

173078

ZOOTEKNİDE KATEGORİK VERİLERİN LOG-LİNEAR  
MODELLER YARDIMIYLA ANALİZİ

SERAP DOĞAN

Ç.Ü.

FEN BİLİMLERİ ENSTİTÜSÜ  
ZOOTEKNİ ANABİLİM DALI

YÜKSEK LİSANS TEZİ

ADANA  
ARALIK-1985

Çukurova Üniversitesi Fen Bilimleri Enstitüsü Müdürlüğüne,

Bu çalışma, jürimiz tarafından Zootekni Anabilim Dalında YÜKSEK LİSANS Tezi olarak kabul edilmiştir.

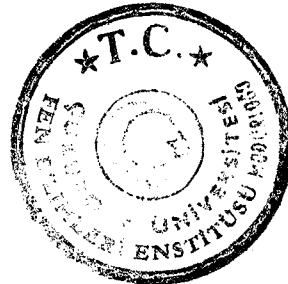
Başkan: Doç.Dr.Yüksel BEK

Üye : Prof.Dr.Erdoğan PEKEL

Üye : Prof.Dr.Fikri AKDENİZ

Kod no: 88

Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım.



Prof.Dr.Ural DİNÇ  
Enstitü Müdürü

## İÇİNDEKİLER

	<u>Sayfa</u>
ÇİZELGE LİSTESİ .....	IV
EKLER LİSTESİ .....	VI
ÖZ .....	VII
ABSTRACT .....	VIII
1. GİRİŞ .....	1
2. ÖNCEKİ ÇALIŞMALAR .....	4
3. MATERYAL VE METOD .....	9
3.1. İki Yönlü Tam Tablolar .....	9
3.1.1. 2x2 Boyutlu Tablolar .....	9
3.1.1.1. Bağımsızlık ve Toplanabi- lirlik .....	19
3.1.1.2. Bağımsızlık Altında Enbü- yük-Olabilirlik Tahminleri .....	24
3.1.1.3. Bağımsızlık Testleri .....	25
3.1.1.4. 2x2 Boyutlu Bir Tabloda İlişkinin Ölçülmesi .....	26
3.1.1.5. Süreklilik Düzeltilmesi ...	29
3.1.2. 2xs Boyutlu Tablolar .....	30
3.1.2.1. Enbüyük-Olabilirlik Tah- minleri .....	31
3.1.2.2. Khi-Kare Testleri .....	31
3.1.2.3. Düzeltilmiş Rezidüaller ..	32
3.1.3. rxs Boyutlu Tablolar .....	32
3.1.3.1. Toplanabilir Log-Linear Modeller .....	33
3.1.3.2. Örneklerin Karşılaştırılma- sı .....	35
3.1.3.3. Bağımsızlık ve Toplanabi- lirlik .....	36

3.1.3.4.	Doymuş Model .....	39
3.1.3.5.	Çapraz-Çarpım Oranları ...	44
3.1.3.6.	Khi-Kare Testleri .....	48
3.1.3.7.	Düzeltilmiş Rezidüyaller .	49
3.1.3.8.	Standardize Parametre Tah- minleri.....	51
3.2.	Üç Yönlü Tam Tablolar .....	51
3.2.1.	2x2x2 Boyutlu Tablolar .....	54
3.2.1.1.	Nisbi Risklerin Karşılaştı- rılması .....	55
3.2.1.2.	Üçlü-Faktör İnteraksiyonu Olmayan Model .....	62
3.2.2.	Üç Yönlü rxsxt Boyutlu Tablolar İçin Mümkün Hiyerarşik Modeller .....	67
3.2.3.	2x3x2 Boyutlu Tablolar .....	85
3.2.3.1.	Marginal ve Kısmi İlişki- ler .....	86
3.2.3.2.	Marginal İlişki İçin Khi- Kare İstatistikleri .....	89
3.2.3.3.	Şartlı Bağımsızlık .....	89
3.2.3.4.	Şartlı Bağımsızlık İçin Enbüyük-Olabilirlik Tahmin- leri ve Khi-Kare Testleri .	91
3.2.3.5.	Logaritmik Çapraz-Çarpım Oranları .....	93
3.2.3.6.	$A_h$ Verildiğinde $B_h$ ve $C_h$ Arasındaki Kısmi İlişkiler.	93
3.2.3.7.	Üçlü-Faktör İnteraksiyonu Olmayan Model .....	96
3.2.4.	Modellerin Seçimi .....	96
3.2.4.1.	Khi-Kare İstatistikleri ...	97
3.2.4.2.	Khi-Kare İstatistiklerinin Parçalanması .....	99

3.2.4.3. Rezidüyal Analiz .....	101
3.2.4.4. Standardize Parametre Tah- minleri .....	104
3.3. Çok-Yönlü Tam Tablolar .....	106
3.3.1. 2x2x3x2 Boyutlu Tablolar .....	107
3.3.1.1. Doymuş Modelle Tahmin .....	107
3.3.1.2. Doymuş Modellerde Enbüyük- Olabilirlik Tahminleri He- saplanması İçin Kapalı Form- da İfadeler .....	112
3.3.1.3. Uyum Testi .....	118
3.3.1.4. Modellerin Yorumu .....	120
3.3.2. 3x3x3x3 Boyutlu Tablolar .....	124
3.3.2.1. Doymuş Model .....	125
3.3.2.2. Khi-Kare İstatistikleri ...	129
3.3.2.3. Rezidüyal Analizi .....	131
3.4. Tam Tablolar İçin Enbüyük-Olabilirlik Tah- minleri .....	137
3.4.1. Üç ve Dört Boyutlu Tablolarda Direkt Tahminler .....	137
3.4.2. Direkt Tahminlerin Varlığını Tayin Etmek İçin Kurallar .....	141
4. ARAŞTIRMA BULGULARI .....	146
Uygulama 1 .....	146
Uygulama 2 .....	148
5. TARTIŞMA VE SONUÇLAR .....	153
ÖZET .....	160
SUMMARY .....	164
KAYNAKLAR .....	168
TEŞEKKÜR .....	171
ÖZGEÇMİŞ .....	172

ÇİZELGE LİSTESİ

Çizelge 1. İkiye İkilik Olasılık Tablosu .....	10
Çizelge 2. İkiye İkilik Gözlenen Değerler Tablosu .	13
Çizelge 3. İki Tane İki Şıklı Değişken İçin Doymuş Modelin $\lambda$ -Terimlerinin İşareti .....	22
Çizelge 4. 2xs Boyutlu Gözlenen Değerler Tablosu ..	30
Çizelge 5. rxs Boyutlu Gözlenen Değerler Tablosu ..	33
Çizelge 6. 2x2x2 Boyutlu Gözlenen Değerler Tablosu.	54
Çizelge 7. Üç Yönlü Tablolar İçin Doymuş Modelin $\lambda$ - Terimlerine Ait Katsayılar .....	63
Çizelge 8. 2x3x2 Boyutlu Gözlenen Değerler Tablosu.	86
Çizelge 9. Üç-Yönlü Tablolar İçin Hiyerarşik Log-Linear' Modellerde Khi-Kare Testlerine Ait Serbestlik Dereceleri .....	99
Çizelge 10. Hiyerarşik Modeller İçin $L^2$ İstatistikleri .....	100
Çizelge 11. Üç-Yönlü Tablolarda Hiyerarşik Modeller İçin Düzeltilmiş Rezidüyaller .....	102
Çizelge 12. Model 1 - Model 6 İçin Beklenen Değerler İçin Kesin İfadeler .....	128
Çizelge 13. Dört Boyutlu Tablolarda Mümkün Direkt Tahmin Tipleri .....	140
Çizelge 14. Dört Boyutlu Tablolar İçin Hiyerarşik Modeller .....	143
Çizelge 15. Ç.Ü.Z.F. Araştırma Çiftliğinde 3 Yılda Doğan Kuzuların Ana Yaşı (2,3,4). Cinsiyet (Erkek-Dişi), Doğum Şekli (Tek, İkiz)'ne Göre Karşılıklı Sınıflanmasıyla Oluşturulan Olabilirlik Tablosu ...	146
Çizelge 16. Çizelge 15 'deki Veriler İçin Uydurulan Hiyerarşik Modeller ve Bunlara Ait Khi-Kare İstatistikleri .....	147

Çizelge 17. Ceylanpınar Tarım İşletmesinde Doğan Buzagıların Doğum Mevsimi, Ana Irkı, Cinsiyetine Göre Karşılıklı Sınıflan- masıyla Elde Edilen Olabilirlik Tablo- su .....	149
Çizelge 18. Çizelge 17 'deki Veriler İçin Uyduru- lan Hiyerarşik Modeller ve Bunlara Ait Khi-Kare İstatistikleri .....	150



EKLER LİSTESİ

EK I- Log-Linear Modeller İin Enbyk-Olabilirlik Tahmin-  
lerini Hesaplayan Program.





## ÖZ

Bu çalışmada, zooteknide sık karşılaşılan kesikli verilerin, çok boyutlu tablolar halinde ifade edilerek log-linear modeller yardımıyla analizi konusu ele alınmış ve bu tür yaklaşımın kesikli değişkenlerin bağımlılık yapısını araştırmada çok faydalı olabileceği gösterilmiştir. Daha çok, model uydurma, model seçimi, uyum sağlayan modellerin yorumu gibi konular üzerinde durularak, yöntem üç değişkenli-iki tabloya uygulanıp sonuçlar tartışılmıştır.



## ABSTRACT

In this study, log-linear approaches to analysis of categorical data in animal science is introduced with the aim of motivating animal scientists to consider using these methods. It is shown that the method of log-linear analysis to investigate the independence structure of categorical data is very useful. Herein the model fitting, selection of most subtle model, and interpretations of the fitted models are also presented. Two data sets involving three variables are analysed to demonstrate the method of analysis and to seek some important points in the interpretation of the results.

## 1. GİRİŞ

Son yıllarda kalitatif verilerin analizinde önemli kuramsal ve yöntemsel gelişmeler yapıldı. Son yıllarda kategorik verilerin analizi, çok ilgi çeken bir araştırma alanı oldu ve çok zengin bir literatüre sahiptir. Bununla beraber zootečniyle ilgili bilimsel yayınların tetkikinde, kalitatif verilerle ilgili daha yeni kuram ve yöntemlerin kullanımının çok az olduğu görülür.

Şimdiye kadar, zooteknide kesikli yani kategorik veriler ilgilenilmeyen etkiler olarak görülmüş ve kovaryant olarak bunların etkileri elemine edilmeye çalışılarak, esas ilgilenilen sürekli özelliğın etkisine ait elde edilen standardize veriler analiz edilmiştir. Kesikli özellikler arasındaki ilişkiler üzerinde detaylı şekilde durulmayarak, bunlar üzerinde şimdiye kadar yapılan analizler, genellikle basit  $x^2$  analizi ile homogenlik testleri şeklinde olmuştur.

Birçok uygulamalı bilim dallarında, araştırmalardan elde edilen verilerin belli özelliklere göre sınıflandırılmaları sık rastlanılan bir durumdur. Özellikle gözlemlerin sayım yoluyla elde edildiğı araştırmalarda sonuçlar bir veya daha çok boyutlu tablolar halinde ifade edilir. Çok değişkenli kalitatif verilerde her birey bir takım özelliklerle tanımlanır. Belli özellikler yönünden aynı tanımlamaya sahip tüm bireyler birer birer sayılır ve bu değerler sonuçta olabirlik tablosunun oluşumunda kullanılır. Bu tür tabloların boyutları cinsiyet, kan grubu, yaş gibi özellikleri temsil eden değişkenler tarafından belirlenir. Araştırmaların çoğunda esas amaç tablonun boyutlarını oluşturan değişkenlerin bağımlılık yapısını ortaya koymaktır.

Tablonun boyut sayısının iki olduğu durumlarda, değişkenlerin bağımlı olup olmadıkları khi-kare testlerine dayanan yöntemler uygulanarak incelenebilir. Ancak, boyut sayısı ikiden fazla ise bağımlılık yapısının incelenmesi daha karmaşık hale gelmektedir. Çok boyutlu tabloların sırasıyla ikişer boyutunun ele alınarak incelenmesi de, her defasında diğer boyutların etkisinin ihmal edilmesini gerektirir ve bu tip testlerde yüksek dereceden interaksiyonlara yer verilmediğinden yanıltıcı olabilmektedir.

Kesikli verilerin analizinde log-linear modeller ve genelleştirilmiş en küçük kareler yaklaşımı kullanılabilir. Log-linear analizi ile ilgili teorik esaslar BISHOP ve ark. (1975) ve HABERMAN (1978 a,b) gibi teorik istatistik kitaplarında yer almaktadır. Son yıllarda bu tip veriler için de varyans analiz metodlarına benzer (ANOVA) metodlar kullanılmaktadır (GOODMAN (1971)). Bu metodlar bağımlı ve bağımsız değişkenler arasındaki ilişkinin nedenini izah etmek için logit modeller kullanırlar. Kesikli verilerin analizinde log-linear modellere alternatif olan, genelleştirilmiş en küçük kareler yaklaşımıdır (GRIZZLE ve ark. (1969)). Bu metodlar, gerçek hücre olasılıklarının, deneme değişkenlerinin linear bir fonksiyonu gibi ifade edilebileceğini varsayarlar. Genel linear model yaklaşımı ile çeşitli modeller gözden geçirilir.

Genelleştirilmiş en küçük kareler yaklaşımı ile minimum khi-kare tahminleri elde edilirken, log-linear yaklaşımı en büyük - olabilirlik tahminleri kullanır. Büyük örnek teorisi bu iki metodun eşdeğer olduğunu verir. Çoğu problemlerde her iki kullanım da uygundur. Her ikisi de, birçok veri grubu için kullanılan diğer metodlardan daha fazla bilgi verir gözükmektedir.

Zooteknistlerin dikkatini bu yöntemler üzerine çek-

mek ve şimdiye kadar Türkiye'de zootekni sahasında kullanılmayan bu analiz tekniğini tanıtmak amacıyla kesikli veya kategorik verilerin analizi konusunda log-linear model yaklaşımı üzerinde durulacaktır.

Log-linear analizin amacı, olabirlik tablolarının tanımlanmasını yapmak ve değişkenler arasındaki ilişkiyi varyans analizine benzer yöntemle testetmektir. Log-linear model bazı özelliklerinden dolayı amacımıza uygun görünmektedir. Bu özellikler:

1- (p) olasılıkları yerine beklenen hücre değerleri, model tanımlamada kullanıldığında, modelin yalnız bir parametresi değiştirilir.

2- Model değişik örnekleme yöntemleri için uygundur.

3- Model terimlerinin hemen izah edilebilir olması, aritmetik ölçekle linear olan diğer modellerde bulunmaz.

Bilgisayar olanaklarının artması, çok boyutlu tabloların daha geçerli analizlerinin yapılmasını sağlayacak yöntemlerin geliştirilmesine yol açmıştır. Özellikle son 20 yıl içinde bu amaçla geliştirilen yöntemler temelde linear modellere dayandıklarından, regresyon analizinin birçok olanaklarının bu alanda kullanılması mümkün olmuştur.

## 2. ÖNCEKİ ÇALIŞMALAR

Kalitatif verilerle çok sık karşılaşılmasına rağmen, 1960'a kadar bu verilerin analizleri için kullanılan istatistiksel metodlar oldukça ilkel kalmıştır. 1960'lardan sonra, log-linear modellerin geliştirilmesi, kalitatif verilerin ayrıntılı olarak analizinde çok hızlı bir gelişmeye neden olmuştur. Log-linear modellerdeki gelişmenin çoğu Goodman ve Frederick Mosteller gibi istatistikçiler tarafından yapılmıştır; buna rağmen birçok sosyal bilimci bu istatistiksel model sınıfıyla ilgili bazı bilgileri geliştirmişlerdir.

GOODMAN (1971a), çok boyutlu olabilirlik tablolarının analizi konulu bir çalışmada, çoklu sınıflandırmada model uydurma için basamaklı işlemleri ve direk tahmin metodlarını tartışmaktadır.

GOODMAN (1971 b), çok boyutlu olabilirlik tablolarında beklenen frekans tahmini ve marginal olasılık tablolarının analizi, khi-karenin parçalanışı konularını izah ediyor.

GRIZZLE ve WILLIAMS (1972), çok yönlü olabilirlik tablolarındaki bağımsızlığı test etmek için faktöryel deneylerin analizi ve marginal bağımsızlık testleri arasındaki uygunluğa dayanan istatistiksel metodlar geliştirmişlerdir. Bu ilişki olasılık tablolarının yorumunu kolaylaştırır ve olasılıklardan bazıları üzerine sıfır kısıtlamasının konmuş olduğu tablolar için basitleştirilmiş testlerin kullanılmasını sağlar. Linear model yaklaşımı ise olasılık tahminlerinin daha kolay hesaplanmasında kullanılır. Bu söz konusu çalışmada, çok yönlü tablo haline getirilmiş çok şıklı verilere log-linear modeller uydurma metodları tanımlanarak, anlatılan metodlarla linear modellerle yapı-

lan sürekli verilerin analizi arasındaki benzerlik ele alınıyor ve iyi bilinen marginal bağımsızlık testleriyle faktöriyel deneme analizi ile ilgili kesin testler arasındaki ilişki tayin ediliyor. BHAPKAR ve KOCH (1968 a, b) ile GOODMAN (1971) tarafından yapılan buna benzer çalışmalarda da burada gösterilen metodların bazıları tartışılmaktadır. Çok şıklı verilere model uydurmada esas sorunun en uygun tahmin metodunu seçmek olduğu belirtiliyor.

HABERMAN (1972), olabilirlik tabloları için log-linear uyum konusunu ele alır.

HABERMAN (1973a), frekans verileri için log-linear modeller adlı çalışmasında uygun istatistikler ve olabilirlik eşitliklerini veriyor.

HABERMAN (1973c), çok boyutlu olasılık tablolarının log-linear modellerle analizi konusunu ele alıyor.

HABERMAN (1974 b), düzenli sınıflanmış frekans tabloları için log-linear modeller konusunu ele alıyor.

BISHOP ve ark. (1975), kesikli çoklu değişken analizini teori ve uygulamalarıyla birlikte vererek log-linear model analizini büyük çapta araştırmacıların hizmetine sunmaktadır.

HABERMAN (1977b), küçük beklenen hücre değerlerine sahip frekans tabloları için log-linear model kullanımını açıklamaktadır.

HABERMAN (1978 a,b), kalitatif verilerin analizi konulu çalışmasında iki ve daha çok boyutlu tam ve tam olmayan olabilirlik tablolarının analizini örnekleriyle birlikte açıklamaktadır.

FALGON (1979), bir ekonomist olarak, uygulamada karşılaşılabilecek bazı durumları içeren kısa bir saha çalışması



sında, kalitatif bir değişkenin iki şıklı veya çok şıklı olmasına göre istatistiksel metodları tanımlıyor. Bu istatistiksel metodlar, linear model, logit/probit-type modeller, discriminant analizi ve çok şıklı logit modeli içermektedir. Ayrıca birkaç kalitatif bağımsız değişken analizi için istatistiksel metodlar yer almaktadır. Olabilirlik tablolarındaki ilişki ölçümüyle ilgili bazı genel incelemeler yapılıyor ve birkaç tip ilişki ölçümü yeniden gözden geçiriliyor. Dış etkilere bağımlı veya bağımlı olmayan olabilirlik tabloları için log-linear modeller tanımlanır. Sonuçta, genelleştirilmiş enküçük kareler tahmini kullanılarak, tahmin edilen karşılıklı bağımlı ikili değişkenler için linear bir model tanıtılıyor.

JAMES (1979), çok değişkenli kesikli verilerin log-linear modeller kullanılarak kovaryans analizi için birleştirilmiş teoriyi vermektedir. Model, vektör uzayındaki nokta terimlerinde bir "coordinate-free" gruplaşması şeklinde tanıtılıyor. Model için marginal ve şartlı dağılış özellikleri geliştirilerek, parametrelerin nasıl tahmin edilebileceği ve hipotez testleri gösteriliyor ve model Filipinlerde modern tarımsal uygulamanın adaptasyonu ile ilgili verilerin analizine uygulanmaktadır.

DANIEL GIANOLA (1979), hayvansal üretimde (hayatta kalma, buzağılama, hastalıklara dayanıklılık gibi) biyo-ekonomik öneme sahip birçok özelliklerin kesikli fenotip dağılış gösterdiğini ve hayvan ıslahçılarının genellikle fonksiyonları tahmin etmek ve kesikli veriler üzerine hipotezleri test etmek amacıyla, çok değişkenli normal dağılışa dayanan basma kalıp tekniklere başvurduklarını belirterek, linear logistik modelin ikili ve çok şıklı tesadüf değişkenlerini tanımlayan sabit modeller için basma kalıp tekniklere bir alternatif teşkil ettiğini öne sürmektedir. Araştırmacı, farelerle yapılan bir hat çaprazlaması deneyin-



den alınan verileri, logistik model uydurmak için hesaplama yöntemlerini tanımlamada kullanıyor ve bunu alışılmış khi-kare teknikleriyle mukayese ediyor.

RUTLEDGE ve GUNSETT (1982), zooteknide kesikli verilerin analizi için log-linear ve genelleştirilmiş en küçük kareler yaklaşımını açıklayarak, analiz metodlarını tanımlamak amacıyla, iki veri grubunu analiz ediyor. Araştırmacılar son yıllarda kalitatif verilerin analizinde önemli kuramsal ve yöntemsel gelişmeler olduğunu, bununla beraber yeni kuram ve yöntemlerin zootekniyle ilgili kullanımının çok az olduğunu belirterek, amaçlarının son birkaç kuramsal çalışma sonucunu uygun örneklerle göstermeye çalışarak zooteknistleri bu metodların benimsenmesi üzerinde düşünmek için harekete geçirmek olduğunu belirtmektedirler.

FIENBERG ve ark. (1983), geçen on yılda çok yönlü olasılık tablolarının log-linear ve logit modeller kullanılarak analizi konusunda çok sayıda kitap ve makaleler yayımlandığına işaret ederek makalelerinde bu çalışmalardan çoğunu destekleyen istatistiksel bir teori özetini vermektedirler. Özellikle psikometris ve ekonometrislerin ilgilendiği metod ve modellerle bir bağlantı kuruyorlar. Tartışma, özellikle büyük çok-yönlü olabilirlik tabloları için enbüyük-olabilirlik tahminleri konusunda geçmişte yapılan ve şimdi yapılmakta olan araştırmaların yeniden gözden geçirilmesini içermektedir.

OKUR (1983), bir çalışmada, özellikle sosyal ve davranış bilimleri alanlarında sık sık karşılaşılan çok boyutlu tabloların log-linear modeller kullanılarak analizi konusunu ele almış, modeldeki parametrelerin en büyük-olabilirlik tahminlerini elde ettikten sonra gözlemlerle beklenen değerlerin uyumunu, olabilirlik-oran ista-

tistiđi,  $L^2$ , kullanılarak test etmektedir. Bu alıřmada, model uydurulması, model seimi, uyumu sađlayan modellerin yorumlanması gibi konular zerinde durularak yntem  ynl bir tabloya uygulanmıř ve ortaya ıkan bazı sorunlar tartıřılmıřtır.



### 3. MATERYAL VE METOD

#### 3.1. İki-Yönlü Tam Tablolar

Kalitatif verilerin analizinde esas amaç, iki çok şıklı değişkenin bağımlı olup olmadıklarını kontrol etmektir. İki çok şıklı değişkenin bağımsızlık modeli bir log-linear modele eşdeğerdedir. Bu log-linear model için en büyük-olabilirlik tahminleri kesin çözümlere sahiptir. Bu nedenle khi-kare testleri ve düzeltilmiş rezidüyaller bağımsızlık modeline kolaylıkla uygulanır. Düzeltilmiş rezidüyallerin tek tek incelenmesi/bağımsızlıktan ayrılış kaynaklarının, bilinen Pearson khi-kare bağımsızlık testinden daha detaylı bir analizini elde etmek mümkündür.

Bağımsızlıktan ayrılışların büyüklüğünün kantitatif olarak bilinmesi en azından mevcut bağımsızlıktan ayrılışların tesbiti kadar önemlidir. Bu tür değerlendirme normal olarak logaritmik çapraz-çarpım oranları olarak bilinen logaritmik hücre ortalamalarının linear kombinasyonlarının kontrolüyle yapılır.

##### 3.1.1. 2x2 Boyutlu Tablolar

En basit olabilirlik tablosu dört hücrelidir ve kategoriler iki değişkenin sıklarından oluşur. Bu dört hücre, A değişkeninin kategorilerine tekabül eden iki satırı ve B değişkeninin kategorilerine tekabül eden iki sütunu olan 2x2 boyutlu bir tablo şeklinde düzenlenir. Diğer bir deyişle, ikiye ikilik olasılık tabloları iki tane iki şıklı değişkenin gözlem değerlerinin karşılıklı sınıflanmasıyla teşkil edilirler. Bu tip tablolar genellikle, şartlı ve marginal olasılıklarla veya eşdeğer olarak, şartlı ve marginal risk oranları ve logaritmik risk oranlarına

başvurularak karakterize edilebilirler. İki değişkenin bağımsızlığı şartlı olasılıklardaki farklarla veya şartlı logaritmik risklerin oranlarındaki farklarla tayin edilebilir. Logaritmik risklerin oranlarındaki bu fark logaritmik çarpaz-çarpım oranına götürür. Logaritmik çarpaz çarpım oranı, HABERMAN (1978)'in bildirdiğine göre YULE (1900 ve 1912)'ın iki şıklı iki değişkenin ilişkisini gösteren klasik istatistiğine eşdeğerdir. Logaritmik çarpaz-çarpım oranı ikiye ikilik olasılık tablosunun hücre ortalamalarının linear bir fonksiyonu olduğundan ve iki tane iki şıklı değişkenin bağımsızlığı da, logaritmik çarpaz çarpım oranının sıfır olmasına eşdeğer olduğundan, bağımsızlık hipotezi toplanabilir log-lienar modele eşdeğer görülebilir. Bu durumda enbüyük - olabilirlik tahminleri ve khi-kare istatistikleri kolayca hesaplanabilir.

Çizelge 1- İkiye-İkilik Olasılık Tablosu.

		B		$\sum$
		1	2	
A	1	$p_{11}$	$p_{12}$	$p_1^A$
	2	$p_{21}$	$p_{22}$	$p_2^A$
	$\sum$	$p_1^B$	$p_2^B$	1

İkili alt indis düzenlemedeki hücrelerin pozisyonunu verir. İlk indis, A değişkeninin ikinci indis, B değişkeninin kategori sayısını belirler,  $1 \leq i \leq 2$ ,  $1 \leq j \leq 2$ , ve iki boyutlu düzenleme iki satırlı ve iki sütunlu bir ızgara görünümündedir.

