized Linear Models**, Sage University Paper, 1994, pp. 1–2.
- Long, J. Scott : **Regression Models for Categorical and Limited Dependent Variables**, USA, Sage Publication, 1997, p. 36.
- Long, J. Scott, Cheng, Simon : **The Handbook of Data Analysis**, Edited by Hardy and Bryman pp. 261–262.
- Long, Michael S., Malitz, Ilen B. : “Investment Patterns and Financial Leverage”, **National Bureau of Economic Research**, Ed. Benjamin M. Friedman, Chapter 9, 1985, pp. 325–352.

- Longford, Nicholas T. : **Handbook of Statistical Modeling for the Social and Behavioral Sciences**, Plenum Press, Edited by Gerhard Arminger, Clifford C. Clogg and Michael E. Sobel, 1995, p. 551.
- Maddala, G. S. : **Limited-Dependent and Qualitative Variables in Econometrics**, Cambridge University Press, 1983, p.39.
- Magee, Lonnie : “ R^2 Measures Based on Wald and Likelihood Ratio Joint Significance Tests”, **The American Statistician**, Vol: 44, 1990, pp. 250–253.
- Marks, S., Dunn, O. J. : “Discriminant Functions When Covariance Matrices are Unequal”, **Journal of the American Statistical Association**, 1974, 69, pp. 555–559.
- McCullagh, P. : “The Conditional Distribution of Goodness-of-fit Statistics for Discrete Data”, **Journal of the American Statistical Association**, Vol: 81, No: 393, 1986, pp. 104–107.
- McCullagh, P., Nelder, J. A. : **Generalized Linear Models**, 2nd Edition, Chapman & Hall, 1989, p. 30.
- McKay, R. J., Campbell, N. A. : **Variable Selection Techniques in Discriminant Analysis I. Description**, British Journal of Mathematics Statistical Psychology, 35, 1982a, pp. 1–29.
- Mickey J., Greenland S. : “A Study of the Impact of Confounder-selection Criteria on Effect Estimation”, **American Journal of Epidemiology**, 129, pp. 125–137.
- Moschos, Demetrios : “Export Expansion, Growth and the Level of Economic Development: An Empirical Analysis”, **Journal of Development Economics**, Vol: 30 Issue 1, 1989, p. 93.
- Morbey, Graham K., Reithner, Robert M. : “How Research and Development Affects Sales Growth, Productivity and Profitability”, **Research-Technology Management**, Vol: 33, 3, 1990, pp. 11–15.
- Mucuk, İ. : **Modern İşletmecilik**, 7. Baskı, Türkmen Kitapevi, İstanbul, 1997, ss. 98–99.
- Myers, Stewart C. : “Capital Structure”, **Journal of Economic Perspectives**, Vol: 15, Number 2, Spring 2001, pp. 81–102.
- Nelder, J. A., Wedderburn, R. W. M. : “Generalized Linear Models”, **Journal of the Royal Statistical Society Series A**, Vol: 135, No: 3, 1972, pp. 370–384.
- Newbold, Paul, Carlson, William L., Thorne, Betty : **Statistics for Business and Economics**, Sixth edition, Prentice Hall 2007, p.539.
- Obuchowski, N. A. : “ROC Analysis”, **American Journal of Roentgenology**, 184, 2005, pp. 364–372.
- Onaolapo, Adekunle A., Kajola, Sunday O. : “Capital Structure and Firm Performance: Evidence from Nigeria”, **European Journal of Economics, Finance and Administrative Sciences**, Issue 25, 2010, pp. 70–82.
- Orhunbilge, Neyran : **Tanımsal İstatistik Olasılık ve Olasılık Dağılımları**, Avcıol Basım, 2000, s. 192.
- Orhunbilge, Neyran : **Uygulamalı Regresyon ve Korelasyon Analizi**, İstanbul, 2002, s. 241.
- Orhunbilge, Neyran : **Çok Değişkenli İstatistik Yöntemler**, İstanbul Üniversitesi Basım ve Yayımevi Müdürlüğü, İstanbul, 2010, s. 191.
- Öğüdücü, Ş. G. : “Veri Madenciliği Farklı Sınıflandırma Yöntemleri”, (Çevirimiçi) http://web.itu.edu.tr/~sgunduz/courses/veri_maden 15 Ağustos 2011.
- Özdamar, Kazım : **Paket Programlar ile İstatistiksel Veri Analizi**, Cilt 2, Yenilenmiş 5. Baskı, 2004, s. 355.