Burada N tane  $(A_h, B_h)$  çifti,  $1 \leq h \leq N$  olmak üzere bağımsız ve benzer şekilde dağılmıştır.  $P_{ij} > 0$ ,  $A_h = i$

ve  $B_n = j$  olma olasılığıdır.  $p_i^A$  ve  $p_j^B$  olasılıkları sırasıyla A değişkeninin i. ve B değişkeninin j. kategorisine ait olan şartsız olasılıklardır. Bunlara marginal olasılıklar adı verilir.

$$p_i^A = p_{i1} + p_{i2} = \sum_{j=1}^2 p_{ij}$$

$A_n = i$  olma olasılığıdır. Bu durumda Çizelge 1'den de görüleceği gibi

$$p_1^A = p_{11} + p_{12} = \sum_{j=1}^2 p_{1j}$$

$$p_2^A = p_{21} + p_{22} = \sum_{j=1}^2 p_{2j}$$

olur. İki marginal olasılık olan  $p_1^A$  ve  $p_2^A$  yı karşılaştırmak için, risk oranları, yani

$$q_{12}^A = \frac{p_1^A}{p_2^A}$$

veya

$$q_{21}^A = \frac{p_2^A}{p_1^A} = \frac{1}{q_{12}^A}$$

kullanılabilir.

$$p_1^A + p_2^A = 1$$

olduğundan,  $q_{12}^A$  hakkındaki bilgi,  $p_1^A$  ve  $p_2^A$  nın bilgisini ihtiva eder.

$$p_1^A = \frac{q_{12}^A}{(q_{12}^A + 1)}$$

ve

$$p_2^A = \frac{1}{(q_{12}^A + 1)}$$

dır. Benzer şekilde,

$$p_j^B = p_{1j} + p_{2j} = \sum_{i=1}^2 p_{ij}$$

$B_h = j$  olma olasılığıdır. Risk oranları,

$$q_{12}^B = \frac{p_1^B}{p_2^B}$$

veya

$$q_{21}^B = \frac{p_2^B}{p_1^B}$$

olur. Bununla da iki marginal olasılık olan  $p_1^B$  ve  $p_2^B$  karşılaştırılır.

$$p_1^B = \frac{p_{12}^B}{(q_{12}^B + 1)}$$

ve

$$p_2^B = \frac{1}{(q_{12}^B + 1)}$$

dir. Risklerin logaritmaları genellikle tahminlerin istatistiksel özelliklerini iyileştirmek ve pozitif değerler ile sınırlanmamış ölçümler elde etmek için kullanılır.

$$\tau_{12}^A = \log q_{12}^A$$

$$\tau_{21}^A = \log q_{21}^A = -\tau_{12}^A$$

$$\tau_{12}^B = \log q_{12}^B$$

$$\tau_{21}^B = \log q_{21}^B = -\tau_{12}^B$$

olur.  $\tau_{12}^A$  ile  $A_h = 2$  den ziyade  $A_h = 1$  olmasının logaritmik riski ve  $\tau_{12}^B$  ile de  $B_h = 2$  den ziyade  $B_h = 1$  olmasının logaritmik riski ifade edilir.

Çizelge 2- İkiye-ikilik Gözlenen Değerler Tablosu.

		B		$\Sigma$
		1	2	
A	1	$n_{11}$	$n_{12}$	$n_1^A$
	2	$n_{21}$	$n_{22}$	$n_2^A$
$\Sigma$		$n_1^B$	$n_2^B$	N

Marginal olasılıkları, marginal risk oranlarını tahmin etmek için tüm formüllerde,  $p_{ij}$  yerine  $n_{ij}/N$  kullanılabilir. Tahmin edilen asimtotik standart sapmalar BISHOP, FIENBERG ve HOLLAND (1975)'in tanımladığı ve  $\delta$  metodu denilen bir istatistiksel işlemle elde edilir.

$p_i^A$  marginal olasılıkları,

$$p_i^A = \frac{n_i^A}{N}$$

ile tahmin edilir. Burada

$$n_i^A = \sum_{j=1}^2 n_{ij} = n_{i1} + n_{i2}$$

dir.  $n_{ij}$ ,  $A_h = i$  ve  $B_h = j$  olan bireylerin gözlenen sayısıdır.  $n_i^A$ ,  $A_h = i$  olan birey sayısıdır. Böylece  $n_i^A/N$ ,  $A_h = i$  olma oranıdır.  $n_i^A/N$  in standart sapması,

$$\sigma(\hat{p}_i^A) = \left[ p_i^A (1-p_i^A)/N \right]^{1/2}$$

dir, ve örnek tahmini,

$$S(\hat{p}_i^A) = \left[ \frac{(n_i^A/N)(1 - \{n_i^A/N\})}{N} \right]^{1/2} = \left( \frac{n_1^A n_2^A}{N^3} \right)^{1/2}$$

ile elde edilir. Risk oranı olan  $q_{12}^A$  nin örnek tahmini,

$$\hat{q}_{12}^A = n_1^A/n_2^A$$

olur. Bunun tahmin edilen asimtotik standart sapması (EASD),

$$S(\hat{q}_{12}^A) = \frac{n_1^A}{n_2^A} \left( \frac{1}{n_1^A} + \frac{1}{n_2^A} \right)^{1/2}$$

dir. Logaritmik risk oranı olan  $\tau_{12}^A$  nin asimtotik standart sapması (EASD),

$$S(\hat{\tau}_{12}^A) = \left( \frac{1}{n_1^A} + \frac{1}{n_2^A} \right)^{1/2}$$

olur.  $\tau_{12}^A$  nin tahmini,

$$\hat{\tau}_{12}^A = \log(n_1^A/n_2^A)$$



dir.

Küçük gözlem değerlerine sahip örneklerde  $n_1^A$  yerine  $n_1^A + \frac{1}{2}$ ,  $n_2^A$  yerine  $n_2^A + \frac{1}{2}$  ve  $N$  yerine  $N+1$  konması tavsiye edilebilir. Bu işlem HALDANE (1955) ve GART ve ZWEIFEL (1967) nin tavsiyelerini yansıtır.

Benzer şekilde,  $p_j^B$  marginal olasılıkları

$$\hat{p}_j^B = n_j^B / N$$

ile tahmin edilebilirler. Burada,

$$n_j^B = \sum_{i=1}^2 n_{ij} = n_{1j} + n_{2j}$$

$B_h = j$  olan birey sayısıdır.  $n_j^B / N$  in tahmin edilen standart sapması,

$$S(\hat{p}_j^B) = \left( \frac{n_1^B n_2^B}{N^3} \right)^{1/2}$$

dir. Marginal olasılıklar için güven aralıkları bulunabilir. Örneğin,  $P_1^B$  için yaklaşık % 95'lik güven aralığı,

$$\hat{p}_1^B - 1.96(S(\hat{p}_1^B))$$

ve

$$\hat{p}_1^B + 1.96(S(\hat{p}_1^B))$$

dir.

Eğer istenirse,  $q_{12}^B$

$$\hat{q}_{12}^B = n_1^B / n_2^B$$

ile tahmin edilir, ve asimtotik standart sapması,

$$S(\hat{q}_{12}^B) = \frac{n_1^B}{n_2^B} \left( \frac{1}{n_1^B} + \frac{1}{n_2^B} \right)^{1/2}$$

dir.  $z_{12}^B$  nin tahmini,

$$\hat{z}_{12}^B = \log (n_1^B/n_2^B)$$

dir, ve

$$s(\hat{z}_{12}^B) = \left( \frac{1}{n_1^B} + \frac{1}{n_2^B} \right)^{1/2}$$

olur.  $q_{12}$  tahmininin 1'e ve  $z_{12}$  tahmininin 0'a yakın olması iki marginal grubun karşılaştırmasının ölçüsüdür.

Bir değişkenin diğer değişken üzerine etkisi ile ilgilenildiğinde şartlı olasılıklar dikkate alınmalıdır.  $B_h = j$  olduğu verildiğinde  $A_h = i$  olmasının şartlı olasılığı

$$P(A_h=i/B_h=j) = p_{i.j}^{A.B} = \frac{p_{ij}}{p_j}$$

dir.  $p_{1.j}^{A.B}$  yi  $p_{2.j}^{A.B}$  ile karşılaştırırken kullanılacak risk oranı,

$$q_{12.j}^{A.B} = \frac{p_{1.j}^{A.B}}{p_{2.j}^{A.B}} = \frac{p_{1j}/p_j^B}{p_{2j}/p_j^B} = \frac{p_{1j}}{p_{2j}} = \frac{Np_{1j}}{Np_{2j}} = \frac{m_{1j}}{m_{2j}}$$

dir, burada  $1 \leq i \leq 2$ ,  $1 \leq j \leq 2$  dir.  $m_{ij} = Np_{ij}$ ,  $n_{ij}$  nin beklenen değeridir. Bu risk oranı ile  $p_{1.j}^{A.B}$  ve  $p_{2.j}^{A.B}$ ,

$$p_{1.j}^{A.B} = \frac{q_{12.j}^{A.B}}{(q_{12.j}^{A.B} + 1)}$$

ve

$$p_{2.j}^{A.B} = \frac{1}{(q_{12.j}^{A.B} + 1)}$$

şeklinde bulunur. Bu risklerin logaritması

$$z_{12.j}^{A.B} = \log q_{12.j}^{A.B} = \log m_{1j} - \log m_{2j}$$

dir. Benzer şekilde,

$$q_{21.j}^{A.B} = \frac{p_{2.j}^{A.B}}{p_{1.j}^{A.B}} = \frac{m_{2j}}{m_{1j}} = \frac{1}{q_{12.j}^{A.B}}$$

ve

$$z_{21.j}^{A.B} = \log p_{21.j}^{A.B} = \log m_{2j} - \log m_{1j} = -z_{12.j}^{A.B}$$

dir.  $p_{i.j}^{A.B}$ ,

$$\hat{p}_{i.j}^{A.B} = \frac{n_{ij}}{n_j^B}$$

ile tahmin edilebilir. Bunun asimtotik standart sapması (EASD),

$$S(\hat{p}_{i.j}^{A.B}) = \left[ \frac{1}{n_j^B} \left( \frac{n_{ij}}{n_j^B} \right) \left( 1 - \frac{n_{ij}}{n_j^B} \right) \right]^{1/2} = \left( \frac{n_{ij} n_{2j}}{(n_j^B)^3} \right)^{1/2}$$

dir.  $q_{12.j}^{A.B}$  yi tahmin etmek için

$$\hat{q}_{12.j}^{A.B} = \frac{n_{1j}}{n_{2j}}$$

kullanılır ve

$$S(\hat{q}_{12.j}^{A.B}) = \left( \frac{n_{1j}}{n_{2j}} \right) \left( \frac{1}{n_{1j}} + \frac{1}{n_{2j}} \right)^{1/2}$$

dir. Benzer şekilde,

$$\hat{z}_{12j}^{A.B} = \log \left( \frac{n_{1j}}{n_{2j}} \right)$$

ve

$$s(\hat{z}_{12j}^{A.B}) = \left( \frac{1}{n_{1j}} + \frac{1}{n_{2j}} \right)^{1/2}$$

olur.

Bütünlüğü sağlamak için,

$$p_{j.i}^{B.A} = \frac{p_{ij}}{p_i^A}$$

$A_n =$  olduğu verildiğinde  $B_n = j$  olmasının şartlı olasılığı olarak belirtilir. Risk oranı ve logaritmik riskler aşağıdaki eşitliklerle tanımlanır:

$$q_{12.i}^{B.A} = \frac{1}{q_{21.i}^{B.A}} = \frac{p_{1.i}^{B.A}}{p_{2.i}^{B.A}} = \frac{m_{i1}}{m_{i2}}$$

ve

$$z_{12.i}^{B.A} = -z_{21.i}^{B.A} = \log q_{12.i}^{B.A} = \log m_{i1} - \log m_{i2}$$

Bu niceliklerin tahmini ve bunlara tekabül eden standart sapmalar (EASD),

$$\hat{p}_{j.i}^{B.A} = \frac{n_{ij}}{n_i^A},$$

$$s(\hat{p}_{j.i}^{B.A}) = \left[ \frac{n_{i1} n_{i2}}{(n_i^A)^3} \right]^{1/2},$$

$$s(\hat{q}_{12.i}^{B.A}) = \frac{n_{i1}}{n_{i2}} \left( \frac{1}{n_{i1}} + \frac{1}{n_{i2}} \right)^{1/2},$$

$$\hat{q}_{12.i}^{B.A} = \frac{n_{i1}}{n_{i2}}$$

$$\hat{z}_{12.i}^{B.A} = \log \left( \frac{n_{i1}}{n_{i2}} \right)$$

$$s(\hat{z}_{12.i}^{B.A}) = \left( \frac{1}{n_{i1}} + \frac{1}{n_{i2}} \right)^{1/2}$$

olur.

### 3.1.1.1. Bağımsızlık ve Toplanabilirlik

Bağımsızlık modeli de bir log-linear modeldir. A ve B değişkeni bağımsız ise

$$p_{ij} = p_i^A p_j^B, \quad 1 \leq i \leq 2, \quad 1 \leq j \leq 2$$

olur. Yani, bireysel olasılık marginal olasılıkların çarpımına eşit olur. Buna eşdeğer olarak,  $B_h = 1$  olduğu verildiğinde  $A_h = i$  olmasının şartlı olasılığı olan  $p_{i.1}^{A.B}$ ,  $B_h = 2$  olduğu verildiğinde  $A_h = i$  olmasının şartlı olasılığı olan  $p_{i.2}^{A.B}$  ile aynıdır. Böylece iki değişkenin bağımsızlığı sadece ve sadece

$$z_{12.1}^{A.B} = z_{12.2}^{A.B} = z_{12}^A$$

ise sağlanır. Logaritmik çapraz-çarpım oranı

$$\begin{aligned} z_{(12)(12)}^{AB} &= z_{(12).1}^{A.B} - z_{(12).2}^{A.B} = z_{(12).1}^{B.A} - z_{(12).2}^{B.A} \\ &= \log m_{11} - \log m_{21} - \log m_{12} + \log m_{22} \end{aligned}$$

bağımsızlıktan ayrılışın bir ölçüsü olarak ifade edilir. Buna göre, bağımsızlık modeli, log-linear modele eşdeğerdir. Burada, bağımsızlığın mevcut oluşu durumunda  $\log m_{ij}$ , aşağıdaki linear eşitlikle sınırlanır:

$$\zeta_{(12)}^{AB}(12) = \log m_{11} - \log m_{21} - \log m_{12} + \log m_{22} = 0$$

Yani,  $\zeta_{(12)}^{AB}(12) = 0$  olduğunda, A ve B değişkeni arasında bağımsızlık vardır. Bu sınırlama, eşitlik (3.1-1)'deki bazı  $\lambda_1^A, \lambda_2^A, \lambda_1^B, \lambda_2^B$  gibi parametrelerin toplanabilir olmalarına eşdeğerdir. Logaritmik ortalamalar yani  $\log m_{ij}$ 'ler

$$\log m_{ij} = \lambda + \lambda_i^A + \lambda_j^B$$

eşitliğini sağlarlar. Burada ki  $\lambda_1^A, \lambda_2^A, \lambda_1^B$  ve  $\lambda_2^B$  parametreleri de

$$\sum \lambda_i^A = \lambda_1^A + \lambda_2^A = 0$$

$$\sum \lambda_j^B = \lambda_1^B + \lambda_2^B = 0$$

$$\sum \lambda_i^A = \sum \lambda_j^B = 0$$

eşitliklerini sağlar.

$\zeta_{(12)}^{AB}(12) = 0$  şartının, toplanabilir log-linear modele tekabül ettiğini ispatlamak için,  $\lambda, \lambda_i^A, \lambda_j^B$  ve  $\lambda_{ij}^{AB}$ ,  $1 \leq i \leq 2, 1 \leq j \leq 2$ , nin tek değerleri ile

$$\log m_{ij} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB}$$

yazılabilir.

$$\sum_{i=1}^2 \lambda_i^A = \lambda_1^A + \lambda_2^A = 0, 1 \leq i \leq 2$$

$$\sum_{j=1}^2 \lambda_j^B = \lambda_1^B + \lambda_2^B = 0, \quad 1 \leq j \leq 2$$

$$\sum_{j=1}^2 \lambda_{ij}^{AB} = \lambda_{i1}^{AB} + \lambda_{i2}^{AB} = 0, \quad 1 \leq i \leq 2$$

$$\sum_{i=1}^2 \lambda_{ij}^{AB} = \lambda_{1j}^{AB} + \lambda_{2j}^{AB} = 0, \quad 1 \leq j \leq 2$$

**TÜRKİYE**  
**BİLİMSEL ve TEKNİK**  
**ARAŞTIRMA KURUMU**  
**KÜTÜPHANESİ**

dir. İkiye ikilik tablodaki, A ve B değişkeninin gerçekten aynı kategorilere sahip olduğunu ve basit olarak, bir değişken üzerinde iki farklı zamanda alınan ölçümler olduğunu kabul edelim. Bunları  $A_1$  ve  $A_2$  olarak isimlendirelim. Bununla ilgili üç farklı düzenleme vardır ve her bir düzenleme de bir çapraz-çarpım oranı elde edilir. Bu üç düzenleme ve bunlara tekabül eden çapraz-çarpım oranları:

		$A_2$		
		1	2	
- Esas tablo	1	$p_{11}$	$p_{12}$	$q_{(12)(12)}^{AB} = \frac{p_{11} p_{22}}{p_{21} p_{12}}$
	2	$p_{21}$	$p_{22}$	

		Aynı indis	Farklı indis	
-	1	$p_{11}$	$p_{12}$	$q_{1(12)(12)}^{AB} = \frac{p_{11} p_{21}}{p_{22} p_{12}}$
	2	$p_{22}$	$p_{21}$	

		$A_2$		
		1	2	
-	Aynı indis	$p_{11}$	$p_{22}$	$q_{2(12)(12)}^{AB} = \frac{p_{11} p_{12}}{p_{21} p_{22}}$
	Farklı indis	$p_{21}$	$p_{12}$	
		$p_{+1}$	$p_{+2}$	

şeklindedir. Bunların logaritması alındığında,

$$\tau_{(12)(12)}^{AB} = \log m_{11} - \log m_{21} - \log m_{12} + \log m_{22}$$

$$\tau_{1(12)(12)}^{AB} = \log m_{11} - \log m_{12} + \log m_{21} - \log m_{22}$$

$$\tau_{2(12)(12)}^{AB} = \log m_{11} + \log m_{12} - \log m_{21} - \log m_{22}$$

olur.

Çizelge 3- İki Tane İki Şıklı Değişken İçin  
Doymuş Modelin  $\lambda$  - Terimlerinin  
İşareti

Hücre	$\lambda$	$\lambda_i^A$	$\lambda_j^B$	$\lambda_{ij}^{AB}$
11	+	+	+	+
12	+	+	-	-
21	+	-	+	-
22	+	-	-	+

Böylece,

$$\lambda = \frac{1}{4} (\log m_{11} + \log m_{21} + \log m_{12} + \log m_{22})$$

$$\begin{aligned} \lambda_1^A = -\lambda_2^A &= \frac{1}{4} (\log m_{11} - \log m_{21} + \log m_{12} - \log m_{22}) \\ &= \frac{1}{4} \tau_{1(12)(12)}^{AB} \end{aligned}$$

$$\begin{aligned} \lambda_1^B = -\lambda_2^B &= \frac{1}{4} (\log m_{11} + \log m_{12} - \log m_{21} - \log m_{22}) \\ &= \frac{1}{4} \tau_{2(12)(12)}^{AB} \end{aligned}$$



$$\begin{aligned} \hat{\lambda}_{11}^{AB} = -\hat{\lambda}_{21}^{AB} = -\hat{\lambda}_{12}^{AB} = \hat{\lambda}_{22}^{AB} = \frac{1}{4} (\log m_{11} - \log m_{12} \\ - \log m_{12} + \log m_{22}) = \frac{1}{4} \tau_{(12)}^{AB} \end{aligned} \quad (12)$$

dir. Böylece sadece ve sadece  $\sum \hat{\lambda}_i^A = \sum \hat{\lambda}_j^B = 0$  olması şartı ile toplanabilir model

$$\log m_{ij} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_j^B \quad (3.1-1)$$

nin bazı  $\hat{\lambda}$ ,  $\hat{\lambda}_i^A$  ve  $\hat{\lambda}_j^B$  ler için sağlanması ile  $\tau_{(12)}^{AB} = 0$  olur.

$\hat{\lambda}_i^A$  parametreleri, A değişkeninin esas etkileridir.  $i \neq i'$  için  $\hat{\lambda}_i - \hat{\lambda}_{i'}^A = 2\hat{\lambda}_i^A$  olur, bu da  $\frac{1}{2} (\tau_{ii'.1}^{A.B} + \tau_{ii'.2}^{A.B})$  dir. Şartlı risklerin oranlarının  $(q_{ii'.j}^{A.B})$  ortalama logaritması  $(\tau_{ii'.j}^{A.B})$  toplanabilir (veya bağımsızlık) model altında

$$\hat{\lambda}_i^A - \hat{\lambda}_{i'}^A = \tau_{ii'}^A$$

dür. Böylece  $\hat{\lambda}_i^A$  iki kategori arasındaki nisbi tercihin bir ölçümünü verir.

Benzer şekilde,  $\hat{\lambda}_j^B$ , B değişkeninin esas etkileridir.  $j \neq j'$  için  $\hat{\lambda}_j - \hat{\lambda}_{j'}^B = 2\hat{\lambda}_j^B$  olur, bu da  $\frac{1}{2} (\tau_{jj'.1}^{B.A} + \tau_{jj'.2}^{B.A})$  olur. Şartlı logaritmik risklerin ortalaması  $\tau_{jj'.i}^{B.A}$ ,  $1 \leq i \leq 2$  dir. Bağımsızlık altında

$$\hat{\lambda}_j^B - \hat{\lambda}_{j'}^B = \tau_{jj'}^B$$

dür.

$\hat{\lambda}_{ij}^{AB}$ , A ve B deęişkenleri arasındaki interaksiyonlardır.  $i \neq i', j \neq j'$  için

$$\hat{\lambda}_{ij}^{AB} - \hat{\lambda}_{i'j}^{AB} - \hat{\lambda}_{ij'}^{AB} + \hat{\lambda}_{i'j'}^{AB} = 4\hat{\lambda}_{ij}^{AB}$$

$$Z_{(ii')(jj')}^{AB} = \log m_{ij} - \log m_{i'j} - \log m_{ij'} + \log m_{i'j'}$$

dür. Bağımsızlık altında  $\hat{\lambda}_{ij}^{AB}$  sıfırdır.

$m_{ij} = \log m_{ij} = \log p_{ij}$  kabul edilirse,  $\hat{\lambda}$ ,  $m_{ij}$ 'lerin logaritmasının genel ortalaması olarak

$$\hat{\lambda} = \frac{1}{4} (\log m_{11} + \log m_{21} + \log m_{12} + \log m_{22})$$

şeklinde ifade edilir.  $\hat{\lambda}_i^A$  ve  $\hat{\lambda}_j^B$ , genel ortalama  $\hat{\lambda}$  dan sapmaları gösterir.

$$\hat{\lambda}_i^A = \frac{m_{i+}}{2} - \frac{m_{++}}{4}$$

$$\hat{\lambda}_j^B = \frac{m_{+j}}{2} - \frac{m_{++}}{4}$$

ve

$$\hat{\lambda}_{ij}^{AB} = m_{ij} - \frac{m_{i+}}{2} - \frac{m_{+j}}{2} + \frac{m_{++}}{4}$$

dür.

### 3.1.1.2. Bağımsızlık Altında Enbüyük-Olabilirlik Tahminleri

Bağımsızlık modeli altında, marginal olasılığın ( $p_i^A$ ), en büyük-olabilirlik tahmini  $\hat{p}_i^A = n_i^A/N$  dir.  $\hat{p}_i^A$ , i. kategorideki bireylerin gözlenen oranıdır. Benzer şekilde,  $\hat{p}_j^B = n_j^B/N$ , j. kategorideki bireylerin gözlenen oranıdır.  $p_{ij}$ 'nin en büyük olabilirlik tahmini

$$\hat{p}_{ij} = \hat{p}_i^A \hat{p}_j^B$$

dir.

$$\hat{p}_i^A \hat{p}_j^B = (n_i^A/N)(n_j^B/N)$$

dir.  $m_{ij}$  nin en büyük-olabilirlik tahmini

$$\hat{m}_{ij} = N\hat{p}_{ij} = \frac{n_i^A n_j^B}{N}$$

dir.  $\hat{p}_i^A$ ,  $\hat{p}_j^B$ ,  $\hat{p}_{ij}$  ve  $\hat{m}_{ij}$ 'ler enbüyük-olabilirlik tahminleri oldukları söylenmeden daha önce kullanıldılar.

### 3.1.1.3. Bağımsızlık Testleri

HABERMAN (1978a)'a göre bağımsızlık modelini test etmek için PEARSON (1904) khi-kare istatistiği  $\chi^2$  veya WILKS (1935) olabilirlik-oran istatistiği  $L^2$  kullanılabilir.

$$\begin{aligned} \chi^2 &= \sum_i \sum_j \left( n_{ij} - \frac{n_i^A n_j^B}{N} \right)^2 / \left( \frac{n_i^A n_j^B}{N} \right) \\ &= \frac{N(n_{11}n_{22} - n_{12}n_{21})^2}{n_1^A n_2^A n_1^B n_2^B} \end{aligned}$$

ve

$$\begin{aligned} L^2 &= 2 \sum_i \sum_j n_{ij} \log \left( \frac{n_{ij}^N}{n_i^A n_j^B} \right) \\ &= 2 \left[ n_{11} \log \left( \frac{n_{11}^N}{n_1^A n_1^B} \right) + n_{12} \log \left( \frac{n_{12}^N}{n_1^A n_2^B} \right) \right. \\ &\quad \left. + n_{21} \log \left( \frac{n_{21}^N}{n_2^A n_1^B} \right) + n_{22} \log \left( \frac{n_{22}^N}{n_2^A n_2^B} \right) \right] \end{aligned}$$

bağıntıları ile ifade edilmektedir.

Bağımsızlık altında, her iki istatistik de bir serbestlik dereceli yaklaşık khi-kare dağılışı gösterir. HABERMAN (1978a)'a göre ilk olarak FISHER (1922) de belirtilmiştir. Hesaplamasının kolay oluşu,  $L^2$  den ziyade  $x^2$  nin kullanılmasına neden olur.

Pearson khi-kare istatistiği gözlenen ilişkinin genel yönünü göstermek için kullanılır. Ancak farklı sınıflara ait ilişkiler incelenmek istendiğinde düzeltilmiş rezidüyal üzerinde durulur. Düzeltilmiş rezidüyaller

$$r_{ij} = (n_{ij} - \hat{m}_{ij}) / (Nm_{ij} m_{i'j'})^{1/2}$$

dir ve

$$r_{11} = n_{11} - \left( \frac{n_{1A} n_{1B}}{N} \right) / \left( \frac{n_{1A} n_{2A} n_{1B} n_{2B}}{N} \right)^{1/2}$$

şeklinde hesaplanır.  $x^2 = r_{11}^2$  eşitliği sağlanır. Bununla ilgili detaylı açıklama daha fazla boyutlu olabilirlik tabloları incelenirken verilecektir.

#### 3.1.1.4. 2x2 Boyutlu Bir Tabloda İlişkinin Ölçülmesi

$x^2$  istatistiği A ve B değişkenlerinin ilişkili olup olmadığını gösterir. HABERMAN (1978a)'a göre ilişkinin büyüklüğünü tayin için, YULE (1900)'nin çapraz çarpım oranı

$$q_{(12)}^{AB}(12) = \frac{m_{11} m_{22}}{m_{21} m_{12}} = \frac{p_{11} p_{22}}{p_{21} p_{12}} = \frac{q_{(12)}.1}{q_{(12)}.2} \quad \begin{matrix} A.B \\ A.B \end{matrix}$$

veya onun logaritması olan

$$\begin{aligned} \tau_{(12)}^{AB}(12) &= \log m_{11} - \log m_{21} - \log m_{12} + \log m_{22} = \tau_{(12)}.1^{AB} \\ &\quad - \tau_{(12)}.2^{AB} \end{aligned}$$

kullanılabilir. Burada logaritma üzerinde durulmasının nedenleri, sınırsız bir aralıkta yer alması ve basit özellikleri sahip olmasındandır.

$$z_{(12)(12)}^{AB} = -z_{(12)(21)}^{AB} = -z_{(21)(12)}^{AB} = z_{(21)(21)}^{AB}$$

özelliğine sahiptir. Böylece kategorilerin sıklıklarının değiştirilmesi sadece logaritmik çarpım oranının işareti etkiler.

Bağımsızlık altında,  $q_{(12)(12)}^{AB} = 1$  ve  $z_{(12)(12)}^{AB} = 0$  dır. Diğer durumlarda  $q_{(12)(12)}^{AB} > 1$  ve  $z_{(12)(12)}^{AB} > 0$  veya  $q_{(12)(12)}^{AB} < 1$  ve  $z_{(12)(12)}^{AB} < 0$  olur. Eşitsizliğin yönü 1. şık 2. şıktan daha fazla olabilir ise  $>$  ve 1. şık 2. şıktan daha az olabilir ise  $<$  şeklinde olur.

İlişkileri tanımlamak için değişik katsayılar tanımlanmıştır. HABERMAN (1978a)'a göre YULE (1900), association (ilişki) katsayısını

$$Q = (q_{(12)(12)}^{AB} - 1) / (q_{(12)(12)}^{AB} + 1)$$

şeklinde tanımlamıştır. Aynı şekilde HABERMAN (1978a)'ın YULE (1912)'ye dayanarak bildirdiğine göre Colligation (Bağlantı) katsayısı

$$Y = (q_{(12)(12)}^{AB \ 1/2} - 1) / (q_{(12)(12)}^{AB \ 1/2} + 1)$$

olarak tanımlamıştır. Bunlar  $z_{(12)(12)}^{AB}$  nin basit fonksiyonlarıdır. Yorumlarının açık olması nedeniyle  $q_{(12)(12)}^{AB}$  ve  $z_{(12)(12)}^{AB}$  ölçümleri üzerinde daha fazla durulmaktadır.

$z_{(12)(12)}^{AB}$  yi tahmin etmek için

$$\hat{\tau}_{(12)(12)}^{AB} = \log n_{11} - \log n_{12} - \log n_{21} + \log n_{22}$$

eşitliği kullanılır. WOOLF (1955) de belirtildiği gibi, bu istatistiğe ait asimtotik varyans

$$\sigma^2(\hat{\tau}_{(12)(12)}^{AB}) = \frac{1}{m_{11}} + \frac{1}{m_{12}} + \frac{1}{m_{21}} + \frac{1}{m_{22}}$$

$$s^2(\hat{\tau}_{(12)(12)}^{AB}) = \frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}$$

ile tahmin edilir.

HALDANE (1955) ve GART ve ZWEIFEL (1967),  $\hat{\tau}_{(12)(12)}^{AB}$  ve  $s^2(\hat{\tau}_{(12)(12)}^{AB})$  yi hesaplamadan önce hücre frekanslarına ( $n_{ij}$ ), 1/2 eklemeyi önermektedirler. Buna göre  $\hat{\tau}_{(12)(12)}^{AB}$  için yaklaşık % 95 lik güven aralığı

$$\hat{\tau}_{(12)(12)}^{AB} - 1.96 (s(\hat{\tau}_{(12)(12)}^{AB}))$$

$$\hat{\tau}_{(12)(12)}^{AB} + 1.96 (s(\hat{\tau}_{(12)(12)}^{AB}))$$

olur.  $q_{(12)(12)}^{AB}$  için yukardaki aralığa tekabül eden sınırlar,

$$e \hat{\tau}_{(12)(12)}^{AB} - 1.96 (s(\hat{\tau}_{(12)(12)}^{AB}))$$

$$e \hat{\tau}_{(12)(12)}^{AB} + 1.96 (s(\hat{\tau}_{(12)(12)}^{AB}))$$

şeklinde ifade edilebilir.

İki değişken arasındaki ilişki, şartlı olasılıklardaki farkların ( $p_{i.j}^{A.B} - p_{i.j}^{A.B}$ ) tahminiyle de araştırılabilir. ( $\hat{p}_{i.j}^{A.B} - \hat{p}_{i.j}^{A.B}$ ) tahmininin standart sapması (EASD),

$$s(\hat{p}_{i.j}^{A.B} - \hat{p}_{i.j}^{A.B}) = \left[ s^2(\hat{p}_{i.j}^{A.B}) + s^2(\hat{p}_{i.j}^{A.B}) \right]^{1/2}$$

$$= \left[ \frac{n_{11}n_{21}}{(n_1^B)^3} + \frac{n_{12}n_{22}}{(n_2^B)^3} \right]^{1/2}$$

olur.

$(\hat{p}_{i.j}^{A.B} - \hat{p}_{i.j}^{A.B})$  için yaklaşık % 95'lik güven aralığı da bulunabilir.

Ancak, şartlı olasılıklardaki farklar çapraz çarpım oranlarının sağlamış olduğu bilgiyi vermezler. Çünkü şartlı olasılıklardaki farklar, karşılaştırılan olasılıkların büyüklüğünü dikkate almazlar. Bundan dolayı HABERMAN (1978a)'ın YULE (1900, 1912)'ye dayanarak bildirdiğine göre karşılaştırılan olasılıklardaki büyüklüğü de yansıtmak için şartlı olasılıklardaki farktan ziyade, çapraz-çarpım oranına  $(\chi_{(12)}^{AB}(12))$  eşdeğer ölçümlerin kullanılması önerilmektedir.

### 3.1.1.5. Süreklilik Düzeltmesi

HABERMAN (1978a)'a göre,  $\chi^2$ 'nin önem seviyesi  $\chi^2$  dağılışı yardımı ile bulunuyorsa, bu yaklaşımın iyileştirilmesi için YATES (1934) süreklilik düzeltmesi kullanmıştır.  $\chi^2$ 'nin bu şekilde modifikasyonu ile önem seviyesi iyileştirilebilmektedir. Yates düzeltmesi,

$$\chi_c^2 = \frac{N \left( |n_{11}n_{22} - n_{12}n_{21}| - \frac{1}{2} N \right)^2}{n_1^A n_2^A n_1^B n_2^B}$$

şeklinde yapılır. COCHRAN (1954) beklenen değerlerden bir kısmının 5'den küçük olması durumunda  $\chi^2$  den ziyade  $\chi_c^2$

kullanımını önermektedir.

### 3.1.2. 2xS Boyutlu Tablolar

Bir iki şıklı ve bir çok şıklı deęişkendeki gözlemlerin karşılıklı sınıflanmasıyla teşkil edilen olasılık tabloları ikiye-ikilik tabloların basitleştirilme özelliklerinin birkaçını bulundurur. Pearson khi-kare hesaplanmasında kullanışlı bazı sadeleştirmeler vardır, ve rezidüyal analizi önem kazanır.

Çizelge 4- 2xs Boyutlu Gözlenen Deęerler Tablosu

		B					$\Sigma$
		1	2	3	....	S	
A	1	$n_{11}$	$n_{12}$	$n_{13}$	...	$n_{1S}$	$n_1^A$
	2	$n_{21}$	$n_{22}$	$n_{23}$	...	$n_{2S}$	$n_2^A$
	$\Sigma$	$n_1^B$	$n_2^B$	$n_3^B$	...	$n_S^B$	N

Burada  $(A_h, B_h)$ ,  $1 \leq h \leq N$ , bağımsız ve benzer dağılmış tesadüfi deęişken çiftleridir.  $p_{ij} > 0$ ,  $A_h = i$  ve  $B_h = j$  olması olasılığıdır.  $n_{ij}$ ,  $A_h = i$  ve  $B_h = j$  olan bireylerin sayısıdır. A ve B deęişkeni arasında ilişki yoksa

$$p_{ij} = p_i^A \cdot p_j^B$$

olur, burada  $1 \leq i \leq 2$ ,  $1 \leq j \leq s$  dir. ~



### 3.1.2.1. Enbüyük-Olabilirlik Tahminleri

Daha önce de eşitlik (3.1-1) de belirtildiği gibi, bağımsızlık modeli olarak, log-linear model

$$\log m_{ij} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_j^B$$

şeklinde ifade edilebilir. Burada  $\sum \lambda_i^A = \sum \lambda_j^B = 0$  dir. Bu model kısım 3.1.3.1 de detaylı olarak incelenecektir. Bilindiği gibi  $m_{ij}$  nin enbüyük-olabilirlik tahmini

$$\hat{m}_{ij} = \frac{n_i^A n_j^B}{N}$$

dir.

### 3.1.2.2. Khi-Kare Testleri

HABERMAN (1978a)'ın PEARSON (1911)'a dayanarak belirttiği gibi

$$\begin{aligned} \chi^2 &= \sum_i \sum_j (n_{ij} - n_i^A n_j^B / N)^2 / (n_i^A n_j^B / N) \\ &= \sum_j \left( \frac{n_1^A n_2^A}{n_j^B} \right) \left( \frac{n_{1j}}{n_j^B} - \frac{n_{2j}}{n_j^B} \right)^2 \end{aligned}$$

olur. Yine HABERMAN (1978a)'ın WILKS (1935)'e dayanarak bildirdiği gibi olabilirlik-oran khi-karesi

$$L^2 = 2 \sum_i \sum_j n_{ij} \log \left( \frac{n_{ij} N}{n_i^A n_j^B} \right)$$

olur. Serbestlik derecesi  $(s-1)$  dir.

### 3.1.2.3. Düzeltilmiş Rezidüaller

Düzeltilmiş rezidüaller

$$r_{ij} = \frac{n_{ij} - (n_i^A n_j^B / N)}{\left\{ n_i^A n_j^B \left[ 1 - (n_i^A / N) \right] \left[ 1 - (n_j^B / N) \right] / N \right\}^{1/2}}$$

bağımsızlık altında yaklaşık standart normal dağılım gösterirler.  $r_{1j} + r_{2j} = 0$  olduğundan  $r_{1j}$  ye önem vermek yeterli olur.

2xS tablosu orta bir durum arzeder. İkiye-ikilik tablolarından daha karışıktır. Evvela, bağımsızlık için khi-kare testleri birden daha fazla serbestlik derecesine sahiptir. Bu nedenle de rezidüyal analizi önemsiz değildir, ve süreklilik düzeltmeleri yapılmaz. Diğer taraftan iki çok şıklı değişkenin karşılıklı sınıflanmasında daha çok düzeltilmiş rezidüyal görünür, ve olabilecek ilişkileri tanımlamada daha çok problemle karşılaşılır.

### 3.1.3. rxs Boyutlu Tablolar

Gözlemler,  $r \geq 3$  kategorili bir çok şıklı değişken ile  $S \geq 3$  kategorili bir çok şıklı değişkende karşılıklı sınıflandığında, bir rxs boyutlu olasılık tablosu ortaya çıkar. Daha basit iki yönlü tablolardaki gibi bağımsızlık hipotezi toplanabilir bir log-linear modele uyar, ve khi-kare testleri ve düzeltilmiş rezidüaller bu hipotezin testinde kullanılabilir. Eğer bağımlılık söz konusu ise, o zaman bağımsızlıktan sapmaların tanımı zor ve önemli bir problem olur. Bağımsızlığın büyüklüğünü tanımlamada, logaritmik çarpım oranları kullanılır.

Çizelge 5- rxs boyutlu Gözlenen Değerler  
Tablosu

		B					
		1	2	3	...	s	$\sum$
A	1	$n_{11}$	$n_{12}$	$n_{13}$	...	$n_{1s}$	$n_1^A$
	2	$n_{21}$	$n_{22}$	$n_{23}$	...	$n_{2s}$	$n_2^A$
	3	$n_{31}$	$n_{32}$	$n_{33}$	...	$n_{3s}$	$n_3^A$
	⋮	⋮	⋮	⋮		⋮	⋮
	r	$n_{r1}$	$n_{r2}$	$n_{r3}$		$n_{rs}$	$n_r^A$
$\sum$	$n_1^B$	$n_2^B$	$n_3^B$	...	$n_s^B$		

A ve B değişkeni arasında bir ilişki olup olmadığını anlamak için,  $(A_h, B_h)$ ,  $1 \leq h \leq N$ , benzer dağılmış bağımsız çiftler olsun. Eğer h. birey i. kategoride ise  $A_h = i$  olur, j. kategoride ise  $B_h = j$  olur.  $p_{ij} > 0$ ,  $A_h = i$  ve  $B_h = j$  olması olasılığıdır.  $n_{ij}$ ,  $A_h = i$  ve  $B_h = j$  olan bireylerin gözlenen sayısıdır.

### 3.1.3.1. Toplanabilir Log-linear Modeller

rxs boyutlu  $n_{ij}$  elemanlı bir n tablosu dikkate alınıyor. Burada  $1 \leq i \leq r$ ,  $1 \leq j \leq s$  ve  $r > 2$  ve  $s > 2$  dir.  $n_{ij}$  nin beklenen değeri  $m_{ij}$  dir. Daha genel toplanabilir bir log-linear modelin

$$\log (m_{ij}/z_{ij}) = \lambda + \lambda_i^A + \lambda_j^B \quad (3.1-2)$$

eşitliğini sağlayacağı varsayılır. Bu eşitlikte,  $z_{ij} > 0$  bilinir.  $\lambda$ ,  $\lambda_i^A$  ve  $\lambda_j^B$  bilinmez, ve  $\sum \lambda_i^A = \sum \lambda_j^B = 0$  dir. DRAPER ve SMITH (1966) deki gibi, eğer bu model regresyon modeli şeklinde düşünülürse,

$$\log (m_{ij}/z_{ij}) = \lambda + \sum_{i'=1}^{r-1} \lambda_{i'}^A x_{ii'}^A + \sum_{j'=1}^{s-1} \lambda_{j'}^B x_{jj'}^B$$

(3.1-3)

olur ve (3.1-2) de sağlanır. Burada x değişkenlerinin elemanları:  $1 \leq i' \leq r-1$  için,

$$\begin{aligned} x_{ii'}^A &= 1, & i &= i' \\ &= -1, & i &\neq i', i < r \\ &= 0, & i &= r \end{aligned}$$

ve  $1 \leq j' \leq s-1$  için

$$\begin{aligned} x_{jj'}^B &= 1, & j &= j' \\ &= -1, & j &= j', j < s \\ &= 0, & j &= s \end{aligned}$$

şeklinde tanımlanır. Örneğin, eğer  $i=j=1$  ise, o zaman  $x_{i1}^A = x_{j1}^B = 1$  ve diğer  $x_{ii'}^A$  ve  $x_{jj'}^B$  sıfırdır. Böylece model

$$\log (m_{11}/z_{11}) = \lambda + \lambda_1^A + \lambda_1^B$$

şekline dönüşür. Genel olarak log-linear model

$$\log (m_{ij}/z_{ij}) = \alpha + \sum_{k=1}^q \beta_k x_{ijk} \quad (3.1-4)$$

şeklinde ifade edilir.  $1 \leq i \leq r$ ,  $1 \leq j \leq s$ ,  $1 \leq k \leq q$  dür.  $z_{ij} > 0$  ve  $x_{ijk}$  bilinmekte ve  $\alpha$  ve  $\beta_k$  bilinmemektedir.

Eşitlik (3.1-4)'de verilen genel bağıntı ile (3.1-3)'de verilen özel durum arasındaki benzerliği görmek için,  $1 \leq k \leq r-1$  için  $x_{ijk} = x_{ik}^A$ ,  $r \leq k \leq r+s-2$  için  $x_{ijk} = x_j^B$  ( $k-r+1$ ),  $q = r+s-2$ ,  $1 \leq k \leq r-1$  için  $\alpha = \lambda$ ,  $\beta_k = \lambda_k^A$ ,

$r \leq k \leq r+s-2$  için,  $\beta_k = \lambda_k^B - r+1$  yazıldığında modellerin aynı model olduğu görülecektir.

### 3.1.3.2. Örneklerin Karşılaştırılması

Şimdiye kadar incelenen iki yönlü tablolar, iki çok şıklı değişken üzerine gözlemlerin karşılıklı sınıflandırmaları ile teşkil edilmişlerdi. Şu ana kadar kullanılan olasılık modeli  $N$  tane bağımsız ve benzer olarak dağılmış,  $(A_h, B_h)$  çiftlerinin gözlenmiş olduğunu varsayar, burada  $1 \leq h \leq N$  dir.  $p_{ij} > 0$ ,  $1 \leq i \leq r$  ve  $1 \leq j \leq s$  için  $A_h = i$  ve  $B_h = j$  olması olasılığıdır. Gözlenen değer  $n_{ij}$ ,  $A_h = i$  ve  $B_h = j$  olan  $h$  sayısıdır,  $1 \leq h \leq N$ . Üzerinde durulan hipotez bağımsızlık hipotezidir. Bağımsızlık altında

$$p_{ij} = p_i^A p_j^B$$

dir, burada  $1 \leq i \leq r$  için,  $p_i^A$ ,  $A_h = i$  olması olasılığıdır, ve  $1 \leq j \leq s$  için  $p_j^B$ ,  $B_h = j$  olması olasılığıdır. Bu bağımsızlık hipotezi toplanabilir log-linear modele eşdeğerdir.

İki-yönlü tablolar için bir diğer olasılık modeli dikkate alınabilir. Bu modelde  $s$  bağımsız örnek karşılaştırılır. Yine  $(A_h, B_h)$  çiftleri gözlenir,  $1 \leq h \leq N$ . Burada yine  $1 \leq A_h \leq r$ ,  $1 \leq B_h \leq s$  dir. Yine  $n_{ij}$ ,  $1 \leq i \leq r$ ,  $1 \leq j \leq s$  olmak üzere  $A_h = i$  ve  $B_h = j$  olan  $h$  sayısıdır. Bununla beraber,  $B_h$ , şimdi  $A_h$ 'ın alındığı örneği gösteren sabit bir değişkendir.  $B_h$  verildiğinde,  $1 \leq h \leq N$  olmak üzere,  $A_h$ 'ın bağımsız olarak dağıldığı farzedilir. Eğer  $B_h = j$  ise, o zaman  $A_h = i$  olma olasılığı,  $p_{i \cdot j}^{A \cdot B}$  dir. Üzerinde durulan hipotez homogenliktir. Homogenlik altında  $A_h$ 'ın dağılımı, çekildiği örnekten yani  $B_h$ 'den bağımsızdır.  $1 \leq i \leq r$ ,  $1 \leq j' \leq s$ ,  $1 \leq j \leq s$  için

$$p_{ij}^{A.B} = p_{i.j}^{A.B}$$

dir.

Homogenlik hipotezi bağımsızlık hipotezinden farklı bir olasılık modelinden ortaya çıkmasına rağmen, her ikisi de aynı toplanabilir modele eşdeğerdir. Sonuç olarak, en büyük olabilirlik tahminleri, khi-kare testleri ve düzeltilmiş rezidüeller, homogenlik modeli için, bağımsızlık modelindekilerle aynıdır. Birçok farklı olasılık modelleri ekseriya aynı log-linear modele ve aynı tahmin ve test yöntemlerine götürür.

$n_j^B$ ,  $B_h=j$  olan birey sayısı olduğuna göre, o zaman  $n_{ij}$ 'nin beklenen değeri

$$m_{ij} = n_j^B p_{i.j}^{A.B}$$

dir. Homogenlik modeli de sadece ve sadece

$$\log m_{ij} = \lambda + \lambda_i^A + \lambda_j^B$$

bazı  $\lambda$ ,  $\lambda_i^A$  ve  $\lambda_j^B$  için,  $\sum \lambda_i^A = \sum \lambda_j^B = 0$  sağlanırsa geçerlidir.

### 3.1.3.3. Bağımsızlık ve Toplanabilirlik

Eğer bağımsız ve benzer şekilde dağılmış çok şıklı  $(A_h, B_h)$  çiftlerinde,  $1 \leq h \leq N$ ,  $A_h$  ve  $B_h$  bağımsız ise, bağımsızlık modeli sağlanır. Eğer  $1 \leq A_h \leq r$ ,  $1 \leq B_h \leq s$  ise  $A_h = i$  ve  $B_h = j$  olma olasılığı  $p_{ij} > 0$  olur ve

$$p_{ij} = p_i^A p_j^B$$

şeklinde ifade edilir.  $n_{ij}$ 'nin ortalaması olan  $m_{ij}$ ,  $m_{ij} = N p_{ij}$  dir. Bu model, toplanabilir modelin, yani

$$\log (m_{ij}/z_{ij}) = \log m_{ij} = \log (Np_{ij}) = \lambda + \lambda_i^A + \lambda_j^B$$

nin

$$z=1 \text{ için, } \lambda = \log N + \frac{1}{r} \sum_{i'} \log p_{i'}^A + \frac{1}{s} \sum_{j'}$$

$$\log p_{j'}^B, \lambda_i^A = \log p_{i'}^A - \frac{1}{r} \sum_{i'} \log p_{i'}^A,$$

$$\lambda_j^B = \log p_{j'}^B - \frac{1}{s} \sum_{j'} \log p_{j'}^B$$

ile sağlandığını ifade eder. Böylece  $\lambda_i^A$ ,  $p_{i'}^A$  marginal olasılıklarının nisbi büyüklüğünü ölçer, ve  $\lambda_j^B$  ise  $p_{j'}^B$  marginal olasılıklarının nisbi büyüklüğünü ölçer.

Toplanabilirliğin bağımsızlığı ifade ettiği, logaritmik çarpım-çarpım oranı kullanılarak da ispatlanabilir.

$$\log \left( \frac{p_{ij} p_{11}}{p_{i1} p_{1j}} \right) = \log \left( \frac{m_{ij} m_{11}}{m_{i1} m_{1j}} \right)$$

eşitliği dikkate alınsın. Eğer toplanabilir model sağlanıyorsa, o zaman

$$\begin{aligned} \log \left( \frac{m_{ij} m_{11}}{m_{i1} m_{1j}} \right) &= \log m_{ij} - \log m_{i1} - \log m_{1j} + \log m_{11} \\ &= (\lambda + \lambda_i^A + \lambda_j^B) - (\lambda + \lambda_i^A + \lambda_1^B) \\ &\quad - (\lambda + \lambda_1^A + \lambda_j^B) + (\lambda + \lambda_1^A + \lambda_1^B) \\ &= 0 \end{aligned}$$

olur. Böylece

$$\frac{p_{ij} p_{11}}{p_{i1} p_{1j}} = 1$$

dir, ve

$$p_{ij} = \frac{p_{i1} p_{1j}}{p_{11}}$$

dir.

$$p_i^A = \sum_j \left( \frac{p_{i1} p_{1j}}{p_{11}} \right) = \frac{p_{i1}}{p_{11}} \sum_j p_{1j} = \frac{p_{i1} p_1^A}{p_{11}}$$

olduğunu belirtelim. Benzer şekilde,

$$p_j^B = \frac{p_1^B p_{1j}}{p_{11}}$$

dir.

$$\sum_i p_i^A = 1$$

olduğundan

$$1 = \sum_i \left( \frac{p_{i1} p_1^A}{p_{11}} \right) = \frac{p_1^A}{p_{11}} \sum_i p_{i1} = \frac{p_1^A p_1^B}{p_{11}}$$

dir. Şöyleki,

$$p_{11} = p_1^A \cdot p_1^B$$

dir. Daha genel olarak

$$\begin{aligned} p_{ij} &= \frac{p_{i1} p_{1j}}{p_{11}} \\ &= \left( \frac{p_{i1} p_1^A}{p_{11}} \right) \left( \frac{p_1^B p_{1j}}{p_{11}} \right) \end{aligned}$$



$$p_{ij} = p_i^A p_j^B$$

olmaktadır. Böylece toplanabilirliğin bağımsızlığı ifade ettiği görülür.

Bağımsızlık modeli altında, en büyük-olabilirlik eşitlikleri, tahmin edilen marginal olasılıkların, gözlenen marginal nisbi frekanslara eşit olduğunu ifade ederler, yani

$$\hat{p}_i^A = \hat{m}_i/N = n_i^A/N$$

ve

$$\hat{p}_j^B = n_j^B/N$$

dir.

#### 3.1.3.4. Doymuş Model

Doymuş model, tablonun sahip olduğu hücre sayısı kadar bağımsız parametreye sahip modeldir. Toplam bağımsız parametre sayısının toplam basit hücre sayısına eşit olduğunu doğrulamak için her bir  $\lambda$ -teriminin katkısı ilave edilir. Her bir terim için parametrelerin sayısı, aşağıda serbestlik dereceleri adı altında verilmiştir. Doymuş modelde parametrelerin toplamı  $rs$ , basit hücreler  $(i,j)$  sayısına eşittir.

$\lambda$ -terimleri	Serbestlik derecesi
$\lambda$	1
$\lambda_i^A$	$r-1$
$\lambda_j^B$	$s-1$
$\lambda_{ij}^{AB}$	$rs-r-s-1$
Toplam	$rs$

Doymuş log-linear bir model

$$\log m_{ij} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_{ij}^{AB} \quad (3.1-5)$$

şeklinde ifade edilir, burada  $\sum_i \lambda_i^A = \sum_j \lambda_j^B = \sum_i \lambda_{ij}^{AB}$

$= \sum_j \lambda_{ij}^{AB} = 0$  dır. GOODMAN (1970) ve BISHOP FIENBERG ve

HOLLAND (1975)'ın belirttiği gibi, bu log-linear modelde hiç bir sınırlama konmamıştır.

$$\lambda = \frac{1}{rs} \sum_i \sum_j \log m_{ij}$$

$$\lambda_i^A = \frac{1}{r} \sum_j \log m_{ij} - \lambda$$

$$\lambda_j^B = \frac{1}{s} \sum_i \log m_{ij} - \lambda$$

ve

$$\lambda_{ij}^{AB} = \log m_{ij} - \lambda - \lambda_i^A - \lambda_j^B$$

şeklinde ifadeler yazılabilir.