- Pampel, Fred C. : **Logistic Regression A Primer**, USA, Sage Publication, 2000, pp. 54–55.
- Panas, Epaminondas, Wamvoukas, George : “Further Evidence on The Export-Led Growth Hypothesis”, **Applied Economics Letters**, 9, 2002, p.731.
- Pandya, Anil M., Rao, Narender V. : “Deversification and Firm Performance an Empirical Evaluation”, **Journal of Financial and Strategic Decisions**, V.II. N. 2, 1998, p. 67.
- Pedduzi, P. N., Concato, J., Kemper, E., Holford, T. R., Feinstein, A., : “A Simulation Study of the Number of Events per Variable in Logistic Regression Analysis”, **Journal of Clinical Epidemiology**, Vol: 49, 1996, pp. 1373–1379.
- Pregibon, D. : “Logistic Regression Diagnostic”, **Annals of Statistics**, Vol: 9, 1981, pp. 705–724.
- Raftery, Adrian E. : “Bayesian Model Selection Social Research”, **Sociological Methodoloji**, Vol: 25, 1995, pp. 111–165.
- Rao, C. R. : “The Utilization of Multiple Measurements in Problem of Biological Classification”, **Journal of the Royal Statistical Society, Series B** (Methodological), 1948, Vol: 10, No. 2, pp. 159–203.
- Narendar V. Rao, Al-Yahyaee, K. H. M., Syed, Lateef A. M. : “Capital Structure and Financial Performance: Evidence from Oman”, **Indian Journal of Economics and Business**, June 2007.
- Rencher, Alvin C., Schaafje, G. Bruce : **Linear Models in Statistics**, A John Willey & Sons, 2008, p. 514.
- Rothenberg, T. J. : “Hypothesis Testing in Linear Models When the Error Covariance Matrix is Nonscalar”, **Econometrica**, 1984, Vol: 52, pp. 827–842.
- Royston, P., Altman, D. G. : “Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling”, **Applied Statistics**, Vol: 43, 1994, pp. 429–467.
- Sauerbrei, W., Meier-Hirmer, C., Benner, A., Royston, P. : “Multivariable Regression Model Building by Using Fractional Polynomials: Description of SAS, STATA and R Programs”, **Computational Statistics & Data Analysis**, Vol: 50, 2006, p. 3466.
- Schmidt, P., Strauss, R. P. : “The Prediction of Occupation Using Multiple Logit Models”, **International Economic Review**, Vol: 16, No: 2, 1975, pp. 471–486.
- Sharma, Subhash : **Applied Multivariate Techniques**, John Wiley & Sons, Inc., 1996, p. 241.
- Silva, A. Pedro Durate, Stam, Antonie : **Reading and Understanding Multivariate Statistics**, Edited by Laurence G. Grimm, Paul R. Yarnold, American Psychological Association Washington, DC., 1994, p. 277.
- Smith, C. A. B. : “Some Examples of Discrimination”, **Annals of Eugenics**, 1947, 13, pp. 272–282.
- SPSS : (Çevirimiçi) http://support.spss.com/ProductsExt/SPSS/Documentation/Statistics/algorithms/14.0/app14_boxs_m.pdf, 10 Eylül 2011.
- Suiçmez, H. : “Türkiye’de Ekonomik Büyüme ve Verimlilik Artış Performansı Işığında Nasıl bir Kalkınma Politikası Benimsenmeli”, **Türkiye İşveren Sendikaları Konfederasyonu**, Ocak, 2007. (Çevirimiçi)
http://www.tisk.org.tr/isveren_sayfa.asp?yazi_id=1618&id=84, 11 Haziran 2011