Model regresyon modeli şeklinde gösterilmek istenirse,

$$\begin{aligned} \log m_{ij} = & \lambda + \sum_{i'=1}^{r-1} \lambda_{i'}^A x_{ii'}^A + \sum_{j'=1}^{s-1} \lambda_{j'}^B x_{jj'}^B \\ & + \sum_{i'=1}^{r-1} \sum_{j'=1}^{s-1} \lambda_{i'j'}^{AB} x_{ii'}^A x_{jj'}^B \end{aligned}$$

yazılabilir. En büyük-olabilirlik eşitlikleri,

$$\sum_i \sum_j \hat{m}_{ij} = \sum_i \sum_j \hat{n}_{ij},$$

$$\sum_i \sum_j x_{ii'}^A \hat{m}_{ij} = \sum_i \sum_j x_{ii'}^A n_{ij}, \quad 1 \leq i' \leq r-1,$$

$$\sum_i \sum_j x_{jj'}^B \hat{m}_{ij} = \sum_i \sum_j x_{jj'}^B n_{ij}, \quad 1 \leq j' \leq s-1,$$

$$\sum_i \sum_j x_{ii'}^A x_{jj'}^B \hat{m}_{ij} = \sum_i \sum_j x_{ii'}^A x_{jj'}^B n_{ij},$$

$$1 \leq i' \leq r-1, \quad 1 \leq j' \leq s-1$$

olur. Tahminler yerine konduğunda model,

$$\log \hat{m}_{ij} = \hat{\lambda} + \sum_{i'=1}^{r-1} \hat{\lambda}_{i'}^A x_{ii'}^A + \sum_{j'=1}^{s-1} \hat{\lambda}_{j'}^B x_{jj'}^B + \sum_{i'=1}^{r-1} \sum_{j'=1}^{s-1} \hat{\lambda}_{i'j'}^{AB} x_{ii'}^A x_{jj'}^B$$

olur. Eğer  $\hat{m}_{ij} = n_{ij}$  ise

$$\hat{\lambda} = \frac{1}{rs} \sum_i \sum_j \log m_{ij}$$

$$\hat{\lambda}_i^A = \frac{1}{r} \sum_j \log m_{ij} - \hat{\lambda}$$

$$\hat{\lambda}_j^B = \frac{1}{s} \sum_i \log m_{ij} - \hat{\lambda}$$

ve

$$\hat{\lambda}_{ij}^{AB} = \log m_{ij} - \hat{\lambda} - \hat{\lambda}_i^A - \hat{\lambda}_j^B$$

olur.

$$\tau = \sum_i \sum_j c_{ij} \hat{\lambda}_{ij}^{AB}$$

şeklinde ifade edilsin. Burada  $c_{ij}$  sabitleri,

$$\sum_{i=1}^r c_{ij} = 0, \quad 1 \leq j \leq s$$

$$\sum_{j=1}^s c_{ij} = 0, \quad 1 \leq i \leq r$$

olur.  $\lambda_{ij}^{AB}$  nin eşitlik (3.1-5) den çekilmesi ile elde edilen kısım bunun yerine konursa o zaman,

$$\begin{aligned} \zeta &= \sum_i \sum_j c_{ij} (\log m_{ij} - \lambda - \lambda_i^A - \lambda_j^B) \\ &= \sum_i \sum_j c_{ij} \log m_{ij} \end{aligned}$$

ve tahmin edici

$$\hat{\zeta} = \sum_i \sum_j c_{ij} \log n_{ij}$$

olur.

Poisson ve multinomial örnekleme altında  $\zeta$ , asimptotik varyansa sahiptir.

$$\sigma^2(\hat{\zeta}) = \sum_i \sum_j c_{ij}^2 / m_{ij}$$

olur. Bu formül satır-multinomial örnekleme altında da uygulanır. O zaman

$$\begin{aligned} \zeta &= \sum_i \sum_j c_{ij} \log (n_i^A p_{j.i}^{B.A}) \\ &= \sum_i \sum_j c_{ij} \log n_i^A + \sum_i \sum_j c_{ij} \log p_{j.i}^{B.A} \\ &= \sum_i \sum_j c_{ij} \log p_{j.i}^{B.A} \end{aligned}$$

olur. Burada  $p_{j.i}^{B.A}$  olasılıklarının bir fonksiyonudur. Benzer şekilde formül sütun-multinomial örnekleme altında uygulanırsa, o zaman

$$\begin{aligned} \tau &= \sum_i \sum_j c_{ij} \log (n_j^B \log p_{i.j}^{A.B}) \\ &= \sum_i \sum_j c_{ij} \log p_{i.j}^{A.B} \end{aligned}$$

olur, yine  $p_{i.j}^{A.B}$  olasılıklarının bir fonksiyonudur.

Bu neticenin önemli bir uygulaması logaritmik çarpaz-çarpım oranı olan

$$\begin{aligned} \tau_{(i' i'')(j' j'')}^{AB} &= \log m_{i' j'} - \log m_{i' j''} - \log m_{i'' j'} \\ &\quad + \log m_{i'' j''} \\ &= \lambda_{i' j'}^{AB} - \lambda_{i' j''}^{AB} - \lambda_{i'' j'}^{AB} + \lambda_{i'' j''}^{AB} \end{aligned}$$

ile ilgilidir, burada  $i' \neq i''$  ve  $j' \neq j''$  dür.

$$\begin{aligned} c_{ij} &= 1 \quad ; \quad i=i'' \text{ ve } j'=j \text{ veya } i=i'' \text{ ve } j=j'', \\ &= -1 \quad ; \quad i=i' \text{ ve } j''=j \text{ veya } i=i' \text{ ve } j=j', \\ &= 0 \quad ; \quad \text{diğer durumlarda} \end{aligned}$$

değerlerini alır. Buna göre, genel gösterimi ile,

$$\begin{aligned} \tau_{(i' i'')(j' j'')}^{AB} &= \sum_i \sum_j c_{ij} \log m_{ij} \\ \sum_i c_{ij} &= 0, \quad 1 \leq j \leq s \end{aligned}$$

ve

$$\sum_j c_{ij} = 0, \quad 1 \leq i \leq r,$$

olur.

Böylece, tahmin edilen logaritmik çapraz-çarpım oranı da

$$\hat{\chi}_{(i'i'')(j'j'')}^{AB} = \log n_{i'j'} - \log n_{i'j''} - \log n_{i''j'} + \log n_{i''j''}$$

olur. Bu tahminin asimtotik varyansı

$$\sigma^2(\hat{\chi}_{(i'i'')(j'j'')}^{AB}) = \frac{1}{m_{ij'}} + \frac{1}{m_{i'j''}} + \frac{1}{m_{i''j'}} + \frac{1}{m_{i''j''}}$$

şeklinde ifade edilebilir.

### 3.1.3.5. Çapraz-Çarpım Oranları

$z_{ij} = 1$  durumunda logaritmik çapraz-çarpım oranlarının tahminleri, toplanabilir modelden sapmaları tanımlamada büyük role sahiptirler. Çapraz-çarpım oranı

$$q_{(ii')(jj')}^{AB} = \frac{m_{ij} m_{i'j'}}{m_{ij'} m_{i'j}} = \frac{p_{ij} p_{i'j'}}{p_{ij'} p_{i'j}}$$

dir. Bunun logaritması olan logaritmik çapraz-çarpım oranı,

$$\begin{aligned} \chi_{(ii')(jj')}^{AB} &= \log \left( \frac{m_{ij} m_{i'j'}}{m_{ij'} m_{i'j}} \right) = \log m_{ij} - \log m_{ij'} \\ &\quad - \log m_{i'j} + \log m_{i'j'} \\ &= 0, \quad i \neq i', \quad j \neq j' \end{aligned} \quad (3.1-6)$$

ise bu durumda toplanabilir model sağlanır ( $i=i'$  veya  $j=j'$  ise, o zaman toplanabilir model sağlansada sağlanmasada  $\tau_{(ii')(jj')}^{AB} = 0$  dır). Eşitlik (3.1-6) nın toplanabilir model için sağlandığını ispat etmek için, bu eşitliğin sağ tarafı

$$\log m_{ij} - \log m_{ij'} - \log m_{i'j} + \log m_{i'j'}$$

alındığında, bunun eşdeğeri olan bağıntı yazılırsa,

$$\begin{aligned} (\lambda + \lambda_i^A + \lambda_j^B) - (\lambda + \lambda_i^A + \lambda_{j'}^B) - (\lambda + \lambda_{i'}^A + \lambda_j^B) \\ + (\lambda + \lambda_{i'}^A + \lambda_{j'}^B) = 0 \end{aligned}$$

olduğu görülür. Diğer taraftan (3.1-6) dan tüm  $i$ 'ler,  $1 \leq i \leq r$ , ve tüm  $j$ 'ler,  $1 \leq j \leq s$ , için  $\log m_{ij}$  çekilirse,

$$\begin{aligned} \log m_{ij} &= \log m_{ij'} + \log m_{i'j} - \log m_{i'j'} \\ &= \lambda + \lambda_i^A + \lambda_j^B \end{aligned}$$

elde edilir. Burada,

$$\begin{aligned} \lambda &= \frac{1}{r} \sum_{i''} \log m_{i''j} + \frac{1}{s} \sum_{j''} \log m_{i'j''} \\ &\quad - \log m_{i'j'} \end{aligned}$$

$$\lambda_i^A = \log m_{ij'} - \frac{1}{r} \sum_{i''} \log m_{i''j'}$$

ve

$$\lambda_j^B = \log m_{i'j} - \frac{1}{s} \sum_{j''} \log m_{i'j''}$$

dür. Böylece (3.1-6) nın toplanabilir modeli ifade ettiği ispatlanmış olur.

$\tau_{(ii')(jj')}^{AB}$  üzerine birçok cebirsel özdeşlikler yazılabilir. Örneğin

$$\begin{aligned}\tau_{(ii')(jj')}^{AB} &= -\tau_{(ii')(j'j)}^{AB} = -\tau_{(i'i)(jj')}^{AB} \\ &= \tau_{(i'i)(j'j)}^{AB}\end{aligned}$$

ve

$$\tau_{(ii)(jj')}^{AB} = \tau_{(ii')(jj)}^{AB} = 0$$

Toplanabilir modelden ayrılmalar,  $\tau_{(ii')(jj')}^{AB}$  da  $\tau_{(ii')(jj')}^{AB}$  in hepsi kullanılmayabilir, ancak tanımlanabilir. Her sabit  $i''$  ve  $j''$  için  $\tau_{(ii'')(jj'')}^{AB}$ ,  $i \neq i''$ ,  $j \neq j''$ , tüm logaritmik çapraz-çarpım oranlarını tarif eder. İspat için,

$$\begin{aligned}\log \left[ \frac{m_{ij} m_{i'j'}}{m_{ij'} m_{i'j}} \right] &= \log \left( \frac{m_{ij} m_{i'j'}}{m_{ij''} m_{i''j}} \right) \\ &+ \log \left( \frac{m_{ij''} m_{i''j}}{m_{ij'} m_{i'j''}} \right)\end{aligned}$$

$$+ \log \left( \frac{m_{i''j} m_{i'j''}}{m_{i''j''} m_{i'j}} \right) + \log \left( \frac{m_{i''j'} m_{i'j'}}{m_{i''j} m_{i'j''}} \right)$$

olduğu belirtilir. Diğer bir ifadeyle,

$$\begin{aligned}\tau_{(ii')(jj')}^{AB} &= \tau_{(ii'')(jj'')}^{AB} + \tau_{(ii'')(j''j')}^{AB} \\ &+ \tau_{(i''i')(jj'')}^{AB} + \tau_{(i''i')(j''j')}^{AB} \\ &= \tau_{(ii'')(jj'')}^{AB} - \tau_{(ii'')(j'j'')}^{AB}\end{aligned}$$



$$- \tau_{(i'i'')(jj'')}^{AB} + \tau_{(i'i'')(j'j'')}^{AB}$$

olur. Toplanabilir modelden sapmaları ortaya koymak için, doymuş model altında tahmin edilen logaritmik çapraz-çarpım oranı olan,

$$\hat{\tau}_{(ii')(jj')}^{AB} = \log n_{ij} - \log n_{ij'} - \log n_{i'j} + \log n_{i'j'}$$

ile, buna tekabül eden tahmin edilen asimtotik varyans,

$$S^2(\hat{\tau}_{(ii')(jj')}^{AB}) = \frac{1}{n_{ij}} + \frac{1}{n_{ij'}} + \frac{1}{n_{i'j}} + \frac{1}{n_{i'j'}}$$

beraber kullanılabilir. Bu tahminler mümkün olan dört örnekleme işlemi altında da uygulanabilir. GART ve ZWEIFEL (1967)'de belirtildiği gibi  $\tau_{(ii')(jj')}^{AB}$  tahmin edilirken,  $n_{ij}$ 'lere  $1/2$  eklenerek asimtotik sapma (biās) yı indirgemek için düzeltme yapılabilir. Böylece,  $\hat{\tau}_{(ii')(jj')}^{AB}$  nün  $1/2$  ye göre düzeltilmiş hali,

$$\begin{aligned} \bar{\tau}_{(ii')(jj')}^{AB} &= \log(n_{ij} + 1/2) - \log(n_{ij'} + 1/2) \\ &\quad - \log(n_{i'j} + 1/2) + \log(n_{i'j'} + 1/2) \end{aligned}$$

şeklinde tanımlanan bir ifade ile tahmin edilir. Bunun asimtotik varyans tahmini  $S^2(\bar{\tau}_{(ii')(jj')}^{AB})$ ,

$$\begin{aligned} S^2(\bar{\tau}_{(ii')(jj')}^{AB}) &= \frac{1}{n_{ij} + 1/2} + \frac{1}{n_{ij'} + 1/2} + \frac{1}{n_{i'j} + 1/2} \\ &\quad + \frac{1}{n_{i'j'} + 1/2} \end{aligned}$$

olur.

### 3.1.3.6. Khi-Kare Testleri

Eğer (3.1-4) ile tanımlanan log-linear model sağlanıyorsa, o zaman Pearson khi-kare istatistiği

$$x^2 = \sum_i \sum_j (n_{ij} - \hat{m}_{ij})^2 / \hat{m}_{ij}$$

dir. Bağımsızlık altında  $\hat{m}_{ij} = n_i^A n_j^B / N$  olduğundan

$$\begin{aligned} x^2 &= \sum_i \sum_j (n_{ij} - n_i^A n_j^B / N)^2 / (n_i^A n_j^B / N) \\ &= N \sum_i (1/n_i^A) \sum_j (1/n_j^B) n_{ij}^2 - N \end{aligned}$$

olur. Olabilirlik-oran istatistiği

$$\begin{aligned} L^2 &= 2 \sum_i \sum_j n_{ij} \log (n_{ij} / \hat{m}_{ij}) \\ &= 2 \sum_i \sum_j n_{ij} \log \left[ (n_{ij} N) / (n_i^A n_j^B) \right] \\ &= 2 \left( \sum_i \sum_j n_{ij} \log n_{ij} - \sum_i n_i^A \log n_i^A - \sum_j n_j^B \right. \\ &\quad \left. \log n_j^B + N \log N \right) \end{aligned}$$

dir. Her ikisi de  $rs - q - 1$  serbestlik dereceli ve yaklaşık khi-kare dağılışı gösterir. Yaklaşım,  $m_{ij}$  ortalamalarının yeteri kadar büyük ve  $\alpha$  ve  $\beta_k$  parametrelerinin,  $1 \leq k \leq q$ , tek olarak  $\log (m_{ij} / z_{ij})$ ,  $1 \leq i \leq r$ ,  $1 \leq j \leq s$  ile tayin edilmiş olmasını gerektirir.

Toplanabilir log-linear model altında bu istatis-

tikler  $(r-1)(s-1)$  serbestlik dereceli khi-kare dağılışı gösterirler.  $rs$  tane  $n_{ij}$  değeri ve  $(r+s-1)$  tane  $\lambda$ -parametresi olduğundan, serbestlik derecesi  $rs - (r+s-1) = (r-1)(s-1)$  olur. MARGOLIN ve LIGHT (1974),  $L^2$  ye kıyasla khi-kare dağılımının,  $\chi^2$  dağılımını daha iyi tanımladığını belirtmektedirler. Ancak bu netice kesin değildir.

### 3.1.3.7. Düzeltilmiş Rezidüaller

Toplanabilir model altında tüm  $z_{ij} = 1$  olması şartıyla HABERMAN (1973b, 1974a) da belirtildiği gibi düzeltilmiş rezidüyal,

$$r_{ij} = \frac{n_{ij} - (n_i^A n_j^B / N)}{\left\{ n_i^A n_j^B \left[ 1 - (n_i^A / N) \right] \left[ 1 - (n_j^B / N) \right] / N \right\}^{1/2}}$$

şeklinde ifade edilir.

### 3.1.3.8. Standardize Parametre Tahminleri

İnteraksiyon etkilerinin ve buna tekabül eden standart sapmaların karşılaştırılması verilerle uyumlu modelin tayin etmede değerli bir araçtır. Bunun içinde etkilerinin sıfırdan farklılığının test edilmesi gerekir. Hesaplanan parametre tahmin değerleri kendi standart hatasına bölünerek standardize edilmiş değerler elde edilir. Bu standardize değerler standart normal dağılışı gösterirler. Test  $\alpha = 0.05$  önem seviyesinde yapılırsa, standardize değerler 1.96 ile karşılaştırılır. Bundan büyükse önemlidir denir ve parametre modele dahil edilir.

Doymuş model altında  $\lambda_{ij}^{AB}$  nin asimtotik varyansı

$$\sigma^2(\hat{\lambda}_{ij}^{AB}) = v_1^2 \frac{1}{m_{ij}} + v_2^2 \sum_{s \neq j} \frac{1}{m_{is}} + v_3^2 \sum_{r \neq i} \frac{1}{m_{rj}} + v_4^2 \sum_{\substack{r \neq i \\ s \neq j}} \frac{1}{m_{rs}}$$

dir ve  $\{m_{ij}\}$  yerine  $\{n_{ij}\}$  konularak tahmin edilir. Burada,

$$v_1 = \frac{(1-r)(1-s)}{rs},$$

$$v_2 = \frac{(1-r)}{rs},$$

$$v_3 = \frac{(1-s)}{rs},$$

$$v_4 = \frac{1}{rs}$$

dir. Test istatistiği,

$$Z_{AB} = \frac{\hat{\lambda}_{ij}^{AB}}{S \cdot (\hat{\lambda}_{ij}^{AB})} \sim N(0,1), \text{ kritik de\u011fer } Z_{0.05} = 1.96$$

dır.

$$H_0: \hat{\lambda}_{ij}^{AB} = 0$$

$$H_1: \hat{\lambda}_{ij}^{AB} \neq 0$$

olur. Bu hipotez her  $i$  ve  $j$  için ayrı ayrı test edilir.

Doymuş modelde standardize  $\hat{\lambda}$ -parametreleri elde etmenin faydaları aşağıdaki şekilde sıralanabilir:

1. İnteraksiyonları büyüklüklerine göre sıralamak.

Bu, tüm değişkenler iki şıklı olduğunda ve her  $\lambda$ -terimi tek mutlak değerle tanımlandığında çok faydalı olur.

2. Bir değişken için kategori sayılarının ne zaman indirgenebileceğini ortaya koymak.

3. İlerde belirli bir doymamış modelin verilere ne kadar iyi uyacağını tayin etmek.

### 3.2. Üç-Yönlü Tam Tablolar

Her birey üzerinde ölçülen özellik sayıları giderek arttığında ve bu özellikler karşılıklı sınıflandırıldığında elde edilen çok boyutlu olasılık tabloları karmaşık bir yapı kazanır.

Üç çok şıklı değişken karşılıklı sınıflandırıldığında değişkenler arasındaki çok sayıda ilişki incelenebilir. Bu tür ilişkileri tanımlamada modele dahil edilecek parametrelerin bir düzen içerisinde modele dahil edilmesini sağlayan, hiyerarşik modeller denen bir model sınıfı kullanılır. Bu modeller üç-yönlü tablolar için uygun olan teklog-linear model değildir. Bununla beraber hiyerarşik modeller iki veya daha fazla değişkenin karşılıklı sınıflandığı olasılık tablolarının kontrolü için genel bir yöntemdir. Bu tür modeller, genellikle faktöriyel tabloların varyans analizinde kullanılan genel bir parametrisasyona dayanır. Log-linear modellerin, üç-yönlü tablolar için aydınlatılması durumunda,  $n_{ijk}$  değerleri,  $1 \leq i \leq r$ ,  $1 \leq j \leq s$ ,  $1 \leq k \leq t$  olmak üzere üç-yönlü bir tabloda düzenlenir. Her  $n_{ijk}$ 'nin pozitif beklenen değere ( $m_{ijk}$ ) sahip olduğu varsayılır. O zaman A, B, C değişkenleri ve bunların interaksyonlarını temsil eden parametreler  $\lambda$ ,  $\lambda_i^A$ ,  $\lambda_j^B$ ,  $\lambda_k^C$ ,  $\lambda_{ij}^{AB}$ ,  $\lambda_{ik}^{AC}$ ,  $\lambda_{jk}^{BC}$  ve  $\lambda_{ijk}^{ABC}$ ,  $1 \leq i \leq r$ ,  $1 \leq j \leq s$ ,  $1 \leq k \leq t$  için olur. Buna göre, genel doymuş model,

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC} \quad (3.2-1)$$

olmaktadır. Parametreler üzerine konması gerekli olan kısıtlamalar ise,

$$\begin{aligned} \sum \lambda_i^A &= \sum \lambda_j^B = \sum \lambda_k^C = \sum_i \lambda_{ij}^{AB} = \sum_j \lambda_{ij}^{AB} = \sum_i \lambda_{ik}^{AC} \\ &= \sum_k \lambda_{ik}^{AC} = \sum_j \lambda_{jk}^{BC} = \sum_k \lambda_{jk}^{BC} = \sum_i \lambda_{ijk}^{ABC} \\ &= \sum_j \lambda_{ijk}^{ABC} = \sum_k \lambda_{ijk}^{ABC} = 0 \end{aligned} \quad (3.2-2)$$

şeklinde ifade edilir.

Varyans analizinde olduğu gibi,  $\lambda_i^A$  parametreleri, A değişkeni için esas etkiler,  $\lambda_{ij}^{AB}$  ler AxB interaksiyon etkisi ve  $\lambda_{ijk}^{ABC}$  ler ise AxBxC interaksiyon etkisi olarak adlandırılır.  $\lambda_{ij}^{AB}$ ,  $\lambda_{jk}^{BC}$ ,  $\lambda_{ik}^{AC}$  ikili-faktör interaksiyonları,  $\lambda_{ijk}^{ABC}$  ler ise üçlü-faktör interaksiyonlarıdır. Mümkün olan  $\lambda$  parametrelerinin hepsini ihtiva eden model (3.2.-1) doymuş bir modeli tanımlar. Doymuş model, hiyerarşik modelde ilk modeldir.

Diğer hiyerarşik modellerde, eşitlik (3.2-1) de verilen doymuş modeldeki bazı  $\lambda$ -parametrelerinin 0'a eşit kabul edilmesi ile elde edilir. Hiyerarşiklik tanımı ile, eğer s üst indisli herhangi  $\lambda$ -parametresi (o) kabul edildiğinde, o zaman s yi içeren aynı veya daha yüksek derecedeki her  $\lambda$ -parametrelerinin de (o) olduğu varsayımı yapılmaktadır. Karşıt olarak eğer herhangi  $\lambda$ -parametresi sıfır değilse onun daha alt derecedeki akrabası da log-linear modelde

bulunmalıdır. Bu esasa göre, eğer  $\lambda_{ij}^{AB} = 0$  ise  $\lambda_{ijk}^{ABC} = 0$  olmalıdır. Aynı şekilde,  $\lambda_{ik}^{AC}$  modelde varsa,  $\lambda_i^A$  ve  $\lambda_k^C$  de modelde bulunmalıdır. Bu durum hiyerarşi prensibini tanımlar. Örneğin,  $\lambda_{ijk}^{ABC}$ 'nin (0) olduğu farzedilebilir. O zaman üçlü-faktör etkisi olmayan indirgenmiş bir modele sahip olunur. Bu model

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$$

dır. Bir alt hiyerarşik modelde ise, tüm i ve j'ler için  $\lambda_{ij}^{AB}$  ve  $\lambda_{ijk}^{ABC}$ 'nin (0) olduğu varsayıldığında, o zaman model

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$$

olur. Bu indirgemede mutlaka hiyerarşiklik esasına uyulması gerekir. Örneğin

$$\log m_{ijk} = \lambda + \lambda_{ij}^{AB} + \lambda_k^C$$

şeklindeki bir indirgeme ile elde edilen model hiyerarşik log-linear model olmaz. Çünkü, bu son modelde  $\lambda_i^A = 0$  varsayıldığından  $\lambda_{ij}^{AB}$  de sıfır olmalıydı. Oysa bu model  $\lambda_{ij}^{AB}$  yi bulundurmamaktadır. Dolayısıyla hiyerarşi prensibi bozulmaktadır.

Hiyerarşik modeli ilginç duruma getiren çeşitli faktörler vardır. BIRCH (1963) ve GOODMAN (1970)'in belirttiği gibi, üçlü-faktör interaksyonu bulunmayan model hariç olmak üzere, üç-yönlü tablolar için her hiyerarşik model bir veya daha fazla bağımsızlık, şartlı bağımsızlık veya eşit olasılık hipotezlerine eşdeğerdir. Üçlü-faktör interaksyonu olmayan model, birkaç iki yönlü tabloda,

çarpaz-çarpım oranlarının karşılaştırılmasında kullanılan terimler cinsinden yorumlanabilir. Böylece hiyerarşik modeller basit yorumlara sahiptirler. Hiyerarşik modellerde bazı uygun hesaplamalar da vardır. Üçlü-faktör interaksyonu bulunmayan model hariç tutulursa, hiyerarşik modeller için enbüyük-olabilirlik tahminleri ve düzeltilmiş rezidü-yaller iterative hesaplama gerekmeden bulunabilir. Üçlü-faktör interaksyonu bulunmayan modelde iterative oransal uygunluk algoritması, enbüyük-olabilirlik tahminlerini hesaplamada Newton-Raphson algoritmasına alternatif teşkil eder. Newton-Raphson algoritması kullanılsa dahi, bazı sadeleştirmeler yapmada fayda vardır. Böylece hiyerarşik modeller, üç-yönlü tablolar için cazip bir model ailesi oluşturur ve <sup>(3.3)</sup> bölümde görüleceği gibi daha fazla-yönlü tablolara kolayca genelleştirilebilirler.

### 3.2.1. 2x2x2 Boyutlu Tablolar

Çizelge 6- 2x2x2 Boyutlu Gözlenen Değerler Tablosu.

		C		$\Sigma$	
A	B	1	2		
1	1	$n_{111}$	$n_{112}$	$n_{11}^B$	$n_1^A$
	2	$n_{121}$	$n_{122}$	$n_{12}^B$	
2	1	$n_{211}$	$n_{212}$	$n_{21}^B$	$n_2^A$
	2	$n_{221}$	$n_{222}$	$n_{22}^B$	
$\Sigma$		$n_1^C$	$n_2^C$	N	

$$n_1^B = n_{11}^B + n_{21}^B$$

$$n_2^B = n_{12}^B + n_{22}^B$$



Burada,  $n_{ijk}$ ,  $A_h=i$ ,  $B_h=j$  ve  $C_h=k$  olan bireylerin sayısıdır.  $2 \times 2 \times 2$  tablosu  $r \times s \times t$  tablosuna karşılık gelen bir pozisyona sahiptir. Ancak  $2 \times 2 \times 2$  tablosunun tanımlanması ve analizi genel  $r \times s \times t$  tablosundan bir dereceye kadar daha basittir. Bununla beraber temel esaslar fazla farklı değildir.  $2 \times 2 \times 2$  tablolarındaki kolaylık sınırlı sayıda logaritmik riskler ve logaritmik çapraz-çarpım oranları kullanılarak tablonun tanımlanabilir oluşudur. Başka bir kolaylığı da üçlü-faktör etkisi bulunmayan modelle ilgili bir kolaylıktır.  $2 \times 2 \times 2$  tabloları için çok basit bir Newton-Raphson algoritması vardır. Ancak  $r \times s \times t$  tabloları için bu algoritma çok daha az kullanışlıdır.

### 3.2.1.1. Nisbi Risklerin Karşılaştırılması

A'nın  $i$ . ve B'nin  $j$ . kategorileri içerisinde C'nin 2. kategorisinden ziyade 1. kategorisinden olmanın nisbi riski  $m_{ij1}/m_{ij2}$  dir. Logaritmik nisbi risk ise

$$\log (m_{ij1}/m_{ij2}) = \tau_{12.ij}^{C.AB}$$

ile gösterilir. Burada 1 ve 2, A değişkeninin  $i$ . değeri ve B değişkeninin  $j$ . değeri verildiğinde C değişkeninin 1 ve 2 değeri için nisbi risklerin karşılaştırıldığını gösterir.  $\tau_{12.ij}^{C.AB}$  nin tahmin edilen değeri

$$\hat{\tau}_{12.ij}^{C.AB} = \log (n_{ij1}/n_{ij2})'$$

dir. Bu, yaklaşık normal dağılım gösterir. Dağılımın asimtotik ortalaması  $\tau_{12.ij}^{C.AB}$  ve asimtotik varyansı,

$$\sigma^2(\hat{\tau}_{12.ij}^{C.AB}) = \frac{1}{m_{ij1}} + \frac{1}{m_{ij2}}$$

dir. Tahmin edilen asimtotik varyans (EASD) ise,

$$s^2(\hat{z}_{12.ij}^{C.AB}) = \frac{1}{n_{ij1}} + \frac{1}{n_{ij2}}$$

olur. Yaklaşımı iyileştirmek için  $n_{ijk}$ 'lara 1/2 değeri eklenebilir. Bununla beraber  $n_{ijk}$  gözlemlerinin büyük değerler aldığı durumda bu düzeltmenin etkisi ihmal edilir.

$\hat{z}_{12.ij}^{C.AB}$  için formüller 3.1 bölümdeki  $z_{12j}^{A.B}$  formüllerine benzemektedir.

Hem A hemde B değişkenine ait varyasyonlar mevcut olduğunda ve A'daki değişkenliğin B'dekinden daha büyük olduğu görülsün. İlk olarak A değişkenine ait varyasyon dikkate alınır.  $B_h$ 'ın j. kategorisi için, A'nın 1 ve 2. kategorilerine ait logaritmik nisbi risklerin farkı

$$\begin{aligned} z_{(12)(12).j}^{AC.B} &= \log \left( \frac{m_{1j1} m_{2j2}}{m_{1j2} m_{2j1}} \right) \\ &= \log \left( \frac{m_{1j1}}{m_{1j2}} \right) - \log \left( \frac{m_{2j1}}{m_{2j2}} \right) \\ &= z_{12.1j}^{C.AB} - z_{12.2j}^{C.AB} \end{aligned}$$

A'nın 1. ve 2. kategorisi için logaritmik nisbi riski verir. Bunun tahmini,

$$\begin{aligned} \hat{z}_{(12)(12).j}^{AC.B} &= \log \left( \frac{n_{1j1} n_{2j2}}{n_{1j2} n_{2j1}} \right) \\ &= \log \left( \frac{n_{1j1}}{n_{1j2}} \right) - \log \left( \frac{n_{2j1}}{n_{2j2}} \right) \end{aligned}$$

$$= \hat{\tau}_{12.1j}^{C.AB} - \hat{\tau}_{12.2j}^{C.AB}$$

olur. Bu tahminin asimtotik ortalaması  $\tau_{(12)(12).j}^{AC.B}$  ve asimtotik varyansı

$$\begin{aligned} \sigma^2(\hat{\tau}_{(12)(12).j}^{AC.B}) &= \frac{1}{m_{1j1}} + \frac{1}{m_{1j2}} + \frac{1}{m_{2j1}} + \frac{1}{m_{2j2}} \\ &= \sigma^2(\hat{\tau}_{12.1j}^{C.AB}) + \sigma^2(\hat{\tau}_{12.2j}^{C.AB}) \end{aligned}$$

olur ve yaklaşık normal dağılım gösterir.

$\tau_{(12)(12).j}^{AC.B}$  sembolü ile ifade edilmek istenen kavram, logaritmik çapraz-çarpım oranının, A değişkeninin ve C değişkeninin 1 ve 2. değerleri alınarak hesaplandığını göstermektedir. Ancak bu logaritmik çapraz-çarpım oranı, B değişkeninin j. değeri üzerine şartlı olarak alınmaktadır.  $\hat{\tau}_{(12)(12).j}^{AC.B}$  ifadesi, A ve C değişkeninin karşılıklı sınıflandığı 2x2 tablosunda B değişkeninin j. değeri için tahmin edilen logaritmik çapraz-çarpım oranının tahminidir.

Eğer  $\tau_{(12).11}^{C.AB} = \tau_{(12).21}^{C.AB}$  ise, o zaman

$$\tau_{(12)(12).1}^{AC.B} = 0 \text{ olur. Böylece } \hat{\tau}_{(12)(12).j}^{AC.B} / s(\hat{\tau}_{(12)(12).j}^{AC.B})$$

standardize değeri yaklaşık  $N(0,1)$  dağılımı gösterir. Bu değerle  $H_0: \tau_{(12)(12).j}^{AC.B} = 0$  hipotezi test edilir. Farkın büyüklüğü değişkendir. % 95 lik yaklaşım güven aralığı bulunabilir.

B değişkeninin 1. ve 2. kategorileri arasındaki karşılaştırmaları açıklamak için A değişkeninin i. değeri

için nisbi riskler arasındaki farklar ele alınır. Bu fark,

$$\begin{aligned}\tau_{(12)(12).i}^{BC.A} &= \log \left( \frac{m_{i11} m_{i22}}{m_{i21} m_{i12}} \right) \\ &= \log \left( \frac{m_{i11}}{m_{i21}} \right) - \log \left( \frac{m_{i12}}{m_{i22}} \right) \\ &= \tau_{12.i1}^{C.AB} - \tau_{12.i2}^{C.AB}\end{aligned}$$

dir. Bunun tahmin edicisi,

$$\begin{aligned}\hat{\tau}_{(12)(12).i}^{BC.A} &= \log \left( \frac{n_{i11} n_{i22}}{n_{i21} n_{i12}} \right) \\ &= \log \left( \frac{n_{i11}}{n_{i21}} \right) - \log \left( \frac{n_{i12}}{n_{i22}} \right) \\ &= \hat{\tau}_{12.i1}^{C.AB} - \hat{\tau}_{12.i2}^{C.AB}\end{aligned}$$

dir. Bu tahminin asimtotik ortalaması  $\tau_{12.i1}^{BC.A}$  ve asimtotik varyansı

$$\begin{aligned}\sigma^2(\hat{\tau}_{(12)(12).i}^{BC.A}) &= \frac{1}{m_{i11}} + \frac{1}{m_{i21}} + \frac{1}{m_{i12}} + \frac{1}{m_{i22}} \\ &= \sigma^2(\hat{\tau}_{12.i1}^{C.AB}) + \sigma^2(\hat{\tau}_{12.i2}^{C.AB})\end{aligned}$$

dir. Dağılışı ise yaklaşık normal dağılıştır. Tahmin edilen asimtotik varyans,

$$s^2(\hat{\tau}_{(12)(12).i}^{BC.A}) = \frac{1}{n_{i11}} + \frac{1}{n_{i21}} + \frac{1}{n_{i12}} + \frac{1}{n_{i22}}$$

dir.  $\hat{\tau}_{(12)(12).i}^{BC.A}$  için % 95'lik yaklaşık güven aralığı bulunabilir

$$\tau_{(12)(12).1}^{AC.B} = \tau_{(12)(12).2}^{AC.B}$$

eşitliğinin sağ tarafıda diğer tarafa taşınırsa

$$\tau_{(12)(12).1}^{AC.B} - \tau_{(12)(12).2}^{AC.B} = 0$$

elde edilir. Bunun beklenen değerler cinsinden yazılımı neticesinde  $\tau_{(12)(12)(12)}^{ABC}$  ye eşit olduğu görülür. Şöyleki,

$$\tau_{(12)(12)(12)}^{ABC} = \log \left( \frac{m_{111} m_{212} m_{122} m_{221}}{m_{112} m_{211} m_{121} m_{222}} \right) = 0$$

yani,

$$\tau_{(12)(12).1}^{BC.A} - \tau_{(12)(12).2}^{BC.A} = \tau_{(12)(12)(12)}^{ABC}$$

olduğundan  $A_h = i$  verildiğinde B değişkeninin 1 ve 2. kategorisi için logaritmik nisbi riskteki  $\tau_{(12)(12).i}^{BC.A}$  farkı  $i=1$  veya  $i=2$  için aynıdır. Böylece logaritmik nisbi riskteki farkı tek bir değer ile ifade edebiliriz.

$\hat{\tau}_{(12)(12)(12)}^{ABC}$  sınırlaması,

$$\begin{aligned} \hat{\tau}_{(12)(12)(12)}^{ABC} &= \log \left( \frac{n_{111} n_{212} n_{122} n_{221}}{n_{112} n_{211} n_{121} n_{222}} \right) \\ &= \hat{\tau}_{(12)(12).1}^{AC.B} - \hat{\tau}_{(12)(12).2}^{AC.B} \end{aligned}$$

$$= \hat{\tau}_{(12)(12).1}^{BC.A} - \hat{\tau}_{(12)(12).2}^{BC.A}$$

ile tahmin edilir. Bu tahminin asimtotik ortalaması

$\tau_{(12)(12)(12)}^{ABC}$  ve asimtotik varyansı

$$\begin{aligned} \sigma^2(\hat{\tau}_{(12)(12)(12)}^{ABC}) &= \sum_i \sum_j \sum_k \frac{1}{m_{ijk}} \\ &= \frac{1}{m_{111}} + \frac{1}{m_{112}} + \frac{1}{m_{211}} + \frac{1}{m_{212}} \\ &\quad + \frac{1}{m_{121}} + \frac{1}{m_{122}} + \frac{1}{m_{221}} + \frac{1}{m_{222}} \\ &= \sigma^2(\hat{\tau}_{(12)(12).1}^{AC.B}) + \sigma^2(\hat{\tau}_{(12)(12).2}^{AC.B}) \\ &= \sigma^2(\hat{\tau}_{(12)(12).1}^{BC.A}) + \sigma^2(\hat{\tau}_{(12)(12).2}^{BC.A}) \end{aligned}$$

dır. Dağılışı ise yaklaşık normal dağılışı gösterir. Tahmin edilen asimtotik varyans (EASD)

$$s^2(\hat{\tau}_{(12)(12)(12)}^{ABC}) = \sum_i \sum_j \sum_k \frac{1}{n_{ijk}}$$

dır. Eğer  $\tau_{(12)(12)(12)}^{ABC}$ , sıfıra eşit ise,  $\tau_{(12)(12).1}^{AC.B}$  ve

$\tau_{(12)(12).2}^{AC.B}$  birbirine eşit olacağından ortak bir  $\hat{\tau}_{(12)(12)}^{AC}$  tahmini ve  $\tau_{(12)(12).1}^{BC.A}$  ve  $\tau_{(12)(12).2}^{BC.A}$  için de yine ortak bir  $\hat{\tau}_{(12)(12)}^{BC}$  tahmini bulunur.

Diğer taraftan C değişkeninin belli kategorilerinde A ve B değişkeninin ilişkisi araştırılabilir. Bunun için

$$\begin{aligned}\tau_{(12)(12).k}^{AB.C} &= \log \left( \frac{m_{11k} m_{22k}}{m_{21k} m_{12k}} \right) \\ &= \log m_{11k} - \log m_{21k} - \log m_{12k} \\ &\quad + \log m_{22k}\end{aligned}$$

şeklinde ifade edilen logaritmik çapraz çarpım oranı dik-kate alınır. Bu oran C değişkeni yönünden k. kategorideki bir bireyin  $A_h = 2$  den ziyade  $A_h = 1$ 'de olması durumuna ait  $B_h = 1$  ve  $B_h = 2$  için nisbi riskleri karşılaştırır. Üçlü-faktör interaksyonu olmayan model altında,

$$\tau_{(12)(12).1}^{AB.C} = \tau_{(12)(12).2}^{AB.C} = \tau_{(12)(12)}^{AB(C)}$$

olur. Ortak tahmin edici olan  $\tau_{(12)(12)}^{AB(C)}$  nin en büyük-olabilirlik tahmini,

$$\hat{\tau}_{(12)(12)}^{AB(C)} = \log \left( \frac{\hat{m}_{111} \hat{m}_{221}}{\hat{m}_{211} \hat{m}_{121}} \right) = \log \left( \frac{\hat{m}_{112} \hat{m}_{222}}{\hat{m}_{212} \hat{m}_{122}} \right)$$

dir. Bu tahminin asimtotik varyansına ait tahmin, iki ayrı varyans tahmininin harmonik ortalaması ile bulunur.

$$s^2 \left( \hat{\tau}_{(12)(12)}^{AB(C)} \right) = \left[ \sum_{k=1}^t \left( \frac{1}{\hat{m}_{11k}} + \frac{1}{\hat{m}_{21k}} + \frac{1}{\hat{m}_{12k}} + \frac{1}{\hat{m}_{22k}} \right)^{-1} \right]^{-1}$$

dir. Eğer,  $B_h$  verildiğinde diğer iki değişken için ortak tahminin varyansının bulunması isteniyorsa toplam, j üzerinden alınır ve  $m_{1j1}$  vs şeklinde paydalar düzenlenir. Aynı şekilde  $A_h$  verildiğinde bu işlem i üzerinden yapılır.  $k=1$  için

$$v_1 = \frac{1}{\hat{m}_{111}} + \frac{1}{\hat{m}_{211}} + \frac{1}{\hat{m}_{121}} + \frac{1}{\hat{m}_{221}}$$

ve  $k = 2$  için

$$v_2 = \frac{1}{\hat{m}_{112}} + \frac{1}{\hat{m}_{212}} + \frac{1}{\hat{m}_{122}} + \frac{1}{\hat{m}_{222}}$$

olduğundan

$$S(\hat{\tau}_{(12)(12)}^{AB(C)}) = \left( \frac{1}{v_1} + \frac{1}{v_2} \right)^{-1/2}$$

olur. Test istatistiği ise  $\hat{\tau}_{(12)(12)}^{AB(C)} / S(\hat{\tau}_{(12)(12)}^{AB(C)})$  şeklinde ifade edilir ve  $\hat{\tau}_{(12)(12)}^{AB(C)}$  nin sıfırdan farklılığı test edilir.  $\hat{\tau}_{(12)(12)}^{AB(C)}$  için veya doğrudan nisbi risk

$$q_{(12)(12)}^{AB(C)} = \frac{m_{111} m_{221}}{m_{211} m_{121}} = \frac{m_{112} m_{222}}{m_{212} m_{122}}$$

için

% 95'lik yaklaşık güven aralığı bulunabilir.

### 3.2.1.2. Üçlü-Faktör İnteraksiyonu Olmayan Model

$\tau_{(12)(12)(12)}^{ABC}$  teriminin sıfır olduğu model, üçlü-faktör interaksiyonu bulunmayan log-linear model olarak adlandırılır. Model log-linear bir modeldir. Çünkü  $m_{ijk}$  beklenen değerlerinin logaritmaları olan  $\mu_{ijk} = \log m_{ijk}$  linear kısıtlamaya maruz kalırlar. Bu linear kısıtlama

$$\tau_{(12)(12)(12)}^{ABC} = 0 \text{ için}$$



$$\begin{aligned} & m_{111}^- / m_{211}^- / m_{121}^+ / m_{221}^- / m_{112}^+ / m_{212}^+ / m_{122} \\ & - m_{222} = 0 \end{aligned}$$

şeklinde ifade edilir. Modelin adı genellikle varyans analizinde görülen toplam varyasyonun parçalanmasından ortaya çıkar. Olabilirlik tablolarının analizinde bu ayrışmanın doğrudan kullanımının BIRCH (1963)'e dayandığı gözükür.  $m_{ijk}$  ortalamalarının tümü pozitif olduğunda,

$$\begin{aligned} \mu_{ijk} = \log m_{ijk} = & \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} \\ & + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC} \end{aligned}$$

yazılabilir.

2x2x2'lik bir log-linear modelde i, j, k'nın 1. seviyeleri için +1 ve 2. seviyeleri için -1 olacak şekilde  $\lambda$ -terimlerinin katsayıları Çizelge 7'deki gibi tesbit edilir.

Çizelge 7- Üçyönlü Tablolar İçin Doymuş Modelin  $\lambda$ -Terimlerine Ait Katsayılar

Hücreler	$\lambda$	$\lambda_i^A$	$\lambda_j^B$	$\lambda_k^C$	$\lambda_{ij}^{AB}$	$\lambda_{ik}^{AC}$	$\lambda_{jk}^{BC}$	$\lambda_{ijk}^{ABC}$
1,1,1	+	+	+	+	+	+	+	+
2,1,1	+	-	+	+	-	-	+	-
1,2,1	+	+	-	+	-	+	-	-
2,2,1	+	-	-	+	+	-	-	+
1,1,2	+	+	+	-	+	-	-	-
2,1,2	+	-	+	-	-	+	-	+
1,2,2	+	+	-	-	-	-	+	+
2,2,2	+	-	-	-	+	+	+	-

Doymuş modelde,

$$\begin{aligned}
 \sum_i \lambda_i^A &= \sum_j \lambda_j^B = \sum_k \lambda_k^C = \sum_j \lambda_{ij}^{AB} = \sum_i \lambda_{ij}^{AB} = \sum_k \lambda_{ik}^{AC} \\
 &= \sum_i \lambda_{ik}^{AC} = \sum_k \lambda_{jk}^{BC} = \sum_j \lambda_{jk}^{BC} = \sum_k \lambda_{ijk}^{ABC} \\
 &= \sum_j \lambda_{ijk}^{ABC} = \sum_i \lambda_{ijk}^{ABC} = 0
 \end{aligned}$$

dır. Buna göre,  $\lambda$ -terimleri,  $2 \times 2 \times 2$  tablosunda  $m_{ijk}$  cinsinden yazılacak olursa,

$$\lambda = \frac{1}{8} (m_{111} + m_{211} + m_{121} + m_{221} + m_{112} + m_{212} + m_{122} + m_{222}),$$

$$\lambda_1^A = -\lambda_2^A = \frac{1}{8} (m_{111} - m_{211} + m_{121} - m_{221} + m_{112} - m_{212} + m_{122} - m_{222})$$

$$\sum_i \lambda_i^A = 0, \lambda_1^A + \lambda_2^A = 0, \lambda_1^A = -\lambda_2^A$$

$$\lambda_1^B = -\lambda_2^B = \frac{1}{8} (m_{111} + m_{211} - m_{121} - m_{221} + m_{112} + m_{212} - m_{122} - m_{222})$$

$$\sum_j \lambda_j^B = 0, \lambda_1^B + \lambda_2^B = 0, \lambda_1^B = -\lambda_2^B$$

$$\lambda_1^C = -\lambda_2^C = \frac{1}{8} (m_{111} + m_{211} + m_{121} + m_{221} - m_{112} - m_{212} - m_{122} - m_{222})$$

$$\sum_k \lambda_k^C = 0, \quad \lambda_1^C + \lambda_2^C = 0, \quad \lambda_1^C = -\lambda_2^C$$

$$\lambda_{11}^{AB} = -\lambda_{21}^{AB} = -\lambda_{12}^{AB} = \lambda_{22}^{AB}$$

$$= \frac{1}{8} (m_{111} - m_{211} - m_{121} + m_{221} + m_{112} - m_{212} - m_{122} + m_{222})$$

$$\lambda_{11}^{AC} = -\lambda_{21}^{AC} = -\lambda_{12}^{AC} = \lambda_{22}^{AC}$$

$$= \frac{1}{8} (m_{111} - m_{211} + m_{121} - m_{221} - m_{112} + m_{212} - m_{122} + m_{222})$$

$$\lambda_{11}^{BC} = -\lambda_{21}^{BC} = -\lambda_{12}^{BC} = \lambda_{22}^{BC}$$

$$= \frac{1}{8} (m_{111} + m_{211} - m_{121} - m_{221} - m_{112} - m_{212} + m_{122} + m_{222})$$

ve

$$\lambda_{111}^{ABC} = -\lambda_{211}^{ABC} = \lambda_{221}^{ABC} = -\lambda_{121}^{ABC}$$

$$= -\lambda_{112}^{ABC} = -\lambda_{212}^{ABC} = \lambda_{122}^{ABC} = -\lambda_{222}^{ABC}$$

$$= \frac{1}{8} (m_{111} - m_{211} - m_{121} + m_{221} - m_{112} + m_{212} + m_{122} - m_{222})$$

$$= \frac{1}{8} \tau_{(12)(12)(12)}^{ABC}$$

olur.

Bu parçalamada  $\lambda_{ijk}^{ABC}$ , ler üçlü-faktör interaksionları bileşkeleri olarak isimlendirilirler. Böylece  $\tau_{(12)(12)(12)}^{ABC} = 0$  varsayımı,  $\lambda_{ijk}^{ABC}$  üçlü-faktör interaksionlarının sıfır varsayılmasına eşdeğerdir. Yani modelin sadece

$$\mu_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$$

olması durumu söz konusudur.

2x2x2 boyutlu tablolarda üçlü-faktör interaksionu bulunmayan modelin belirgin özellikleri şöyle açıklanabilir: 2x2x2 boyutlu tabloların analizini üçlü-faktör interaksionu olmayan model denen bir hiyerarşik model aracılığıyla açıkladık. Belirgin birkaç özellik ortaya çıkmaktadır. Genel rxsxt tablolarındaki duruma benzer olarak, üçlü-faktör interaksionu bulunmayan model iki tane 2x2 tablosu için çapraz-çarpım oranlarının eşitliği modeline indirgenebildiği için fevkalade basit bir yoruma sahiptir. Newton-Raphson algoritması 2x2x2 tablolarında son derece basittir. 2x2x2 tablolarında böyle bir modelin testinde  $x^2$ 'ye ait serbestlik derecesi bir olduğundan rezidual analizi önemsizdir. Üçlü-faktör interaksionu bulunmayan/üç-yönlü tablolar için uydurulabilecek diğer hiyerarşik modellerden daha özel bir durumu vardır. Çünkü beklenen değerlerin enbüyük-olabilinlik tahminlerini bulmak için iterative hesaplamalar gerekir, diğerlerinde ise böyle bir gereksinme yoktur.

### 3.2.2. Üç-Yönlü rxsxt Boyutlu Tablolar İçin Mümkün Hiyerarşik Modeller

Daha özel tablolarla ilgili detay incelemeleri yapmadan, genel olarak üç boyutlu tablolarda mümkün olan hiyerarşik modeller nelerdir ve bunların genel yorumu nasıl yapılıyor? Bu hususları açıklayalım.

$n_{ijk}$  elemanlı ve  $m_{ijk}$  beklenen değerlerine sahip rxsxt boyutlu bir tablo dikkate alındığında,  $1 \leq i \leq r$ ,  $1 \leq j \leq s$ ,  $1 \leq k \leq t$ ,  $n_{ijk}$ ,  $A_h = i$ ,  $B_h = j$  ve  $C_h = k$  olan gözlem sayısıdır, ve  $(A_h, B_h, C_h)$ ,  $1 \leq h \leq N$ , bağımsız ve benzer dağılımlı üçlü değişken grubudur.  $p_{ijk} > 0$  olasılıkları,  $A_h = i$ ,  $B_h = j$  ve  $C_h = k$  olması olasılığıdır.  $A_h$ 'ın 1'den r'ye,  $B_h$ 'ın 1'den s'ye ve  $C_h$ 'ın 1'den t'ye kadar tamsayı değerler aldığı varsayılır.

Burada, N sabit farzedilmektedir, bu nedenle de  $\mathbf{n}$  tablosu, örnek büyüklüğü N ve olasılıkları  $p_{ijk}$  olan multinomial dağılımı gösterir.

Diğer örnekleme işlemleri de dikkate alınabilir. Marginal tablolarla ilgili bazı basit eşitlikler birçok defa kullanılır.  $A_h = i$  olan fertlerin sayısı  $n_i^A$  değeri  $A_h = i$  ve  $B_h = j$ ,  $1 \leq i \leq r$ ,  $1 \leq j \leq s$  için, olan fertlerin sayıları toplamıdır.

$$n_i^A = \sum_{j=1}^s n_{ij}^{AB}$$

dir. Benzer şekilde,

$$n_i^A = \sum_{k=1}^t n_{ik}^{AC} = \sum_j \sum_k n_{ijk},$$

$$n_{ij}^{AB} = \sum_{k=1}^t n_{ijk}$$

$$N = \sum_{i=1}^r n_i^A = \sum_{i=1}^r \sum_{j=1}^s n_{ij}^{AB} = \sum_{i=1}^r \sum_{j=1}^s \sum_{k=1}^t n_{ijk}$$

dır.

GOODMAN (1970)'e göre hiyerarşik modeller dokuz sınıfa ayrılır. Bunlardan özellikle ilk beşi daha önemlidir.

1. Doymuş model: Burada  $\log m_{ijk}$ 'lar üzerine hiç bir kısıtlama konulmaz, dolayısıyla de model,

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

olur. Burada,

$$\begin{aligned} \sum_i \lambda_i^A &= \sum_j \lambda_j^B = \sum_k \lambda_k^C = \dots = \sum_i \lambda_{ijk}^{ABC} = \sum_j \lambda_{ijk}^{ABC} \\ &= \sum_k \lambda_{ijk}^{ABC} = 0 \end{aligned}$$

dir. Bu modelde "generating" sınıf (modeli tanımlayan sınıf)  $\mathcal{E}$ , ABC'dir. Modeli tanımlayan sınıfın tanımı yapılırken üzerinde durulan husus: Bu sınıfın bulundurduğu harfleri taşıyan ve bu sınıfın kendi içerisinde mevcut harflerden meydana gelebilecek harf gruplarını taşıyan parametreleri ihtiva eden model bu sınıf tarafından tanımlanır. Örneğin: Modeli tanımlayan sınıf ABC'dir, denildiği zaman, modelde A, B, C, AB, AC, BC ve ABC harflerini taşıyan parametreler mevcut olacak demektir. Dolayısıyla de

$$\hat{m}_{ijk} = \hat{m}_{ijk}^{ABC} = n_{ijk}$$

olur.

2. Üçlü-faktör interaksyonu olmayan model: Burada model,

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$$

dır. Daha önce belirtildiği gibi,  $\mathcal{E}$  (modeli tanımlayan sınıf) AB, AC ve BC'den meydana gelir.

$$\hat{m}_{ij}^{AB} = m_{ij}^{AB}$$

$$\hat{m}_{ik}^{AC} = n_{ik}^{AC}$$

$$\hat{m}_{jk}^{BC} = n_{jk}^{BC}$$

tahminleri yukardaki modelde yerine konursa

$$\log \hat{m}_{ijk} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_j^B + \hat{\lambda}_k^C + \hat{\lambda}_{ij}^{AB} + \hat{\lambda}_{ik}^{AC} + \hat{\lambda}_{jk}^{BC}$$

elde edilir. Burada  $\hat{\lambda}$ - parametreleri,  $\lambda$ - parametreleriyle aynı kısıtlamaları sağlar. Bu eşitliklerin çözümü Newton-Raphson veya iterative oransal uydurma algoritması gibi iterative metodları gerektirir.