- Szymanski, David M., Bharadwaj, Sundar G., Varadarjan, P. Rajan : “An Analysis of the Market Share-Profitability Relationship”, **The Journal of Marketing**, Vol: 57, No. 3, 1993, pp. 1–18.
- Tacq, Jacques : **Multivariate Analysis Techniques in Social Science Research**, Sage Publication, 1997, p. 248.
- Tatlıdil, Hüseyin : **Uygulamalı Çok Değişkenli İstatistiksel Analiz**, Cem Ofset Ltd. Şti., Ankara, 1996, s.293.
- Tekin, M. : “Sanayi İşletmelerinde Verimlilik ve Önemi”, **Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi**, Sayı: 1, 1992, s. 170.
- Tezeren, A. : **İmalat Sanayinde Verimliliği Etkileyen Faktörler**, 319 Nolu Milli Produktivite Merkezi Yayınları, Ankara, 1985.
- Titman, Sheridan, Wessels, Roberto : “The Determinants of Capital Structure Choice”, **The Journal of Finance**, Vol: XLIII, No. 1, March 1988, pp. 1–19.
- Tükenmez, M., vd. : **Finansal Yönetim**, Vizyon Eğitim ve Danışmanlık Ltd. Şti., İzmir 1999, ss. 413–414.
- Wald, John K. : “How Firm Characteristics Affect Capital Structure: An International Comparison”, **The Journal of Financial Research**, Vol: XXII, No. 2, Summer 1999, pp. 161–187.
- Wehrens R., Putter, H., Buydens, L.M.C. : “The Bootstrap: A Tutorial”, **Chemometrics and Intelligent Laboratory Systems**, Vol: 54, 2000, pp. 35–52.
- Whittemore, Alice S.
: “Sample Size for Logistic Regression with Small Response Probability”, **Journal of the American Statistical Association**, Vol: 76, 1981, pp. 27–32.
- Wilks, S.S. : **Certain Generalizations in the Analysis of Variance**, *Biometrika*, Vol. 24, No. 3/4, 1932, pp. 471–494.
- Wiwattanakantang, Y. : “An Empirical Study on the Determinants of the Capital Structure of Thai Firms”, **Pacific-Basin Finance Journal**, Vol: 7, 1999, pp. 371–403.
- Wright, Raymond E. : **Reading and Understanding Multivariate Statistics**, American Psychological Association, Edited By Laurence G. Grimm and Paul R. Yarnold, 1994, p. 220.
- Yılmaz, Hüseyin : “Finansal Verimlilik Artışının Finans Maliyetlerine Etkisini Belirlemeye Yönelik Bir İnceleme”, **V. Ulusal Üretim Araştırmaları Sempozyumu**, İstanbul Ticaret Üniversitesi, 25–27 Kasım 2005, s. 484.
: **İstanbul Sanayi Odası Dergisi**, “Türkiye’nin 500 Büyük Sanayi Kuruluşu 2007”, Özel Sayı, Sayı 509, Ağustos 2008, s. 165.
: **Sermaye Piyasası ve Borsa Temel Bilgiler Kılavuzu**, “İstanbul Menkul Kıymetler Borsası Yayınları”, 17. Basım, Temmuz 2002, s.141.
- Açık Gazete (Çevirimiçi) : “Yabancı Sermaye Nedir? Hazırlayan Faruk Eskioğlu”, <http://www.acikgazete.com/ozel-dosyalar/2009/03/24/yabanci-sermaye-nedir.htm?aid=27995>, 01 Temmuz 2011
- Bölüm–1 : “Niçin İhracat”, **100 Soruda Dış Ticaret, İhracat Geliştirme Etüd Merkezi**, Ekim, 2006, (Çevirimiçi)
http://www.dkib.org.tr/dosya/pratik_100soru.pdf, 10 Temmuz 2011.
- Bölüm–10 : “Teknolojik Yenilik ve Ekonomik Performans”, (Çevirimiçi)
http://www.inovasyon.org/pdf/blm10.pdf_01.07.2011.
(Çevirimiçi) : <http://notoku.com/pims-analizi/>, 12. Temmuz 2011
(Çevirimiçi) : <http://www.havvatunc.com/ticaret-acigi-buyuyor-ihracatin-ithalati-karsima-orani-kuculuyor-1107.html>, 10 Eylül 2011.

- (Çevirimiçi) : <http://www.bilgaz.net/dosyalar/OranAnalizi.pdf>, 10Ağustos 2011.
(Çevirimiçi) : <http://www.bacegrup.com/3-4-oran-analizi.ntml>, 10Temmuz 2011.
(Çevirimiçi) : “Yatırım Ortamı Değerlendirmesi Krizden Özel Sektör Öncülüğünde Büyümeye”. <http://www.hazine.gov.tr>, 07Temmuz 2011.