Model, ancak,  $C_n$  verildiğinde  $A_n$  ve  $B_n$  interaksyonunu ölçen şartlı çapraz-çarpım oranlarının

$$q(ii') (jj').k = \frac{m_{ijk} m_{i'j'k}}{m_{ij'k} m_{i'jk}}$$

k'dan bağımsız olması durumunda sağlanır. Bu durumda loga-

ritmik çapraz-çarpım oranı,

$$\begin{aligned} \tau_{(ii')(jj')(kk')}^{ABC} &= \log \left( \frac{m_{ijk} m_{i'j'k} m_{ij'k'} m_{i'jk'}}{m_{ijk'} m_{i'jk} m_{ij'k} m_{i'j'k'}} \right) \\ &= \tau_{(ii')(jj').k}^{AB.C} - \tau_{(ii')(jj').k'}^{AB.C} \\ &= 0 \end{aligned}$$

olur.

3. Şartlı bağımsızlık: Bu sınıf üç model içerir. Üç tanımlayan sınıf vardır. Bunlar, AB ve AC, AB ve BC ve AC ve BC den ibarettir. AB ve AC tanımlayan sınıfının kullanılması durumunda model,

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC}$$

olur. Böylece BC ve ABC interaksiyonları, yani modeldeki  $\lambda_{jk}^{BC}$  ve  $\lambda_{ijk}^{ABC}$  terimleri sıfır kabul edilir.

BIRCH (1963) de belirtildiği gibi, bu model  $A_n$  verildiğinde,  $B_n$  ve  $C_n$ 'in şartlı olarak bağımsız olması hipotezine eşdeğerdir. Burada şartlı bağımsızlığın sağlanması için,

$$p_{jk.i}^{BC.A} = p_{j.i}^{B.A} * p_{k.i}^{C.A}$$

olması gerekir. Bu gösterilişte,

$$p_{jk.i}^{BC.A} = p_{ijk} / p_i^A$$



olur.  $p_{ijk}$  hücre olasılıkları ve  $p_i^A$ ,  $A_h = i$  olmasının marginal olasılığıdır. Eşitliğin sağ tarafındaki şartlı olasılıklarda,  $p_{ij}^{AB}$ ,  $A_h = i$  ve  $B_h = j$  olmasının marginal olasılığı cinsinden tanımlanabilir.  $p_{jk.i}^{BC.A}$ ,  $A_h = i$  olduğu verildiğinde,  $B_h = j$  ve  $C_h = k$  olmasının şartlı olasılığıdır.

$$p_{j.i}^{B.A} = p_{ij}^{AB} / p_i^A ,$$

$A_h = i$  olduğu verildiğinde  $B_h = j$  olmasının şartlı olasılığı ve

$$p_{k.i}^{C.A} = p_{ik}^{AC} / p_i^A ,$$

$A_h = i$  olduğu verildiğinde  $C_h = k$  olmasının şartlı olasılığıdır.

Çalışılan hiyerarşik modeller ve şartlı bağımsızlık arasındaki ilişkiyi göstermek için

$$\begin{aligned} \tau(jj')(kk').i &= \log \left( \frac{m_{ijk} m_{ij'k'}}{m_{ijk'} m_{ij'k}} \right) \\ &= \log \left( \frac{p_{ijk} p_{ij'k'}}{p_{ijk'} p_{ij'k}} \right) \\ &= \log \left[ \frac{(p_{ijk}/p_i^A)(p_{ij'k'}/p_i^A)}{(p_{ijk'}/p_i^A)(p_{ij'k}/p_i^A)} \right] \\ &= \log \left( \frac{p_{jk.i}^{BC.A} p_{j'k'.i}^{BC.A}}{p_{jk'.i}^{BC.A} p_{j'k.i}^{BC.A}} \right) \end{aligned}$$

Çapraz-çarpım oranı dikkate alınır. Eğer parantez içindeki şartlı olasılıklar,

$$p_{j'k'.i}^{BC.A} = p_{j'.i}^{B.A} p_{k'.i}^{C.A}$$

şeklinde ifade edilebilirse, tüm  $\tau_{(jj')(kk').i}^{BC.A}$ 'ler sıfır olur.

$$\begin{aligned} \tau_{(jj')(kk').i}^{BC.A} &= \log m_{ijk} - \log m_{ijk'} - \log m_{ij'k} \\ &\quad + \log m_{ij'k'} \end{aligned}$$

$$\begin{aligned} &= (\lambda_{jk}^{BC} - \lambda_{jk'}^{BC} - \lambda_{j'k}^{BC} + \lambda_{j'k'}^{BC}) \\ &\quad + (\lambda_{ijk}^{ABC} - \lambda_{ijk'}^{ABC} - \lambda_{ij'k}^{ABC} \\ &\quad \quad + \lambda_{ij'k'}^{ABC}) \end{aligned}$$

olduğundan, eğer tüm BC interaksiyonları ( $\lambda_{j'k'}^{BC}$ ) ve ABC interaksiyonları ( $\lambda_{ij'k'}^{ABC}$ ) sıfır ise, o zaman  $\tau_{(jj')(kk').i}^{BC.A}$ 'ler sıfırdır. Bu da,  $A_h$  verildiğinde,  $B_h$  ve  $C_h$ 'ların şartlı olarak bağımsızlığını gösterir. Diğer taraftan, eğer tüm  $\tau_{(jj')(kk').i}^{BC.A}$ 'ler sıfır ise  $\lambda$ -parametreleri üzerine konan kısıtlamalar,

$$\frac{1}{rst} \sum_i \sum_j \sum_k \tau_{(jj')(kk').i}^{BC.A} = \lambda_{j'k'}^{BC} = 0$$

ve

$$\frac{1}{st} \sum_j \sum_k \tau_{(jj')(kk').i}^{BC.A} = \tilde{\lambda}_{j'k'}^{BC} + \lambda_{ij'k'}^{ABC} = 0$$

olduğunu gösterir. Böylece  $\lambda_{jk}^{BC}$  ve  $\lambda_{ijk}^{ABC}$ 'ler 0'dır,

ve model

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC}$$

olur.

$m_{ijk}$  nin enbüyük-olabilirlik tahmini olan  $\hat{m}_{ijk}$ ,

$$\hat{m}_{ijk} = n_{ij}^{AB} n_{ik}^{AC} / n_i^A = N \hat{p}_i^A \hat{p}_{j.i}^{B.A} \hat{p}_{k.i}^{C.A}$$

eşitliğini sağlar. Burada

$$\hat{p}_i^A = n_i^A / N$$

$$\hat{p}_{j.i}^{B.A} = n_{ij}^{AB} / n_i^A$$

ve

$$\hat{p}_{k.i}^{C.A} = n_{ik}^{AC} / n_i^A$$

dır. Bunu doğrudan yazacak olursak,

$$m_{ijk} = N p_{ijk} = N p_i^A p_{jk.i}^{BC.A} = N p_i^A p_{j.i}^{B.A} p_{k.i}^{C.A}$$

dır. Burada,  $\hat{p}_i^A$ ,  $A_h = i$  olan gözlemlerin oranıdır.  $\hat{p}_{j.i}^{B.A}$ ,

$A_h = i$  olduğu verildiğinde  $B_h = j$  olan gözlemlerin oranıdır.

$\hat{p}_{k.i}^{C.A}$ ,  $A_h = i$  olduğu verildiğinde  $C_h = k$  olan gözlemlerin oranıdır.

$\hat{m}_{ijk}$  'nin enbüyük-olabilirlik tahmini olması için

$$\hat{m}_{ij}^{AB} = n_{ij}^{AB}$$

$$\hat{m}_{ik}^{AC} = n_{ik}^{AC}$$

$$\log \hat{m}_{ijk} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_j^B + \hat{\lambda}_k^C + \hat{\lambda}_{ij}^{AB} + \hat{\lambda}_{ik}^{AC}$$

kısıtlamaları sağlanmalıdır.

İlk kısıtlamayı ispat için,  $\sum_k m_{ijk} = m_{ij}^{AB}$  olduğundan,

$$\begin{aligned} \sum_k n_{ij}^{AB} n_{ik}^{AC} / n_i^A &= (n_{ij}^{AB} / n_i^A) \sum_k n_{ik}^{AC} = (n_{ij}^{AB} / n_i^A) n_i^A \\ &= n_{ij}^{AB} \end{aligned}$$

olduğu belirtilir. İkinci kısıtlama da benzer şekilde ispat edilir.

$C_h$  verildiğinde  $A_h$  ve  $B_h$ 'ın,  $B_h$  verildiğinde  $A_h$  ve  $C_h$ 'ın şartlı bağımsızlığı aynı şekilde gösterilebilir.

Şartlı bağımsızlık modeli altında  $C_h$  verildiğinde  $A_h$  ve  $B_h$  için kısmi ilişki katsayısı

$$\tau_{(ii')(jj').k}^{AB.C} = \log \left( \frac{m_{ijk} m_{i'j'k}}{m_{ij'k} m_{i'jk}} \right)$$

$A_h$  ve  $B_h$  için marginal ilişki katsayısına

$$\tau_{(ii')(jj')}^{AB} = \log \left( \frac{m_{ij}^{AB} m_{i'j'}^{AB}}{m_{ij'}^{AB} m_{i'j}^{AB}} \right)$$

eşittir.

$$m_{ijk} = N p_{ij}^{AB} p_{ik}^{AC} / p_i^A$$

ve

$$m_{ij}^{AB} = N p_{ij}^{AB}$$

eşitlikleri yerine konduğunda

$$\tau_{(ii')(jj').k}^{ABC} = \log \left( \frac{p_{ij}^{AB} p_{i'j'}^{AB}}{p_{ij'}^{AB} p_{i'j}^{AB}} \right) = \tau_{(ii')(jj')}^{AB}$$

olur.

4. İki değişkenin üçüncüden bağımsız olması durumu (Kısmi bağımsızlık): 3. sınıfta olduğu gibi, burada da üç ayrı model tanımlayan sınıf vardır. Bunlar AB ve C, A ve BC veya AC ve B'den oluşur. Dolayısıyla de üç değişik model yazılabilir. Sonuçları tanımlamak için,  $\zeta$ 'nin AB ve C'den oluştuğunu kabul edersek, model,

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB}$$

olur. Bu model, ancak  $(A_h, B_h)$  çiftleri  $C_h$ 'dan bağımsız ise sağlanır. Eğer

$$p_{ijk} = p_{ij}^{AB} p_k^C$$

ise,  $C_h, (A_h, B_h)$ , çiftlerinden bağımsız olur. Eşdeğer olarak, her k ve k' için

$$\tau_{kk'.ij}^{C.AB} = \log \left( \frac{p_{k.ij}^{C.AB}}{p_{k'.ij}^{C.AB}} \right) = \log \left( \frac{m_{ijk}}{m_{ijk'}} \right),$$

i ve j den bağımsızdır. Burada,

$$p_{k.ij}^{C.AB} = p_{ijk} / p_{ij}^{AB}$$

dir. Eğer

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB}$$

ise, o zaman

$$\tau_{kk'.ij}^{C.AB} = \log m_{ijk} - \log m_{ijk'} = \lambda_k^C - \lambda_{k'}^C$$

i ve j den bağımsızdır. Böylece, bağımsızlık sağlanır. Diğer taraftan eğer  $\tau_{kk'.ij}^{C.AB}$ , i ve j den bağımsız ise, o zaman

$$\tau_{(jj')(kk').i}^{BC.A} = \tau_{kk'.ij}^{C.AB} - \tau_{kk'.ij'}^{C.AB}$$

$$= \log \left( \frac{m_{ijk} m_{ij'k'}}{m_{ijk'} m_{ij'k}} \right) = 0$$

olur. 3. sınıfta belirtildiği gibi  $\lambda_{jk}^{BC}$  ve  $\lambda_{ijk}^{ABC}$ 'lerin sıfır oldukları görülür. Benzer şekilde,

$$\tau_{(ii')(kk').j}^{AC.B} = 0$$

ve  $\lambda_{ik}^{AC}$  ve  $\lambda_{ijk}^{ABC}$ 'ler sıfırdır. Böylece

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB}$$

olur.

Enbüyük-olabilirlik eşitlikleri,

$$\hat{m}_{ij}^{AB} = n_{ij}^{AB}$$

$$\hat{m}_k^C = n_k^C$$

$$\log \hat{m}_{ijk} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_j^B + \hat{\lambda}_k^C + \hat{\lambda}_{ij}^{AB}$$

dir. Eşitlikler

$$\hat{m}_{ijk} = n_{ij}^{AB} n_k^C / N = N \hat{p}_{ij}^{AB} \hat{p}_k^C$$

ile sağlanır. Burada,

$$\hat{p}_{ij}^{AB} = n_{ij}^{AB} / N,$$

$A_h = i$  ve  $B_h = j$  olan gözlemlerin oranıdır, ve

$$\hat{p}_k^C = n_k^C / N,$$

$C_h = k$  olan gözlemlerin oranıdır.

$$\hat{\chi}_{kk',ij}^2 = \log \left( \frac{N \hat{p}_{ij}^{AB} \hat{p}_k^C}{N \hat{p}_{ij}^{AB} \hat{p}_{k'}^C} \right) = \log \left( \frac{\hat{p}_k^C}{\hat{p}_{k'}^C} \right),$$

$i$  ve  $j$  den bağımsızdır.

$$\hat{m}_{ij}^{AB} = \sum_k \hat{m}_{ijk} = (n_{ij}^{AB} / N) \sum_k n_k^C = (n_{ij}^{AB} / N) N = n_{ij}^{AB},$$

ve

$$\hat{m}_k^C = \sum_i \sum_j \hat{m}_{ijk} = (n_k^C / N) \sum_i \sum_j n_{ij}^{AB} = (n_k^C / N) N = n_k^C$$

dir.

3. sınıfta olduğu gibi,  $C_h$  verildiğinde  $A_h$  ve  $B_h$  için kısmi ilişki ölçümü

$$\tau_{(ii')(jj').k}^{AB} = \log \left( \frac{m_{ijk} m_{i'j'k}}{m_{ij'k} m_{i'jk}} \right)$$

ile,  $A_h$  ve  $B_h$  için, marginal ilişki ölçümü

$$\tau_{(ii')(jj')}^{AB} = \log \left( \frac{m_{ij}^{AB} m_{i'j'}^{AB}}{m_{ij'}^{AB} m_{i'j}^{AB}} \right)$$

eşittir. Bu iddia ile hem  $\tau_{(ii')(jj').k}^{AB.C}$  hem de  $\tau_{(ii')(jj')}^{AB}$

$$m_{ijk} = N p_{ij}^{AB} p_k^C$$

ve

$$m_{ij}^{AB} = N p_{ij}^{AB}$$

cinsinden yazılarak sonuçta,

$$\log \left( \frac{p_{ij}^{AB} p_{i'j'}^{AB}}{p_{ij'}^{AB} p_{i'j}^{AB}} \right)$$

elde edilir.

5. Tüm değişkenlerin karşılıklı olarak bağımsız olması durumu: Burada model tanımlayan sınıf A, B ve C den meydana gelir, ve

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C$$

olur. Buna göre tüm ikili-faktör ve üçlü-faktör interaksiyonları sıfır varsayılır. Bu model, ancak  $A_h$ ,  $B_h$  ve  $C_h$  karşılıklı bağımsız ise sağlanır. Bunu da



$$p_{ijk} = p_i^A p_j^B p_k^C$$

şeklinde ifade edebiliriz.

Bağımsızlık altında,

$$\begin{aligned} \log m_{ijk} &= \log N + \log p_i^A + \log p_j^B + \log p_k^C \\ &= \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C \end{aligned}$$

olduğu belirtilir, burada

$$\lambda = \log N + \frac{1}{r} \sum_f \log p_f^A + \frac{1}{s} \sum_g \log p_g^B + \frac{1}{t} \sum_h \log p_h^C,$$

$$\lambda_i^A = \log p_i^A - \frac{1}{r} \sum_f \log p_f^A,$$

$$\lambda_j^B = \log p_j^B - \frac{1}{s} \sum_g \log p_g^B$$

ve

$$\lambda_k^C = \log p_k^C - \frac{1}{t} \sum_h \log p_h^C$$

dir. Diğer taraftan, eğer

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C$$

ise o zaman

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB}$$

ve

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AC}$$

olması doğrudur. Bundan dolayı, 4. sınıftaki neticeler sağlanırsa

$$p_{ijk} = p_{ij}^{AB} p_k^C = p_{ik}^{AC} p_j^B$$

olduğu izlenir. Buradan,

$$p_{ij}^{AB} = \sum_k p_{ijk} = p_j^B \sum_k p_{ik}^{AC} = p_i^A p_j^B$$

olduğundan,

$$p_{ijk} = p_i^A p_j^B p_k^C$$

yazılabilir.

Enbüyük-olabilirlik tahmini,

$$\hat{m}_{ijk} = n_i^A n_j^B n_k^C / N^2 = N \hat{p}_i^A \hat{p}_j^B \hat{p}_k^C$$

dir. Burada

$$\hat{p}_i^A = n_i^A / N,$$

$$\hat{p}_j^B = n_j^B / N,$$

ve

$$\hat{p}_k^C = n_k^C / N$$

dir.

$$\hat{m}_i^A = n_i^A,$$

$$\hat{m}_j^B = n_j^B,$$

$$\hat{m}_k^C = n_k^C$$

ve

$$\log \hat{m}_{ijk} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_j^B + \hat{\lambda}_k^C$$

eşitlikleri sağlanır.

$$\text{Eğer } \hat{m}_j^B = n_j^B, \hat{m}_k^C = n_k^C \text{ ve } \hat{m}_i^A = n_i^A$$

ve

$$\hat{\lambda}_i^A = \log n_i^A - \frac{1}{s} \sum_{i'} \log n_{i'}^A,$$

$$\hat{\lambda}_j^B = \log n_j^B - \frac{1}{s} \sum_{j'} \log n_{j'}^B,$$

$$\hat{\lambda}_k^C = \log n_k^C - \frac{1}{t} \sum_{k'} \log n_{k'}^C$$

ve

$$\hat{\lambda} = \frac{1}{r} \log n_{i'}^A + \frac{1}{s} \sum_{j'} \log n_{j'}^B + \frac{1}{t} \sum_{k'} \log n_{k'}^C - 2 \log N$$

ise, o zaman

$$\begin{aligned} \log \hat{m}_{ijk} &= \log n_{i'}^A + \log n_{j'}^B + \log n_{k'}^C - 2 \log N \\ &= \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_j^B + \hat{\lambda}_k^C \end{aligned}$$

olur.

6. İki değişken verildiğinde diğer bir değişkenin tüm kategorilerinin aynı derecede olası olması: Bu sı-

nıftaki üç modelde model tanımlayan sınıf, AB, AC veya BC gibi bir tek çiftinden meydana gelir. Burada AC'den oluşan model tanımlayan sınıf dikkate alınmaktadır. Model,

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_k^C + \lambda_{ik}^{AC}$$

olur. Model, ancak  $A_h$  ve  $C_h$  verildiğinde  $B_h$ 'in her kategorisinin aynı derecede olası olması durumunda sağlanır. Böylece,

$$p_{j.ik}^{B.AC} = 1/s$$

olur.

$$\hat{m}_{ijk} = \frac{1}{s} n_{ik}^{AC},$$

enbüyük olabirlik tahminidir ve

$$\hat{m}_{ik}^{AC} = n_{ik}^{AC}$$

ve

$$\log \hat{m}_{ijk} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_k^C + \hat{\lambda}_{ik}^{AC}$$

eşitliklerini sağlar.

7. İki değişken verildiğinde ve bu iki değişken birbirinden bağımsız olduğunda, diğer bir değişken kategorilerinin aynı derecede olası olması: Üç model, bu sınıfı temsil eder. Model tanımlayan sınıf  $\tilde{A}$  ve B veya A ve C veya B ve C'den meydana gelebilir. Tanımlayıcı amaçlar için  $\mathcal{E}$ 'nin A ve C den meydana geldiği durumu inceleyelim. Buna göre model,

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_k^C$$

dir. Model, ancak  $A_h$  ve  $C_h$  bağımsızsa ve  $A_h$  ve  $C_h$  verildiğinde  $B_h$ 'ın her kategorisinin aynı derecede olası olması durumunda sağlanır. Dolayısıyla,

$$p_{ik}^{AC} = p_i^A p_k^C$$

ve

$$p_{j.ik}^{B.AC} = \frac{1}{s}$$

olur. Enbüyük olabilirlik tahmini,

$$\hat{m}_{ijk} = s^{-1} n_i^A n_j^B / N = \frac{1}{s} - N \hat{p}_i^A \hat{p}_k^C$$

dir. Burada,

$$\hat{p}_i^A = n_i^A / N$$

ve

$$\hat{p}_k^C = n_k^C / N$$

dir. Aşağıdaki eşitlikler sağlanır;

$$\hat{m}_i^A = n_i^A$$

$$\hat{m}_k^C = n_k^C$$

ve tahmin eşitliği

$$\log \hat{m}_{ijk} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_k^C$$

olur.

8. Bir değişken verildiğinde, diğer iki değişke-

nin kategorilerinin tüm kombinasyonlarının aynı derecede olası olması: Bu sınıfta da üç model vardır. Model tanımlayan sınıf A, B veya C tek elemanlarından meydana gelir. Eğer E, A'yı bulunduruyorsa, o zaman model,

$$\log m_{ijk} = \lambda + \lambda_i^A$$

dır.  $A_h$  verildiğinde,  $B_h$  ve  $C_h$ 'ın kategorilerinin tüm kombinasyonları aynı derecede olası ise bu model sağlanır. Buna göre,

$$p_{jk.i}^{BC.A} = \frac{1}{st}$$

olur. Enbüyük-olabilirlik tahmini,

$$\hat{m}_{ijk} = \left(\frac{1}{st}\right) n_i^A / N$$

dir ve

$$\hat{m}_i^A = n_i^A$$

ve

$$\log \hat{m}_i = \hat{\lambda} + \hat{\lambda}_i^A$$

eşitliklerini sağlar.

9. Üç değişkenin tüm kombinasyonlarının aynı derecede olası olması: Burada modeli tanımlayan sınıf hiçbir alt indis içermez ve model,

$$\log m_{ijk} = \lambda$$

dır.  $A_h$ ,  $B_h$  ve  $C_h$  in tüm kategori kombinasyonları aynı derecede olası olduğunda bu modeller sağlanır. Burada,

$$p_{ijk} = \frac{1}{rst}$$

dir. Enbüyük-olabilirlik tahmini

$$\hat{m}_{ijk} = \frac{N}{rst}$$

dir ve

$$\hat{M} = \sum_i \sum_j \sum_k \hat{m}_{ijk} = N$$

kısıtlamasını sağlar.

Üçlü-faktör interaksyonu bulunmayan model dışında,  $\hat{m}_{ijk}$ 'lerin hesaplanması kesin ifadeler üzerinden yapılabilir. Üçlü-faktör interaksyonu bulunmayan modelde  $\hat{m}_{ijk}$ 'lerin hesaplanması, Newton-Raphson algoritması veya Deming-Stephan algoritması kullanılarak yapılır.

### 3.2.3. 2x3x2 Boyutlu Tablolar

Burada, 2x3x2 boyutlu olabilirlik tablosu inceleniyor ve üçlü-faktör interaksyonu bulunmayan model yanında şartlı ve marginal bağımsızlık modellerine de önem verilmektedir. Bu modeller arasındaki ilişkiler, ortalamalar, logaritmik çapraz-çarpım oranları, enbüyük-olabilirlik tahminleri, khi-kare testleri ve düzeltilmiş rezidüyal terimleri cinsinden açıklanmaktadır.

Burada  $n_{ijk}$ ,  $A_h = i$ ,  $B_h = j$  ve  $C_h = k$  olan gözlem sayısıdır,  $1 \leq h \leq N$ . Örnekleme işlemi  $A_h = i$  olanların sayısını ( $n_i^A$ ) sabit varsayar. Yani sonuçlar  $n_1^A$  ve  $n_2^A$  üzerine şartlıdır.  $A_h = i$  olduğu verildiğinde,  $B_h = j$  ve  $C_h = k$  olması olasılığı  $p_{jk.i}^{BC.A} > 0$  dır.  $A_h$  verildiğinde  $(B_h, C_h)$ ,  $1 \leq h \leq N$ , çiftleri bağımsız olarak dağılır.  $n_{ijk}$ 'nin beklenen değeri,

o zaman  $n_i^A p_{jk.i}^{BC.A}$  dır. Her  $n_{ijk}$  değerler tablosu bağımsız multinomial dağılımı gösterir,  $1 \leq j \leq 3, 1 \leq k \leq 2$ .

Çizelge 8 - 2x3x2 Boyutlu Gözlenen Değerler Tablosu.

A \ B	1	2	$\Sigma$		
1	1	$n_{111}$	$n_{112}$	$n_{11}^B$	$n_1^A$
	2	$n_{121}$	$n_{122}$	$n_{12}^B$	
	3	$n_{131}$	$n_{132}$	$n_{13}^B$	
2	1	$n_{211}$	$n_{212}$	$n_{21}^B$	$n_2^A$
	2	$n_{221}$	$n_{222}$	$n_{22}^B$	
	3	$n_{231}$	$n_{232}$	$n_{23}^B$	
$\Sigma$	$n_1^C$	$n_2^C$	$N$		

$$\begin{aligned} n_1^B &= n_{11}^B + n_{21}^B \\ n_2^B &= n_{21}^B + n_{22}^B \\ n_3^B &= n_{31}^B + n_{32}^B \end{aligned}$$

### 3.2.3.1. Marginal ve Kısmi İlişki

A ve C değişkenleri arasındaki ilişki iki yönden incelenebilir. Birincisi, marginal ilişkinin sadece,  $A_h = i$  olduğu verildiğinde  $C_h = k$  olmasının marginal olasılıklarını ( $p_{k.i}^{C.A}$ ) dikkate alınması durumudur. Burada arada kalan B değişkeni ihmal edilir. İkincisi,  $A_h = i$  ve  $B_h = j$  olduğu verildiğinde  $C_h = k$  olmasının şartlı olasılıklarının ( $p_{k.ij}^{C.AB}$ ) kısmi ilişki şeklinde dikkate alınması durumudur.