EKLER

EK.1 Karlılık Oranları

Kendall Uygunluk Katsayısı

Test Statistics

N	8
Kendall's W ^a	.802
Chi-square	930.582
df	145
Asymp. Sig.	.000

a. Kendall's Coefficient
of Concordance

Correlations

Spearman's rho	OzSermaye	Ekonomik Rant	AktifKar	FaalKar Marj	BrutKar Marj	NetKarMarj	VOKarMarj	Faiz, Karsilama
OzSermaye Correlation Coefficient	1.000	.823**	.907**	.743**	.462**	.852**	.840**	.747**
Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000
N	146	146	146	146	146	146	146	146
EkonomikR ant Correlation Coefficient	.823**	1.000	.896**	.767**	.530**	.832**	.832**	.672**
Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.000
N	146	146	146	146	146	146	146	146
AktifKar Correlation Coefficient	.907**	.896**	1.000	.810**	.542**	.953**	.945**	.805**
Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.000
N	146	146	146	146	146	146	146	146
FaalKarMarj Correlation Coefficient	.743**	.767**	.810**	1.000	.686**	.852**	.859**	.871**
Sig. (2-tailed)	.000	.000	.000		.000	.000	.000	.000

N	146	146	146	146	146	146	146	146	146	146	146
BrutKarMarj	.462**	.530**	.542**	.686**	1.000	.628**	.652**	.550**			
Correlation Coefficient											
Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
N	146	146	146	146	146	146	146	146	146	146	146
NetKarMarj	.852**	.832**	.953**	.852**	.628**	1.000	.991**	.809**			
Correlation Coefficient											
Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
N	146	146	146	146	146	146	146	146	146	146	146
VOKarMarj	.840**	.832**	.945**	.859**	.652**	.991**	1.000	.816**			
Correlation Coefficient											
Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
N	146	146	146	146	146	146	146	146	146	146	146
FaizKarsila	.747**	.672**	.805**	.871**	.550**	.809**	.816**	1.000			
Correlation Coefficient											
Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
N	146	146	146	146	146	146	146	146	146	146	146

** . Correlation is significant at the 0.01 level (2-tailed).

EK.2 Büyüme Oranları

Kendall Uygunluk Katsayısı

Test Statistics

N	4
Kendall's W ^a	.501
Chi-square	290.488
df	145
Asymp. Sig.	.000

a. Kendall's Coefficient
of Concordance

Spearman Sıra Korelasyon Değerleri

Correlations

Spearman's rho		Satartsireel	Karartsireel	Ozsermayearts	Varlikarts
Satartsireel	Correlation Coefficient	1.000	.208*	.280**	.374**
	Sig. (2-tailed)	.	.012	.001	.000
	N	146	146	146	146
Karartsireel	Correlation Coefficient	.208*	1.000	.275**	.327**
	Sig. (2-tailed)	.012	.	.001	.000
	N	146	146	146	146
Ozsermayearts	Correlation Coefficient	.280**	.275**	1.000	.544**
	Sig. (2-tailed)	.001	.001	.	.000
	N	146	146	146	146
Varlikarts	Correlation Coefficient	.374**	.327**	.544**	1.000
	Sig. (2-tailed)	.000	.000	.000	.
	N	146	146	146	146

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

EK.3 Piyasa Performans Değerleri

Kendall Uygunluk Katsayısı

Test Statistics

N	3
Kendall's W ^a	.485
Chi-square	210.915
df	145
Asymp. Sig.	.000

a. Kendall's Coefficient
of Concordance

Spearman Sıra Korelasyon Değerleri

Correlations

Spearman's rho		Hissebaskar	Fiyatkazanc	PiyasaDefDeg
Hissebaskar	Correlation Coefficient	1.000	.319**	.111
	Sig. (2-tailed)	.	.000	.182
	N	146	146	146
Fiyatkazanc	Correlation Coefficient	.319**	1.000	.253**
	Sig. (2-tailed)	.000	.	.002
	N	146	146	146
PiyasaDefDeg	Correlation Coefficient	.111	.253**	1.000
	Sig. (2-tailed)	.182	.002	.
	N	146	146	146

** . Correlation is significant at the 0.01 level (2-tailed).

EK.4 Model 1 için Diskriminant Analizi Sonuçları

Eşit Varyanslılık Varsayımını test etmek için Enter model kullanılarak hesaplanan Box's M test sonucu,

Log Determinants

y	Rank	Log Determinant
0	24	36,640
1	24	34,862
Pooled within-groups	24	59,410

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

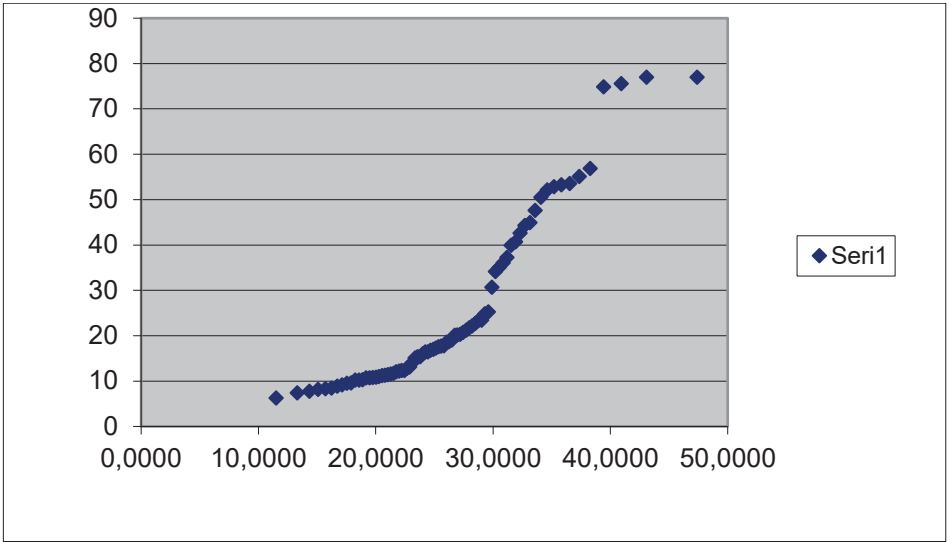
Test Results

Box's M	1835,091
F	Approx. 3,937
	df1 300
	df2 13603,759
	Sig. ,000

Tests null hypothesis of equal population covariance matrices.

Çoklu Normallik Varsayımı için yapılan testler,

Sıralanmış Mahalanobis uzaklıkları ile χ^2 değerlerinin serpilme diyagramı,



Ek.4'ün Devamı

Analyze > Descriptive > Explore

y	Kolmogorov-Smirnov ^a		Shapiro-Wilk	
	Statistic	df	Statistic	Sig.
Likidite1	,285	32	,582	,000
1	,215	47	,737	,000
Likidite2	,287	32	,580	,000
1	,207	47	,700	,000
Likidite3	,308	32	,520	,000
1	,278	47	,519	,000
Likidite4	,177	32	,875	,002
1	,113	47	,936	,012
Likidite5	,135	32	,921	,023
1	,160	47	,912	,002
FinYap1	,124	32	,976	,688
1	,108	47	,968	,221

FinYap2	0	,084	32	,200*	,975	32	,653
	1	,128	47	,051	,933	47	,009
FinYap3	0	,447	32	,000	,234	32	,000
	1	,141	47	,020	,885	47	,000
FinYap4	0	,228	32	,000	,719	32	,000
	1	,104	47	,200*	,976	47	,445
FinYap5	0	,110	32	,200*	,956	32	,218
	1	,091	47	,200*	,966	47	,182
Faaliyet1	0	,355	32	,000	,416	32	,000
	1	,471	47	,000	,161	47	,000
Faaliyet2	0	,388	32	,000	,297	32	,000
	1	,210	47	,000	,734	47	,000
Faaliyet3	0	,453	32	,000	,225	32	,000
	1	,235	47	,000	,632	47	,000
Faaliyet4	0	,171	32	,018	,873	32	,001
	1	,132	47	,041	,919	47	,003
Faaliyet5	0	,244	32	,000	,765	32	,000
	1	,145	47	,015	,879	47	,000
Faaliyet6	0	,180	32	,010	,859	32	,001

	1		,154	47	,007	,866	47	,000
Faaliyet7	0		,229	32	,000	,785	32	,000
	1		,120	47	,086	,910	47	,002
SermayeYog1	0		,219	32	,000	,766	32	,000
	1		,197	47	,000	,760	47	,000
SermayeYog2	0		,226	32	,000	,818	32	,000
	1		,186	47	,000	,856	47	,000
İhracat	0		,404	32	,000	,399	32	,000
	1		,404	47	,000	,349	47	,000
YabSermaye	0		,512	32	,000	,404	32	,000
	1		,323	47	,000	,713	47	,000
Yenilik	0		,269	32	,000	,701	32	,000
	1		,530	47	,000	,135	47	,000
PazarPay	0		,203	32	,002	,750	32	,000
	1		,285	47	,000	,554	47	,000