A ve C değişkeninin marginal ilişkisi, çapraz-çarpım oranı

$$q_{(12)(12)}^{AC} = \frac{m_{11}^{AC} m_{22}^{AC}}{m_{12}^{AC} m_{21}^{AC}} = \frac{N_1^A p_{1.1}^{C.A} N_2^A p_{2.2}^{C.A}}{N_1^A p_{1.2}^{C.A} N_2^A p_{2.1}^{C.A}}$$



$$= \frac{p_{1.1}^{C.A} p_{2.2}^{C.A}}{p_{1.2}^{C.A} p_{2.1}^{C.A}}$$

veya bunun logaritması olan

$$\zeta_{(12)(12)}^{AC} = \log q_{(12)(12)}^{AC} = \log\left(\frac{p_{1.1}^{C.A}}{p_{2.1}^{C.A}}\right) - \log\left(\frac{p_{1.2}^{C.A}}{p_{2.2}^{C.A}}\right)$$

değerleri ile ölçülebilir. Burada,  $A_h = i$  ve  $C_h = k$  olan bireylerin beklenen sayısı  $m_{ik}^{AC}$ ,  $B_h = j$  için,  $1 \leq j \leq 3$ ,  $m_{ijk}$  ortalamalarının toplamıdır. Eğer

$$p_{i.1}^{C.A} = p_{i.2}^{C.A}, \quad 1 \leq i \leq 2$$

ise, A ve C değişkeni marginal olarak bağımsızdır. Bu durumda,

$$q_{(12)(12)}^{AC} = 1$$

ve dolayısıyla

$$\zeta_{(12)(12)}^{AC} = 0$$

olur.

$B_h = j$  olduğu verildiğinde, A ve C değişkeni arasındaki kısmi ilişki çapraz çarpım oranı

$$q_{(12)(12).j}^{AC.B} = \frac{m_{1j1} m_{2j2}}{m_{1j2} m_{2j1}} = \frac{p_{1.1j}^{C.AB} p_{2.2j}^{C.AB}}{p_{1.2j}^{C.AB} p_{2.1j}^{C.AB}}$$

veya bunun logaritması olan

$$\begin{aligned} \zeta_{(12)(12).j}^{AC.B} = \log q_{(12)(12).j}^{AC.B} &= \log \left( \frac{p_{1.1j}^{C.AB}}{p_{2.1j}^{C.AB}} \right) \\ &- \log \left( \frac{p_{2.2j}^{C.AB}}{p_{1.2j}^{C.AB}} \right) \end{aligned}$$

değeri ile ölçülebilir. Eğer

$$p_{i.1j}^{C.AB} = p_{i.2j}^{C.AB}, \quad 1 \leq i \leq 2, \quad 1 \leq j \leq 3$$

şartı  $B_h = j$  olduğu verildiğinde sağlanıyorsa, A ve C şartlı olarak bağımsızdır.

Kısmi ilişki ölçümlerinin ( $\zeta_{(12)(12).j}^{AC.B}$ ),  $1 \leq j \leq 3$ , marginal ilişki ölçümü  $\zeta_{(12)(12)}^{AC}$  ile ilişkisi karışıktır. İlgili tartışmalar BISHOP, FIENBERG ve HOLLAND (1975)'da bulunabilir.  $A_h$  verildiğinde  $C_h$  ve  $B_h$  şartlı olarak bağımsız ise veya  $C_h$  verildiğinde  $A_h$  ve  $B_h$  şartlı olarak bağımsız ise

$$\zeta_{(12)(12).j}^{AC.B} = \zeta_{(12)(12)}^{AC}$$

eşitliği sağlanır. Bu durum, her  $\zeta_{(12)(12).1}^{AC.B}$ ,  $\zeta_{(12)(12).2}^{AC.B}$ ,  $\zeta_{(12)(12).3}^{AC.B}$  ve  $\zeta_{(12)(12)}^{AC}$  kombinasyonu için mümkündür.

$\zeta_{(12)(12).j}^{AC.B} = -\zeta_{(12)(12)}^{AC} \neq 0$ ,  $1 \leq j \leq 3$ , gibi değişik durumlar marginal ve kısmi ilişki katsayıları karşıt işaretli olduğu durumlarda ortaya çıkar.

Marginal bağımsızlık hipotezini test edebilmek için

$$n_{ik}^{AC} = \sum_j n_{ijk} = n_{i1k} + n_{i2k}$$

değerlerinin marginal tablosu ( $n^{AC}$ ) dikkate alınır. Burada  $n_{ik}^{AC}$ ,  $A_h = i$  ve  $C_h = k$  olan bireylerin sayısıdır. Marginal bağımsızlığı test etmede  $n^{AC}$ 'ye 2x2 tabloları için alışılmış khi-kare istatistikleri uygulanabilir.

### 3.2.3.2. Marginal İlişki İçin Khi-Kare İstatistikleri

Khi-kare istatistiklerini tanımlamak için, N tablosundaki birey sayısıdır,  $n_i^A$ ,  $A_h = i$  olan ve  $n_k^C$ ,  $B_h = k$  olan birey sayısıdır. O zaman Pearson khi-kare istatistiği

$$\begin{aligned} \chi_{AC}^2 &= \sum_i \sum_k (n_{ik}^{AC} - n_i^A n_k^C / N)^2 / (n_i^A n_k^C / N) \\ &= \frac{(n_{11}^{AC} n_{22}^{AC} - n_{12}^{AC} n_{21}^{AC})^2 N}{n_1^A n_2^A n_1^C n_2^C} \end{aligned}$$

dir ve olabilirlik-oran khi-karesi

$$L_{AC}^2 = 2 \sum_i \sum_k n_{ik}^{AC} \log \left( \frac{n_{ik}^{AC} N}{n_i^A n_k^C} \right)$$

dir. Her iki istatistik için de serbestlik derecesi birdir.

### 3.2.3.3. Şartlı Bağımsızlık

Şartlı bağımsızlık modelini log-linear model terimleri cinsinden yorumlamak için, ortalamaların tek gösterime sahip olduğu belirtilir. Bu gösterimde

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

dir. Burada,

$$\begin{aligned} \sum_i \lambda_i^A &= \sum_j \lambda_j^B = \sum_k \lambda_k^C = \sum_i \lambda_{ij}^{AB} = \sum_j \lambda_{ij}^{AB} = \dots \\ &= \sum_i \lambda_{ijk}^{ABC} = \sum_j \lambda_{ijk}^{ABC} = \sum_k \lambda_{ijk}^{ABC} = 0 \end{aligned}$$

olur.

Kısmi ilişki katsayısı,

$$\begin{aligned} \tau_{(12)(12).j}^{AC.B} &= \log m_{1j1} - \log m_{1j2} - \log m_{2j1} + \log m_{2j2} \\ &= (\lambda_{11}^{AC} + \lambda_{1j1}^{ABC}) - (\lambda_{12}^{AC} + \lambda_{1j2}^{ABC}) \\ &\quad - (\lambda_{21}^{AC} + \lambda_{2j1}^{ABC}) + (\lambda_{22}^{AC} + \lambda_{2j2}^{ABC}) \\ &= 4 (\lambda_{11}^{AC} + \lambda_{1j1}^{ABC}) \end{aligned}$$

olur. Burada,

$$\lambda_{11}^{AC} = -\lambda_{21}^{AC} = -\lambda_{12}^{AC} = \lambda_{22}^{AC}$$

ve

$$\lambda_{1j1}^{ABC} = -\lambda_{1j2}^{ABC} = -\lambda_{2j1}^{ABC} = \lambda_{2j2}^{ABC}$$

dir. Eğer  $\lambda_{ik}^{AC}$  ve  $\lambda_{ijk}^{ABC}$  sıfır ise, her  $\tau_{(12)(12).j}^{AC.B}$  logaritmik çapraz-çarpım oranı sıfırdır. Böylece  $B_h$  verildiğinde  $A_h$  ve  $C_h$ 'in şartlı bağımsızlığı, eğer model

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{jk}^{BC}$$

olursa, sağlanır. Yani hem A hemde C içeren  $\lambda$ -parametreleri sıfır olur.

### 3.2.3.4. Şartlı Bağımsızlık İçin Enbüyük-Olabilirlik Tahminleri ve Khi-Kare Testleri

Khi-kare testleri ve düzeltilmiş rezidüyaller,  $B_h$  için kontrol edilen  $A_h$  ve  $C_h$ 'in bağımsızlığı modeli için kolayca hesaplanırlar.  $m_{ijk}$  enbüyük-olabilirlik tahmini ( $\hat{m}_{ijk}$ ),

$$\hat{m}_{ijk} = n_{ij}^{AB} n_{jk}^{BC} / n_j^B$$

dir. Burada  $n_j^B$ ,  $B_h = j$  olan bireylerin sayısıdır.  $n_{ij}^{AB}$ ,  $A_h = i$  ve  $B_h = j$  olan bireylerin sayısıdır.  $n_{jk}^{BC}$ ,  $B_h = j$  ve  $C_h = k$  olan bireylerin sayısıdır. Böylece Pearson khi-kare istatistiği ve olabilirlik-oran khi-kare istatistiği elde edilir. Bunlar sırasıyla

$$\begin{aligned} \chi_{AC,B}^2 &= \sum_i \sum_j \sum_k (n_{ijk} - n_{ij}^{AB} n_{jk}^{BC} / n_j^B)^2 \\ &\quad / (n_{ij}^{AB} n_{jk}^{BC} / n_j^B) \\ &= \sum_j \frac{(n_{1j1} n_{2j2} - n_{1j2} n_{2j1})^2 n_j^B}{n_{1j}^{AB} n_{2j}^{AB} n_{j1}^{BC} n_{j2}^{BC}} \end{aligned}$$

ve

$$L_{AC.B}^2 = 2 \sum_i \sum_j \sum_k n_{ijk} \log \frac{n_{ijk} n_j^B}{n_{ij}^{AB} n_{jk}^{BC}}$$

dir. Bu istatistikler için serbestlik derecesi  $s(r-1)$   $(t-1)$  dir.

Kısmi ilişkinin varlığını tetkik için başka bir yolda şartlı bağımsızlık için düzeltilmiş rezidüyallerin

$$r_{ijk} = \frac{n_{ijk} - n_{ij}^{AB} n_{jk}^{BC} / n_j^B}{\left[ n_{ij}^{AB} n_{jk}^{BC} (1 - n_{ij}^{AB} / n_j^B) (1 - n_{jk}^{BC} / n_j^B) / n_j^B \right]^{1/2}}$$

kontrol edilmesidir.  $2 \times s \times 2$  boyutlu bir tabloda düzeltilmiş rezidüyaller COCHRAN (1954)'daki gibi

$$r_{1j1} = -r_{2j1} = -r_{1j2} = r_{2j2}$$

$$= \frac{n_{1j1} n_{2j2} - n_{1j2} n_{2j1}}{\left( n_{1j}^{AB} n_{2j}^{AB} n_{j1}^{BC} n_{j2}^{BC} / n_j^B \right)^{1/2}}$$

dir.

$$x^2 = r_{111}^2 + r_{121}^2 + r_{131}^2$$

dir.  $r_{1j1}^2$  terimleri  $x^2$ 'nin unsurları olarak kabul edilebilir. Şartlı bağımsızlık modeli altında, her unsur bir serbestlik dereceli yaklaşık khi-kare dağılışı gösterir. Her  $r_{1j1}$  de yaklaşık  $N(0,1)$  dağılışı gösterir.

Bu kısımdaki sonuçlar  $2 \times s \times 2$  boyutlu tablolar içinde kullanılabilir. Burada,  $1 \leq i \leq 2$ ,  $1 \leq j \leq s$ ,  $1 \leq k \leq 2$  dir.

### 3.2.3.5. Logaritmik Çapraz-Çarpım Oranları

$\hat{\tau}_{(12)(12).j}^{AC.B}$  için % 95'lik güven aralığı

$$\hat{\tau}_{(12)(12).j}^{AC.B} \pm 1.96 s(\hat{\tau}_{(12)(12).j}^{AC.B})$$

dir. Burada

$$\hat{\tau}_{(12)(12).j}^{AC.B} = \log \left( \frac{n_{1j1} n_{2j2}}{n_{1j2} n_{2j1}} \right)$$

ve

$$s^2(\hat{\tau}_{(12)(12).j}^{AC.B}) = \frac{1}{n_{1j1}} + \frac{1}{n_{1j2}} + \frac{1}{n_{2j1}} + \frac{1}{n_{2j2}}$$

dir.  $q_{(12)(12).j}^{AC.B}$ 'ye tekabül eden güven sınırları da bulunabilir.

### 3.2.3.6. $A_h$ Verildiğinde $B_h$ ve $C_h$ Arasındaki Kısmi İlişki

A değişkeni verildiğinde B ve C değişkeni arasındaki kısmi ilişki çapraz-çarpım oranları ve log-linear modellerle tanımlanabilir.  $1 \leq j < j' \leq 3$  için,

$$q_{(jj')(12).i}^{BC.A} = \frac{m_{ij1} m_{ij'2}}{m_{ij2} m_{ij'1}} = \frac{p_{1.ij}^{C.AB} p_{2.ij'}^{C.AB}}{p_{2.ij}^{C.AB} p_{1.ij'}^{C.AB}},$$

$A_h=i$  ve  $B_h=j$  olan bireyler arasındaki uyumun risklerini

$$p_{1.ij}^{C.AB} / p_{2.ij}^{C.AB},$$

bunlara tekabül eden  $A_h=i$  ve  $B_h=j'$  olan bireyler için bulunan riskler

$$p_{1.ij'}^{C.AB} / p_{2.ij'}^{C.AB}$$

ile karşılaştırmaktadır.

$$\tau_{(jj')(12).i}^{BC.A} = \log q_{(jj')(12).i}^{BC.A}$$

olsun, bu durumda,

$$\begin{aligned} \tau_{(jj')(12).i}^{BC.A} &= (\lambda_{j1}^{BC} - \lambda_{j2}^{BC} - \lambda_{j'1}^{BC} + \lambda_{j'2}^{BC}) \\ &\quad + (\lambda_{ij1}^{ABC} - \lambda_{ij2}^{ABC} - \lambda_{ij'1}^{ABC} \\ &\quad \quad \quad + \lambda_{ij'2}^{ABC}) \\ &= 2(\lambda_{j1}^{BC} - \lambda_{j'1}^{BC}) + 2(\lambda_{ij1}^{ABC} - \lambda_{ij'1}^{ABC}) \end{aligned}$$

olur.

Eğer,  $A_h$  verildiğinde  $B_h$  ve  $C_h$  şartlı olarak bağımsız ise, o zaman tüm  $q_{(jj')(12).i}^{BC.A}$  'ler 1 dir. Tüm  $\tau_{(jj')(12).i}^{BC.A}$  'ler 0 ve  $\lambda_{ij}^{BC}$  ve  $\lambda_{ijk}^{ABC}$  'lerin hepsi 0 dir.

$A_h$  verildiğinde  $B_h$  ve  $C_h$  'ın bağımsızlığı ancak

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC}$$

ise sağlanır.



Bu model için enbüyük-olabilirlik tahminleri ,

$$\hat{m}_{ijk} = n_{ij}^{AB} n_{ik}^{AC} / n_i^A$$

eşitliğiyle bulunur.

İlgili khi-kare istatistikleri

$$\chi_{BC.A}^2 = \sum_i \sum_j \sum_k \frac{(n_{ijk} - n_{ij}^{AB} n_{ik}^{AC} / n_i^A)^2}{(n_{ij}^{AB} n_{ik}^{AC} / n_i^A)}$$

ve

$$L_{BC.A}^2 = 2 \sum_i \sum_j \sum_k n_{ijk} \log \left( \frac{n_{ijk} n_i^A}{n_{ij}^{AB} n_{ik}^{AC}} \right)$$

dir. Serbestlik derecesi  $r(s-1)(t-1)$  olur.

B ve C değişkeni arasındaki kısmi ilişkinin büyüklüğü hakkında bilgi edinmek için

$$\hat{\tau}_{(jj')(12).i}^{BC.A} = \log \left( \frac{n_{ij1} n_{ij'2}}{n_{ij2} n_{ij'1}} \right)$$

$$s \left( \hat{\tau}_{(jj')(12).i}^{BC.A} \right) = \left( \frac{1}{n_{ij1}} + \frac{1}{n_{ij2}} + \frac{1}{n_{ij'1}} + \frac{1}{n_{ij'2}} \right)^{1/2}$$

ve

$$\hat{q}_{(jj')(12).i}^{BC.A} = \frac{n_{ij1} n_{ij'2}}{n_{ij2} n_{ij'1}}$$

tahminleri dikkate alınır.

Eğer üçlü-faktör interaksyonu ( $\hat{\gamma}_{ijk}^{ABC}$ ) yoksa, o zaman

$$\tau_{(jj')}(12).1 = \tau_{(jj')}(12).2$$

olur.

### 3.2.3.7. Üçlü-Faktör İnteraksiyonu Olmayan Model

Eğer üçlü-faktör interaksiyonu olmayan model verilere uygun olursa, o zaman B verildiğinde A ve C arasında kısmi ilişki vardır ve  $\tau_{(jj')}(12).i$ , A'ya bağımlı olur.

Üçlü-faktör interaksiyonu bulunmayan model için enbüyük-olabilirlik tahminlerinin hesaplanması Newton-Raphson algoritmasının kullanılmasıyla olur. HABERMAN (1978)'e göre, başka bir alternatif DEMING-STEPHAN (1940)'in iterative oransal uydurma algoritmasının kullanılmasıdır. Bu algoritmanın üçlü-faktör interaksiyonu bulunmayan model için enbüyük-olabilirlik tahminlerini hesaplamada kullanılan şekli BISHOP (1969) ve BISHOP ve MOSTELLER (1969)'e dayanır. Algoritma sürekli olarak ve büyüdükçe  $\hat{m}_{ijk}$ 'ya yaklaşan  $m_{ijkv}$ 'ler verir.

Daha sonra  $x_{ABC}^2$  ve  $L_{ABC}^2$  istatistikleri hesaplanır. Serbestlik derecesi  $(r-1)(s-1)(t-1)$  dir.  $L_{BC.A}^2 - L_{ABC}^2$ , B ve C değişkeni arasındaki ilişkinin ve  $L_{AC.B}^2 - L_{ACB}^2$  istatistiği, A ve C değişkeni arasındaki ilişkinin ilave bir ispatını verir.

### 3.2.4. Modellerin Seçimi

Eğer belirli bir hiyerarşik modelle ilgileniliyorsa, o zaman model testi doğrudan doğruya yapılan bir işlemdir. Khi-kare istatistikleri halihazırda elde edilmiş durumda ve düzeltilmiş rezidüyaller için formüller

üçlü-faktör interaksyonu olmayan model dışında doğrudan işleme konacak türdendir. Eğer veriler önceden seçilmiş herhangi bir belirli hiyerarşik modele göre analiz edilmediyse, o zaman  $R^2$  istatistikleri, F istatistikleri, parametre tahminlerinin standardize değerleri ve olabilirlik-oran khi-karesinin parçalanması önem kazanır.

### 3.2.4.1. Khi-Kare İstatistikleri

$\chi^2$  ve  $L^2$ 'nin hiyerarşik modellerin testinde kullanılmasında özel bir sorunla karşılaşılmaz. Her zaman olduğu gibi, khi-kare yaklaşımının kesinliği hakkında çok bilgi yoktur. Khi-kare istatistikleri için serbestlik dereceleri, sıfır varsayılan  $\lambda$ -parametreleri için tüm  $b(\tau)$  terimlerinin toplamıdır.

$A_h, B_h$  ve  $C_h, 1 \leq h \leq N$  değişkenlerindeki gözlemlerin karşılıklı sınıflanması ile elde edilen rxsxt tablosunda,

$$\begin{aligned} b(A) &= r-1, \\ b(B) &= s-1, \\ b(C) &= t-1, \\ b(AB) &= (r-1)(s-1), \\ b(AC) &= (r-1)(t-1), \\ b(BC) &= (s-1)(t-1), \end{aligned}$$

ve

$$b(ABC) = (r-1)(s-1)(t-1)$$

olur. Üçlü-faktör interaksyonu olmayan modelde sadece  $\lambda_{ijk}^{ABC}$  sıfır varsayıldığından,

$$b(ABC) = (r-1)(s-1)(t-1)$$

serbestlik derecesi vardır. Yani üçlü-faktör interaksyonu

olmayan bir modeli test için hesaplanan  $X^2$  ve  $L^2$  istatistiklerine ait serbestlik dereceleri  $(r-1)(s-1)(t-1)$  dir.

Eğer incelenen model

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC}$$

ise bu modelde  $\lambda_{jk}^{BC}$  ve  $\lambda_{ijk}^{ABC}$  sıfır varsayılır. Böylece hesaplanan  $X^2$  ve  $L^2$  istatistikleri ,

$$\begin{aligned} b(BC) + b(ABC) &= (s-1)(t-1) + (r-1)(s-1)(t-1) \\ &= r(s-1)(t-1) \end{aligned}$$

serbestlik derecesine sahiptir. Böylece  $x^2$  dağılışından faydalanarak elde edilen aynı serbestlik derecesindeki kritik cetvel değeri ile karşılaştırılarak testler yapılır. Modelin önemini test için kullanılan diğer istatistikler  $F$  ve  $R^2$  istatistikleridir.  $F$  istatistiği

$$F_i = \left( \frac{1}{SD_i} \right) (L_i)$$

şeklinde ifade edilir.  $SD_i$  serbestlik derecesi ve  $L_i$  test edilecek khi-kare değeridir.  $F$  istatistiği regresyon analizi veya varyans analizindeki  $F$  istatistiği ile benzerdir ve yaklaşık uygun serbestlik dereceli ( $SD_i$ )  $F$  dağılışı gösterir.  $F$  istatistiği 1'e yaklaşırsa uyum iyidir.  $R^2$  istatistiği de belirli bir modelin diğer modellere üstünlüğünü test etmek amacıyla kullanılabilir.

$$R^2 = (L_1 - L_2) / L_1$$

şeklinde ifade edilir. Regresyon analizindeki  $R^2$  istatistiğine benzer.  $R^2$  istatistiğine ait serbestlik derecesi bulunurken istatistiği hesaplamada kullanılan işlem yapılır.

Örneğin 1. modele ait serbestlik derecesi ile 2. modele ait serbestlik derecesi farkı 1. modele ait serbestlik derecesine bölünerek, 1. modelin 2. modele üstünlüğünü test etmede kullanılacak kritik cetvel değerini bulmada faydalanacağımız serbestlik derecesi elde edilir.

Modeli tanımlayan sınıflara göre serbestlik dereceleri Çizelge 9 'daki gibi yazılabilir.

Çizelge 9 - Üç-Yönlü Tablolar İçin Hiyerarşik Log-Linear Modellerde Khi-Kare Testlerine Ait Serbestlik Dereceleri.

Modeli tanımlayan sınıf	Serbestlik derecesi
ABC	0
AB, AC, BC	$(r-1)(s-1)(t-1)$
AB, AC	$r(s-1)(t-1)$
AB, C	$(rs-1)(t-1)$
A, B, C	$rst-r-s-t+2$
AB	$rs(t-1)$
A, B	$rst-r-s+1$
A	$r(st-1)$
Boş	$rst-1$

#### 3.2.4.2. Khi-Kare İstatistiklerinin Parçalanması

Bu işleme verilere en iyi uyan modeli tesbit etmek amacıyla başvurulur. Bunun için hiyerarşik model tanımlama esasları içerisinde modeldeki her terimi test etmek için gerekli olan birkaç model uydurulur ve herbirisine ait olabilirlik-oran khi-karesi ve serbestlik dereceleri hesaplanır. Daha sonra  $L^2$ 'ler arasındaki farklar test edilerek en uygun model bulunmaya çalışılır.

Anlatılanlar Çizelge 10 'da özetlenmektedir.

Çizelge 10- Hiyerarşik modeller için  $L^2$  istatistikleri.

Model numarası	Model tanımlayan sınıf	$L^2$	SD	Khi-kare-deki indirgeme	SD'deki indirme	$F^a$	$R^{2b}$
1	AB,C	$L_1^2$	$(rs-1)(t-1)$				
2	AB,AC	$L_2^2$	$r(s-1)(t-1)$	$L_1^2 - L_2^2$	$SD_1 - SD_2$		
3	AB,BC	$L_3^2$	$s(r-1)(t-1)$	$L_1^2 - L_3^2$	$SD_2 - SD_3$		
4	AB,AC,BC	$L_4^2$	$(r-1)(s-1)(t-1)$	$L_3^2 - L_4^2, L_2^2 - L_4^2$	$SD_3 - SD_4, SD_2 - SD_4$		

a,  $b_F$  ve  $R^2$  değerleri formülde belirtildiği şekilde bulunabilirler.

$$L_1^2 = \sum_i \sum_j \sum_k n_{ijk} \log \left( \frac{n_{ijk}}{n_{ij} \frac{n_{.k}}{N}} \right)$$

$$L_2^2 = 2 \sum_i \sum_j \sum_k n_{ijk} \log \left( \frac{n_{ijk}}{n_{ij} \frac{n_{ik}}{N}} \right)$$

$$L_3^2 = 2 \sum_i \sum_j \sum_k n_{ijk} \log \left( \frac{n_{ijk}}{n_{ij} \frac{n_{jk}}{N}} \right)$$

ve

$$L_4^2 = 2 \sum_i \sum_j \sum_k n_{ijk} \log \left( \frac{n_{ijk}}{\hat{m}_{ijk4}} \right)$$

olur.

1. Model, C'nin A ve B'den bağımsız olduğunu, 2.

model, A verildiğinde B ve C'nin bağımsız olduklarını, 3. model, B verildiğinde, A ve C'nin bağımsız olduklarını varsayar, 4. model ise üçlü-faktör interaksyonu olmayan modeldir.

$L_1^2-L_2^2$  farkı AC interaksyonunun sıfır olması hipotezini test etmede  $L_1^2-L_3^2$  farkı, BC interaksyonu test etmede kullanılır.  $L_3^2-L_4^2$ , AC ve  $L_2^2-L_4^2$  ise yine BC interaksyonunu test etmede kullanılır.

### 3.2.4.3. Rezidüyal Analiz

Modeldeki başarısızlığın neden kaynaklandığını araştırmak amacıyla, diğer bir ifade ile bağımlılığın hangi kategorilerden geldiğini bulmak için düzeltilmiş rezidüyal analizine gerek duyulur.

Rezidüyal analizinde Çizelge 11'de gösterilen  $r_{ijk}$  formülleri kullanılır. Üçlü-faktör interaksyonu olmayan model için rezidüyal analizi yapılmaz. Çünkü  $r_{ijk}$ 'lerin hesaplanması genellikle enbüyük-olabilirlik tahminlerinin bulunmasında kullanılan Newton-Raphson algoritmasının kullanılmasında elde edilen  $\hat{S}$  matrisinin inversine gerek duyar.

Bu modelle de kabada olsa bir yorum yapabilmek için, bazen standardize rezidüyaller

$$S_{ijk} = (n_{ijk} - \hat{m}_{ijk}) / \hat{m}_{ijk}^{1/2}$$

kullanılması faydalı olur. Üçlü-faktör interaksyonu olmayan model altında,  $S_{ijk}$ 'lar, asimtotik ortalaması 0 ve asimtotik varyansı 1'den daha küçük olan yaklaşık normal dağılım gösterirler. Eğer r, s ve t büyükse ve tüm  $m_{ijk}$ 'lar büyüklük olarak karşılaştırılabilir durumda iseler,

o zaman  $S_{ijk}^*$ 'lar  $r_{ijk}$ 'lara oldukça iyi bir yaklaşım gösterirler. Aynı amaç için, daha iyi düzeltilmiş ve aşağıdaki şekilde hesaplanan

$$S_{ijk}^* = \frac{(n_{ijk} - \hat{m}_{ijk})}{\left[ \hat{m}_{ijk} \left(1 - \frac{n_{ij}^{AB}}{n_i^A}\right) \left(1 - \frac{n_{ik}^{AC}}{n_i^A}\right) \right]^{1/2}}$$

standart rezidüyal de kullanılabilir.  $S_{ijk}^*$ 'nın asimtotik varyansı 1'den küçük, fakat  $S_{ijk}$ 'ninkinden daha büyüktür.