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Ek.4'ün Devamı

Discriminant

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
Likidite1	,978	1,759	1	77	,189
Likidite2	,977	1,836	1	77	,179
Likidite3	,943	4,697	1	77	,033
Likidite4	,967	2,631	1	77	,109
Likidite5	,920	6,668	1	77	,012
FinYap1	,820	16,956	1	77	,000
FinYap2	,815	17,437	1	77	,000
FinYap3	,974	2,032	1	77	,158
FinYap4	,974	2,071	1	77	,154
FinYap5	,979	1,675	1	77	,199
Faaliyet1	,991	,669	1	77	,416
Faaliyet2	,990	,809	1	77	,371
Faaliyet3	,993	,577	1	77	,450
Faaliyet4	1,000	,000	1	77	,999
Faaliyet5	,970	2,401	1	77	,125
Faaliyet6	,917	6,963	1	77	,010
Faaliyet7	,994	,499	1	77	,482
SermayeYog1	,861	12,480	1	77	,001
SermayeYog2	,828	16,017	1	77	,000
İhracat	,995	,349	1	77	,557
YabSermaye	,940	4,936	1	77	,029
Yenilik	,991	,669	1	77	,416
PazarPay	,966	2,726	1	77	,103
Buyukluk	,901	8,507	1	77	,005

Ek.4'ün Devamı
Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
dimension 0	,874 ^a	100,0	100,0	,683

a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
dimension 0 1	,534	46,462	6	,000

Canonical Discriminant

Function Coefficients

	Function
	1
Likidite5	4,206
FinYap2	3,786
SermayeYog1	,000
SermayeYog2	,000
YabSermaye	-,015
Buyukluk	-1,444
(Constant)	-,871

Unstandardized coefficients

Classification Statistics

Prior Probabilities for Groups

y	Prior	Cases Used in Analysis	
		Unweighted	Weighted
0	,405	32	32,000
1	,595	47	47,000
Total	1,000	79	79,000

Classification Results^a

y			Predicted Group Membership		Total
			0	1	
Original	Count	0	23	9	32
		1	6	41	47
	%	0	71,9	28,1	100,0
		1	12,8	87,2	100,0

a. 81,0% of original grouped cases correctly classified.

EK.5 Model 1 için Lojistik Regresyon Analizi Sonuçları**STATA Çıktısı**

```
. sw logit y Likidite1 Likidite2 Likidite3 Likidite4 Likidite5 FinYap1 FinYap2
FinYap3 FinYap4 FinYap5 Faaliyet1 Faaliyet2 Faaliyet3 Faaliyet4 Faaliyet5
Faaliyet6 Faaliyet7 SermayeYog1 SermayeYog2 Ihracat YabSermaye Yenilik
PazarPay Buyukluk, forward pe(.1) pr(.2)
```

```
begin with empty model
```

```
p = 0.0005 < 0.1000 adding FinYap1
p = 0.0032 < 0.1000 adding SermayeYog2
p = 0.0077 < 0.1000 adding Buyukluk
p = 0.0583 < 0.1000 adding YabSermaye
p = 0.0989 < 0.1000 adding Likidite5
p = 0.0427 < 0.1000 adding FinYap5
p = 0.0349 < 0.1000 adding FinYap2
p = 0.5408 >= 0.2000 removing FinYap1
```

```
Logistic regression          Number of obs =      79
LR chi2(6) = 56.84           Prob > chi2 = 0.0000
Log likelihood = -24.904914   Pseudo R2 = 0.5330
```

```

y |   Coef.   Std. Err.   z   P>|z|   [95% Conf. Interval]
-----+-----
FinYap2 | -9.932728   3.326845   -2.99   0.003  -16.45322  -3.412232
SermayeYog2 | 5.45e-06   1.64e-06    3.32   0.001   2.23e-06   8.67e-06
Buyukluk | 3.737337   1.389887    2.69   0.007   1.013209   6.461466
YabSermaye | .0415742   .0161685    2.57   0.010   .0098845   .0732638
Likidite5 | -10.85362   4.46874    -2.43   0.015  -19.61219  -2.095047
FinYap5 | 8.50793    3.335086    2.55   0.011   1.971282   15.04458
_cons | -1.688702   1.251986   -1.35   0.177   -4.14255   .7651456
```

Ek.5'in Devamı

SPSS Çıktısı

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	56.842	6	.000
	Block	56.842	6	.000
	Model	56.842	6	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	49.810 ^a	.513	.693

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	2.523	8	.961

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)	
Step 1 ^a	FinYap2	-9.933	3.327	8.914	1	.003	.000
	SermayeYog2	.000	.000	11.026	1	.001	1.000
	Buyukluk	3.737	1.390	7.230	1	.007	41.986
	YabSermaye	.042	.016	6.612	1	.010	1.042
	Likidite5	-10.854	4.469	5.899	1	.015	.000
	FinYap5	8.508	3.335	6.508	1	.011	4953.902
	Constant	-1.689	1.252	1.819	1	.177	.185

a. Variable(s) entered on step 1: FinYap2, SermayeYog2, Buyukluk, YabSermaye, Likidite5, FinYap5.

Classification Table^a

Observed			Predicted		
			y		Percentage Correct
	0	1	0	1	
Step 1	y	0	25	7	78.1
		1	5	42	89.4
Overall Percentage					84.8

a. The cut value is .500

EK.6 Kesirli Polinomlar Çıktısı**(FinYap2 için)**

```
fracpoly logit y FinYap2 SermayeYog2 Buyukluk YabSermaye Likidite5
FinYap5, compare
```

```
.....
-> gen double IFinY__1 = FinYap2^3-.0277117306 if e(sample)
```

```
-> gen double IFinY__2 = FinYap2^3*ln(FinYap2)+.0331238266 if e(sample)
```

```
Iteration 0: log likelihood = -53.325894
```

```
Iteration 1: log likelihood = -31.470079
```

```
Iteration 2: log likelihood = -26.261177
```

```
Iteration 3: log likelihood = -23.069584
```

```
Iteration 4: log likelihood = -21.620449
```

```
Iteration 5: log likelihood = -21.326689
```

```
Iteration 6: log likelihood = -21.30554
```

```
Iteration 7: log likelihood = -21.305182
```

```
Iteration 8: log likelihood = -21.305182
```

```
Logistic regression                Number of obs =      79
```

```
LR chi2(7) =      64.04                Prob > chi2 =    0.0000
```

```
Log likelihood = -21.305182          Pseudo R2 =    0.6005
```

```
      y |   Coef.   Std. Err.   z   P>|z|   [95% Conf. Interval]
```

```
-----+-----
```

IFinY__1	-51.9244	29.92416	-1.74	0.083	-110.5747	6.725868
IFinY__2	-46.33207	50.87512	-0.91	0.362	-146.0455	53.38134
ISerm__1	8.77e-06	2.87e-06	3.06	0.002	3.14e-06	.0000144
Buyukluk	3.613839	1.518983	2.38	0.017	.6366866	6.590991
IYabS__1	.0677053	.0248941	2.72	0.007	.0189136	.1164969
ILiki__1	-8.967307	4.247005	-2.11	0.035	-17.29128	-.6433298
IFinYa__1	8.770861	3.452947	2.54	0.011	2.003209	15.53851
_cons	2.305496	.9231263	2.50	0.013	.4962014	4.11479

```
-----
```

Fractional polynomial model comparisons:

```
FinYap2      df      Deviance      Gain P(term) Powers
```

```
-----
```

Not in model	0	63.389	--	--	
Linear	1	49.810	0.000	0.000	1
m = 1	2	43.716	6.094	0.014	3
m = 2	4	42.610	7.199	0.575	3 3

Ek.6'nın Devamı**(Buyukluk için)**

fracpoly logit y Buyukluk YabSermaye Likidite5 FinYap5 FinYap2
SermayeYog2, compare

.....
-> gen double IBuyu__1 = X^-2 if e(sample)
-> gen double IBuyu__2 = X^-2*ln(X) if e(sample)
(where: X = (Buyukluk+1))

note: IBuyu__2 dropped due to collinearity

Iteration 0: log likelihood = -53.325894

Iteration 1: log likelihood = -32.455609

Iteration 2: log likelihood = -27.261292

Iteration 3: log likelihood = -25.209538

Iteration 4: log likelihood = -24.912528

Iteration 5: log likelihood = -24.904921

Iteration 6: log likelihood = -24.904914

Logistic regression	Number of obs =	79
LR chi2(6) =	56.84	Prob > chi2 = 0.0000
Log likelihood = -24.904914	Pseudo R2 =	0.5330

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
IBuyu__1	-4.983117	1.853182	-2.69	0.007	-8.615287	-1.350946
IYabS__1	.0415742	.0161685	2.57	0.010	.0098845	.0732638
ILiki__1	-10.85362	4.46874	-2.43	0.015	-19.61219	-2.095047
IFinY__1	8.50793	3.335086	2.55	0.011	1.971282	15.04458
IFinYa__1	-9.932728	3.326845	-2.99	0.003	-16.45322	-3.412232
ISerm__1	5.45e-06	1.64e-06	3.32	0.001	2.23e-06	8.67e-06
_cons	5.586161	1.853692	3.01	0.003	1.95299	9.219331

Fractional polynomial model comparisons:

Buyukluk	df	Deviance	Gain	P(term)	Powers
Not in model	0	60.355	--	--	
Linear	1	49.810	0.000	0.001	1
m = 1	2	49.810	0.000	1.000	-2
m = 2	4	49.810	0.000	1.000	-2 -2

Ek.6'nın Devamı**(YabSermaye için)**

fracpoly logit y YabSermaye Likidite5 FinYap5 FinYap2 SermayeYog2
Buyukluk, compare

.....
-> gen double IYabS__1 = X^-2-.4906096535 if e(sample)

-> gen double IYabS__2 = X^3-2.910019277 if e(sample)

(where: X = (YabSermaye+.0599999986588955)/10)

Iteration 0: log likelihood = -53.325894

Iteration 1: log likelihood = -31.54844

Iteration 2: log likelihood = -26.249092

Iteration 3: log likelihood = -24.007801

Iteration 4: log likelihood = -23.61483

Iteration 5: log likelihood = -23.599609

Iteration 6: log likelihood = -23.599576

Logistic regression	Number of obs =	79
LR chi2(7) =	59.45	Prob > chi2 = 0.0000
Log likelihood = -23.599576	Pseudo R2 =	0.5574

y	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
IYabS__1	-.0000794	.0000379	-2.10	0.036	-.0001536	-5.17e-06
IYabS__2	.0021787	.0024727	0.88	0.378	-.0026677	.0070252
ILiki__1	-11.41502	4.780794	-2.39	0.017	-20.7852	-2.044838
IFinY__1	9.133216	3.585204	2.55	0.011	2.106344	16.16009
IFinYa__1	-10.5042	3.568437	-2.94	0.003	-17.49821	-3.510192
ISerm__1	5.72e-06	1.73e-06	3.31	0.001	2.33e-06	9.12e-06
Buyukluk	3.798341	1.469575	2.58	0.010	.9180261	6.678655
_cons	2.054038	.9676219	2.12	0.034	.1575337	3.950542

Fractional polynomial model comparisons:

YabSermaye	df	Deviance	Gain	P(term)	Powers
Not in model	0	58.557	--	--	
Linear	1	49.810	0.000	0.003	1
m = 1	2	48.102	1.708	0.191	-2
m = 2	4	47.199	2.611	0.637	-2 3

Ek.6'nın Devamı

ROC curve çizmek için Stata Kodu ve Sonuçları

lroc

Logistic model for y

number of observations = 79

area under ROC curve = 0.9402

ÖZGEÇMİŞ

Serap Şahin ilk, orta ve lise öğrenimini Ankara’da sırası ile 13 Ekim İlköğretim, Mehmet Akif Ersoy ve Aydınlikevler Lisesin’de tamamlamıştır.

1993’de Hacettepe Üniversitesi, Fen Fakültesi, İstatistik bölümünü bitirmiştir. 1995’de Milli Eğitim Bakanlığı’nın Bursu ile Amerika’ya gitmiştir. 1997 yılında Virginia Tech Üniversitesi İstatistik Bölümünden yüksek lisans derecesini almıştır.

1998 – 2004 yılları arasında Kırıkkale Üniversitesin’de araştırma görevlisi olarak çalışmıştır. 2005 – 2011 yılları arasında İstanbul Üniversitesi İşletme Fakültesi, Sayısal Yöntemler Anabilim Dalı’nda doktora tezini tamamlamıştır. 2012 yılından bu yana Kırıkkale Üniversitesi İktisadi ve İdari Bilimler Fakültesi, İşletme Bölümü, Sayısal Yöntemler Anabilim Dalı’nda Öğretim Üyesi olarak çalışmaya devam etmektedir.

**İKİ KATEGORİLİ
LOJİSTİK REGRESYON
ve DİSKRİMİNANT ANALİZİ**
İşletme Performans Tahmini Üzerine Bir Uygulama

Dr. Serap ŞAHİN

 **ÖZGÜR**
YAYINLARI

ISBN 978-975-447-705-4

9 789754 4477054