Rezidüyal analizin esas önemi, sadece sınırlı sayıda hücre içeren bir modelden sapmaların bulunmasındadır. Bu özellik cetvelde görüldüğü şekilde belli model tanımlayan sınıflar için rezidüyalenin hesaplanması ve önem testlerinin yapılması ile açıklanabilir.

Çizelge 11- Üç-Yönlü Tablolarda Hiyerarşik Modeller İçin Düzeltilmiş Rezi-düyalen.

Model tanımlayan sınıf	$r_{ijk}$
AB, AC	$\frac{n_{ijk} - (n_{ij}^{AB} n_{ik}^{AC} / n_i^A)}{\left[ \frac{n_{ij}^{AB} n_{ik}^{AC}}{n_i^A} \left(1 - \frac{n_{ij}^{AB}}{n_i^A}\right) \left(1 - \frac{n_{ik}^{AC}}{n_i^A}\right) \right]^{1/2}}$
AB, C	$\frac{n_{ijk} - (n_{ij}^{AB} n_k^C / N)}{\left[ \frac{n_{ij}^{AB} n_k^C}{N} \left(1 - \frac{n_{ij}^{AB}}{N}\right) \left(1 - \frac{n_{ik}^{AC}}{n_i^A}\right) \right]^{1/2}}$



Çizelge 11- Devam.

A, B, C

$$\frac{n_{ijk} - (n_i^A n_j^B n_k^C / N^2)}{\left[ \frac{n_i^A n_j^B n_k^C}{N} \left( 1 - \frac{n_i^A}{N} \frac{n_j^B}{N} \frac{n_k^C}{N} - \frac{n_j^B n_k^C}{NN} + \frac{2n_i^A n_j^B n_k^C}{NNN} \right) \right]^{1/2}}$$

AB

$$\frac{n_{ijk} - \left( \frac{1}{t} \right) n_{ij}^{AB}}{\left[ \frac{1}{t} \left( 1 - \frac{1}{t} \right) n_{ij}^{AB} \right]^{1/2}}$$

A, B

$$\frac{n_{ijk} - \left( \frac{1}{t} n_i^A n_j^B / N \right)}{\left\{ \frac{1}{t} \frac{n_i^A n_j^B}{N} \left[ 1 - \frac{1}{t} + \frac{1}{t} \left( 1 - \frac{n_i^A}{N} \right) \left( 1 - \frac{n_j^B}{N} \right) \right] \right\}^{1/2}}$$

A

$$\frac{n_{ijk} - \left( \frac{1}{st} \right) n_i^A}{\left[ \frac{1}{st} \left( 1 - \frac{1}{st} \right) n_i^A \right]^{1/2}}$$

Boş

$$\frac{n_{ijk} - \left( \frac{1}{rst} \right) N}{\left[ \frac{1}{rst} \left( 1 - \frac{1}{rst} \right) N \right]^{1/2}}$$

Rezidüyal analizi yeniden enbüyük-olabilirlik tahminleri hesaplanmadan sınırlı sayıdaki sapmaları tesbit için kullanılır. Modelden daha genel olan sapmalar rezidüyal analizi ile ortaya konur. Yine de, üzerinde durulan modelin daha genel modellerle karşılaştırılması, olabilirlik

oran istatistiğinin parçalanması ve parametre tahminlerinin kendi tahmin edilen asimtotik standart sapmaları ile karşılaştırılması gibi yollarla daha geniş bir karşılaştırma şeklinde yapılabilir.

#### 3.2.4.4. Standardize Parametre Tahminleri

Verilere oldukça iyi uyum sağlayabilen hiyerarşik modellerin tesbitinde, parametre tahminlerinin ve bunlara ait tahmin edilen standart sapmaların karşılaştırılması iyi bir yoldur. Uygulamada böyle karşılaştırmalar genellikle doymuş modeldeki parametre tahminlerine dayanır. Hesaplamadaki basitlik ve doymuş modelin genelliği bu yaklaşımı ilginçleştirir.

Örnek olarak 2x2x2 tablosu dikkate alınır, doymuş model altında  $\lambda, \lambda_1^A, \lambda_1^B, \lambda_1^C, \lambda_{11}^{AB}, \lambda_{11}^{AC}, \lambda_{11}^{BC}$  ve  $\lambda_{111}^{ABC}$  parametrelerinin tahmini kolaylıkla elde edilebilir. Eğer

$$\hat{m}_{ijk} = \log \hat{m}_{ijk} = \log n_{ijk}$$

ise, o zaman

$$\hat{\lambda}_{111}^{ABC} = \frac{1}{8} (\hat{m}_{111} - \hat{m}_{211} - \hat{m}_{121} + \hat{m}_{221} - \hat{m}_{112} + \hat{m}_{212} + \hat{m}_{122} - \hat{m}_{222})$$

olur.  $\hat{\lambda}_{111}^{ABC}$ 'nin tahmin edilen asimtotik standart sapması

$$s(\hat{\lambda}_{111}^{ABC}) = \frac{1}{8} \left( \sum_i \sum_j \sum_k 1/n_{ijk} \right)^{1/2}$$

olur. Böylece

$$\begin{aligned} Z_{ABC} &= \frac{\hat{\lambda}_{111}^{ABC}}{S(\hat{\lambda}_{111}^{ABC})} \\ &= \frac{\hat{m}_{111} - \hat{m}_{211} - \hat{m}_{121} + \hat{m}_{221} - \hat{m}_{112} + \hat{m}_{212} + \hat{m}_{122} - \hat{m}_{222}}{\left(\frac{1}{n_{111}} + \frac{1}{n_{211}} + \frac{1}{n_{121}} + \frac{1}{n_{221}} + \frac{1}{n_{112}} + \frac{1}{n_{212}} + \frac{1}{n_{122}} + \frac{1}{n_{222}}\right)^{1/2}} \end{aligned}$$

standardize test istatistiği  $\lambda_{111}^{ABC}$ 'nin sıfır olması hipotezinin testini sağlar. Tüm  $\lambda_{ijk}^{ABC}$ 'ler  $\lambda_{111}^{ABC}$  veya  $-\lambda_{111}^{ABC}$ 'ye eşit olduğundan,  $\lambda_{ijk}^{ABC}$ 'lerin sıfır olup olmadıkları da test edilmiş olur. Diğer  $\lambda$ -parametrelerinin sıfır olup olmadıklarına ait hipotezler doymuş modeldeki tahminlere dayanan benzer testlerle kontrol edilebilir.

2x2x2 tablosu dikkate alındığından, hesaplamalar büyük çapta basitleştirilebilir.

$$\hat{\lambda}_{1(A)}^A = \hat{\lambda}_1^A$$

$$\hat{\lambda}_{1(A)B}^{AB} = \hat{\lambda}_{11}^{AB}$$

$$\lambda_{1(ABC)}^{ABC} = \hat{\lambda}_{111}^{ABC} \text{ vs olsun. O zaman } T \text{ üst indisi}$$

için

$$S(\hat{\lambda}_1^T(T)) = \frac{1}{8} \left( \sum_i \sum_j \sum_k \frac{1}{n_{ijk}} \right)^{1/2}$$

olur. Böylece tahmin edilen asimtotik standart sapmaların elde edilmesi oldukça kolaydır. Hesaplanan standardize değerlerin önemlilik olasılıkları  $2 [1 - \Phi(|Z_T|)]$  ile veya hesaplanan  $Z_T$ 'ye tekabül eden Z cetvel değerinin  $P(Z \geq Z_c)$ , iki katı alınarak bulunur. Bu şekilde bulunan değer 0.01'den

küçük ise o terim önemli olarak kabul edilir.

### 3.3. Çok-Yönlü Tam Tablolar

Üç-yönlü tablolardaki hiyerarşik modellerin kullanımı ile ilgili esaslar dört veya daha fazla değişkenin karşılıklı sınıflanması şeklinde oluşturulan çok-boyutlu tabloların incelenmesinde de uygulanabilir. Değişken sayısı arttıkça üç esas değişiklik ortaya çıkar; mümkün olan hiyerarşik modellerin sayısı çok hızlı bir şekilde artar; hiyerarşik modellerden gittikçe azalan sayıda bazı modeller enbüyük-olabilirlik tahminleri için kesin ifadelerle sahip olurlar; ve hiyerarşik modellerin yorumu değişken sayısı arttıkça karmaşık bir hal alır.

Bu tip modellerin analizinde de hiyerarşik modeller yorumlanırken logaritmik riskler, şartlı logaritmik riskler, çapraz-çarpım oranları ve şartlı çapraz-çarpım oranları kullanılır. Yine model seçiminde standardize parametre tahminleri büyük bir araçtır.

Bu bölümde, enbüyük-olabilirlik eşitlikleri kapalı-form tahminleri ve iterative oransal uydurma yöntemleri kullanılarak bulunur. Bunların kullanılmasının başlıca iki nedeni vardır. İlki, birçok model, parametrelerin tahmin edilen asimtotik standart sapmaları dikkate alınmadan karşılaştırılabilir. İkincisi, en çok ilgilenilen modeller kapalı-formda enbüyük-olabilirlik tahminlerine sahip olduğundan, gerek duyulan asimtotik standart sapmalar, Newton-Raphson algoritmasına yeniden başvurulmadan elde edilebilir. Newton-Raphson algoritmasının daha çok boyutlu tabloların analizi için uygun olmadığı sanılmamalıdır.

### 3.3.1. 2x2x3x2 Boyutlu Tablolar

Burada 2x2x3x2 boyutlu tablo için uygun hiyerarşik modeller açıklanıyor. Uygun model sınıfını tesbit etmek için standardize parametre tahminleri kullanılır. O zaman iterative oransal uydurma bir başka model için kullanılırken, ilgilenilen bazı hiyerarşik modellerin enbüyük-olabilirlik tahminleri için kapalı-formda ifadeler bulunur. Daha sonra, olabilirlik-oran khi-kare istatistikleri, hangi modelin verilerle uyumlu olduğunu tayin etmek için kullanılır. O zaman cazip olan çoğu modellerin parametreleri tahmin edilir ve yorumlanır.

Burada  $n_{ijkl}$ ,  $A_h = i$ ,  $B_h = j$ ,  $C_h = k$ ,  $D_h = l$  olan bireylerin sayısıdır. Örneklememe işlemi, sonuçların  $A_h = i$ ,  $B_h = j$  olan birey sayısı için gözlenen  $n_{ij}^{AB}$  marginal toplamları üzerine şartlı olduğunu varsayar.  $A_h = i$  ve  $B_h = j$  olduğu verildiğinde  $(C_h, D_h)$  çiftlerinin,  $A_h = i$  ve  $B_h = j$  olduğu verildiğinde,  $C_h = k$  ve  $D_h = l$  olmasının şartlı olasılığı olan  $P_{kl.ij}^{CD.AB} > 0$  olasılığına sahip bağımsız dağılmış çiftler olduğu varsayılır. Böylece  $n_{ijkl}$ 'nin beklenen değeri olan  $m_{ijkl} = n_{ij}^{AB} P_{kl.ij}^{CD.AB}$  olur.

#### 3.3.1.1. Doymuş Modelle Tahmin

Bu tablo için olası hiyerarşik modeller,

$$\begin{aligned} \mu_{ijkl} = \log m_{ijkl} = & \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} \\ & + \lambda_{il}^{AD} + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD} + \lambda_{ijk}^{ABC} \\ & + \lambda_{ijl}^{ABD} + \lambda_{ikl}^{ACD} + \lambda_{jkl}^{BCD} + \lambda_{ijkl}^{ABCD} \end{aligned}$$

şeklindeki bir modelin parçalanmasına dayanır. Burada

$$\begin{aligned}
 \sum_i \lambda_i^A &= \sum_i \lambda_j^B = \sum_k \lambda_k^C = \sum_l \lambda_l^D = \sum_i \lambda_{ij}^{AB} = \sum_j \lambda_{ij}^{AB} \\
 &= \dots = \sum_i \lambda_{ijk}^{ABC} = \sum_j \lambda_{ijk}^{ABC} = \sum_k \lambda_{ijk}^{ABC} \\
 &= \dots = \sum_i \lambda_{ijkl}^{ABCD} = \sum_j \lambda_{ijkl}^{ABCD} \\
 &= \sum_k \lambda_{ijkl}^{ABCD} = \sum_l \lambda_{ijkl}^{ABCD} = 0
 \end{aligned}$$

olur.

Dikkate alınması gereken olası hiyerarşik modelleri belirlemek için hiçbir  $\lambda$ -parametresinin sıfır varsayılmadığı doymuş model altında,  $\lambda$ -parametrelerini ve bunlara tekabül eden asimtotik standart sapmaları tahmin etmek faydalıdır. Bu modelde  $m_{ijk}$ 'ler  $n_{ijkl}$  enbüyük-olabilirlik tahminine sahiptir ve  $m_{ijkl} = \log m_{ijkl}$ ,  $\hat{m}_{ijkl} = \log n_{ijkl}$  enbüyük-olabilirlik tahminine sahiptir.

$\lambda_v^T$  parametrelerinin tahmin eşitlikleri genel olarak

$$\hat{\lambda}_v^T = \frac{1}{2.2.3.2} \sum_i \sum_j \sum_k \sum_l c_{ijkl} (\lambda_v^T) / \hat{m}_{ijkl} \quad (3.3-1)$$

şeklinde ifade edilir. Burada,  $c_{ijkl} (\lambda_v^T)$ 'lar değişkenin seviye sayısı ve  $\bar{f}$  değerlerine göre negatif veya pozitif belli değerler alırlar. Örneğin;  $\hat{\lambda}_{11}^{CD}$  için katsayı =  $k' = 1' = 1$  dir, dolayısıyla katsayı  $(3\bar{f}_{k2}-1)(2\bar{f}_{11}-1)$  olur. Eğer  $k=1=1$  ise  $\bar{f}_{kl} = \bar{f}_{11} = 1$  dir. Dolayısıyla da katsayı  $(3*2-1)(2*1-1) = 2$  olur. Eğer  $k > 1$ , fakat  $l=1$  olsaydı  $\bar{f}_{kl} = 0$

$f_{11} = 1$  olacaktı. Katsayı da  $(3 \times 0 - 1)(2 \times 1 - 1) = -1$  olur. Eğer  $k=1, l=2$  ise,  $f_{k1} = 1, f_{11} = 0$  olur ve katsayı da  $(3 \times 1 - 1)(2 \times 0 - 1) = -2$  olur. Son durum  $k > 1$  ve  $l=2$  ise  $f_{k1} = 0, f_{11} = 0$  ve katsayının  $(3 \times 0 - 1)(2 \times 0 - 1) = 1$  olması durumudur.

(3.3-1) deki genel ifade açılırsa,

$$\hat{\lambda} = \frac{1}{2.2.3.2} \sum_i \sum_j \sum_k \sum_l \hat{m}_{ijkl},$$

$$\hat{\lambda}_{1', D} = \frac{1}{2.2.3.2} \sum_i \sum_j \sum_k \sum_l (2 f_{11', -1}) \hat{m}_{ijkl},$$

$$\hat{\lambda}_{k', 1'}^{CD} = \frac{1}{2.2.3.2} \sum_i \sum_j \sum_k \sum_l (3 f_{kk', -1}) (2 f_{11', -1}) \hat{m}_{ijkl},$$

$$\hat{\lambda}_{j'k', 1'}^{BCD} = \frac{1}{2.2.3.2} \sum_i \sum_j \sum_k \sum_l (2 f_{jj', -1}) (3 f_{kk', -1}) (2 f_{11', -1}) \hat{m}_{ijkl}$$

ve

$$\hat{\lambda}_{i'j'k', 1'}^{ABCD} = \frac{1}{2.2.3.2} \sum_i \sum_j \sum_k \sum_l (2 f_{ii', -1}) (2 f_{jj', -1}) (3 f_{kk', -1}) (2 f_{11', -1}) \hat{m}_{ijkl}$$

olur. Burada, eğer  $x=y$  ise  $f_{xy} = 1$  ve  $x \neq y$  ise  $f_{xy} = 0$

olur.

Parametrelerin asimtotik varyans tahminleri hesaplanırken, eğer  $n^T$  marginal toplamları örnekleme yöntemiyle sabitleştirilmemişse, o zaman

$$s^2(\hat{\lambda}_v^T) = \left(\frac{1}{2.2.3.2}\right)^2 \sum_i \sum_j \sum_k \sum_l \left[ c_{ijkl}(\hat{\lambda}_v^T) \right]^2 / n_{ijkl}$$

olur. Örneğin,  $D=1$ ' olanların toplamı olan  $n_1^D$ , sabit değildir. O zaman

$$\begin{aligned} s^2(\hat{\lambda}_{1'}^D) &= \left(\frac{1}{2.2.3.2}\right)^2 \sum_i \sum_j \sum_k \sum_l (2 f_{11',-1})^2 / n_{ijkl} \\ &= \frac{1}{576} \sum_i \sum_j \sum_k \sum_l 1 / n_{ijkl} \end{aligned}$$

olur. Burada  $(2 f_{11',-1}) = 1$  veya  $-1$  dir. Benzer şekilde,  $n_{k'1}^{CD}$  toplamı sabit değildir. O zaman

$$\begin{aligned} s^2(\hat{\lambda}_{k'1}^{CD}) &= \left(\frac{1}{2.2.3.2}\right)^2 \sum_i \sum_j \sum_k \sum_l (3 f_{kk',-1})^2 \\ (2 f_{11',-1})^2 / n_{ijkl} &= \frac{1}{576} \sum_i \sum_j \sum_k \sum_l \\ (3 f_{kk',-1})^2 / n_{ijkl} \end{aligned}$$

olur.

Bu formüller,  $\hat{\lambda}_{i'j}^{AB}$ 'ye uygulanmaz. Çünkü örnekleme işlemi ile  $n_{i'j}^{AB}$ 'ye sınırlama konmuştur.

Eğer  $n_{ijkl}$ 'ler çok büyük değilse, o zaman  $\hat{m}_{ijkl}$ '-



leri ve asimtotik standart sapmaların tahminlerini hesaplamadan önce her  $n_{ijkl}$ 'ye 1/2 eklemek uygun olur. GOODMAN (1970), çok boyutlu tablolar için doymuş modellerde,  $n_{ijkl}$ 'lere 1/2 eklemeyi tavsiye etmektedir.

$S(\hat{\lambda}_1^A)$ ,  $S(\hat{\lambda}_1^B)$  ve  $S(\hat{\lambda}_{11}^{AB})$ ,  $n_{ij}^{AB}$  ye dayandığından hesaplanmaz.  $2 \times 2 \times 3 \times 2$  boyutlu tablo için,

$$\hat{\lambda}_2^A = -\hat{\lambda}_1^A, \quad \hat{\lambda}_2^B = -\hat{\lambda}_1^B, \quad \hat{\lambda}_2^C = -\hat{\lambda}_1^C - \hat{\lambda}_3^C,$$

$$\hat{\lambda}_2^D = -\hat{\lambda}_1^D,$$

$$\hat{\lambda}_{22}^{AB} = -\hat{\lambda}_{12}^{AB} = -\hat{\lambda}_{21}^{AB} = \hat{\lambda}_{11}^{AB}, \quad \hat{\lambda}_{2k'}^{AC} = -\hat{\lambda}_{1k'}^{AC},$$

$$\hat{\lambda}_{22}^{AD} = -\hat{\lambda}_{12}^{AD} = -\hat{\lambda}_{21}^{AD} = \hat{\lambda}_{11}^{AD}, \quad \hat{\lambda}_{2k'}^{BC} = -\hat{\lambda}_{1k'}^{BC},$$

$$\hat{\lambda}_{22}^{BD} = -\hat{\lambda}_{12}^{BD} = -\hat{\lambda}_{21}^{BD} = \hat{\lambda}_{11}^{BD}, \quad \hat{\lambda}_{k'2}^{CD} = -\hat{\lambda}_{k'1}^{CD}$$

$$\hat{\lambda}_{22k'}^{ABC} = -\hat{\lambda}_{12k'}^{ABC} = -\hat{\lambda}_{21k'}^{ABC} = \hat{\lambda}_{11k'}^{ABC}, \quad \hat{\lambda}_{222}^{ABD} = -\hat{\lambda}_{122}^{ABD}$$

$$= -\hat{\lambda}_{212}^{ABD} = -\hat{\lambda}_{221}^{ABD} = \hat{\lambda}_{112}^{ABD} = \hat{\lambda}_{121}^{ABD} = \hat{\lambda}_{211}^{ABD} = -\hat{\lambda}_{111}^{ABD},$$

$$\hat{\lambda}_{2k'2}^{ACD} = -\hat{\lambda}_{1k'2}^{ACD} = -\hat{\lambda}_{2k'1}^{ACD} = \hat{\lambda}_{1k'1}^{ACD}, \quad \hat{\lambda}_{2k'2}^{BCD} = -\hat{\lambda}_{1k'2}^{BCD}$$

$$= -\hat{\lambda}_{2k'1}^{BCD} = \hat{\lambda}_{1k'1}^{BCD},$$

$$\hat{\lambda}_{22k'2}^{ABCD} = -\hat{\lambda}_{12k'2}^{ABCD} = -\hat{\lambda}_{21k'2}^{ABCD} = -\hat{\lambda}_{22k'1}^{ABCD} = \hat{\lambda}_{11k'2}^{ABCD}$$

$$= \hat{\lambda}_{12k'1}^{ABCD} = \hat{\lambda}_{21k'1}^{ABCD} = -\hat{\lambda}_{11k'1}^{ABCD}$$

Eğer  $\lambda_V^T$  parametresi sıfır ve eğer  $n^T$  marginal toplamlar tablosu tabitleştirilmemiş ise o zaman  $\hat{\lambda}_V^T / \sqrt{S} (\hat{\lambda}_V^T)$  standardize değeri yaklaşık normal dağılış gösterir. Böyle bir test sonucu  $\lambda_1^C, \lambda_2^C, \lambda_1^D, \lambda_{12}^{BC}, \lambda_{11}^{CD}$  ve  $\lambda_{31}^{CD}$  için hesaplanan değerler normal dağılış tablo değerlerinden büyük olduğu durumda bu parametrelerin sıfır olmadığı varsayılır. Dolayısıyla bu parametrelerden hiçbirini sıfır kabul etmeyen en basit hiyerarşik model,

$$\log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{jk}^{BC} + \lambda_{kl}^{CD} \quad (3.3-2)$$

olur.  $\lambda_i^A, \lambda_j^B$  ve  $\lambda_{ij}^{AB}$  terimleri örnekleme işleminde  $n_{ij}^{AB}$  marginal toplamları üzerindeki kısıtlamaları yansıtmak için modele ilave edilir.  $\hat{\lambda}_{111}^{BCD}$  ve  $\hat{\lambda}_{131}^{BCD}$ , sıfır varsayılmazsa,  $\lambda_{jkl}^{BCD}$  terimi de ilave edilerek model,

$$\log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{jk}^{BC} + \lambda_{kl}^{CD} + \lambda_{j1}^{BD} + \lambda_{jkl}^{BCD} \quad (3.3-3)$$

olur. Bu şekilde başka modellerde belirlenebilir.

### 3.3.1.2. Doymamış Modellerde Enbüyük-Olabilirlik Tahminleri Hesaplanması İçin Kapalı-Formda İfadeler

Hem (3.3-2) hem de (3.3-3),  $\hat{m}_{ijkl}$  enbüyük-olabi-

lik tahminlerinin kapalı-formda ifade edilebildiği hiyerarşik modellerdir. (3.3-2) de

$$\hat{m}_{ijkl} = \frac{n_{ij}^{AB} n_{jk}^{BC} n_{kl}^{CD}}{n_j^B n_k^C} \quad (3.3-4)$$

olur. (3.3-3) de

$$\hat{m}_{ijkl} = \frac{n_{ij}^{AB} n_{jkl}^{BCD}}{n_j^B} \quad (3.3-5)$$

olur.

Eğer enbüyük-olabilirlik tahminleri için kapalı-formda bir ifade uygun ise, o zaman iterative oransal aydurma algoritması bu tahminleri, normal olarak ilk döngü sonunda verir. Eğer  $\hat{m}_{ijkl}$  için kapalı form ifadeleri mevcut ise, bu durumda başka faydalarda sağlanabilir. Bunlar; Model, bağımsızlık, şartlı bağımsızlık ve eşit olasılık hipotezleri terimleri cinsinden yorumlanabilir. Tahmin edilen asimtotik standart sapmalar, iterative hesaplamalar veya Newton-Raphson algoritması kullanılmadan bulunabilir. Düzeltilmiş rezidüeller basit formda ifade edilebilir. Bazı model karşılaştırmaları, tabloların tamamından ziyade marginal tabloların tamamından ziyade marginal tablolar kullanarak yapılabilir.

Kapalı-form tahminleri birçok farklı alt indis içerebilir. Notasyonu basitleştirmek için (3.3-4) ifadesindeki indisler kaldırılarak,

$$\hat{m} = \frac{n^{AB} \times n^{BC} \times n^{CD}}{n^B \times n^C}$$

gibi kısaltılmış bir ifade kullanılabilir. Benzer şekilde (3.3-5) de

$$\hat{m} = \frac{n^{AB} \times n^{BCD}}{n^B}$$

olarak yazılabilir.

$\hat{m}_{ijkl}$ 'lerin kesin ifadeli formülleri belirleme de, model tanımlayan sınıfın  $T(a)$  elemanlarından meydana gelmiş olduğu kabul edilirse, burada  $1 \leq a \leq b$  dir. Doymuş modelde  $b=1$  ve  $T(1)= ABCD$  olur. (3.3-2) de,  $b= 3$  dür.  $T(1)= AB$ ,  $T(2)= BC$ , ve  $T(3)= CD$  olur. (3.3-3) de,  $b= 2$  dir.  $T(1)= AB$ ,  $T(2)= BCD$  olur.

$U$ , bir  $T(a)$ ,  $1 \leq a \leq b$ , daki tüm harflerden meydana gelmiş olsun. Doymuş modelde, (3.3-2) veya (3.3-3) de  $U= ABCD$  dir. Her değişken  $U$ 'daki bir harfe tekabül ederse,  $d= 1$  olur. Böylece (3.3-2) (3.3-3) veya doymuş modelde  $d= 1$  dir.

Diğer durumlarda,  $U$ 'da gösterilmeyen değişkenlerin kategorisi sayılarının çarpımı  $d$ 'yi verir. Örneğin eğer  $U= AB$  ise, o zaman  $d= 3 \times 2= 6$  olur. Burada 3, C'nin kategori sayısı ve 2, D'nin kategori sayısıdır. Eğer  $a > 1$  ise,  $V(a)$ , bir  $T(a')$ ,  $a' < a$ , da da olan  $T(a)$  daki tüm harflerden ibarettir, ve  $C(a)$ ,  $V(a)$  daki her harfin  $T(C(a))$  da da olduğu en küçük tam sayısıdır. (3.3-2) de  $V(2)= B$ ,  $C(2)=1$ ,  $V(3)= C$  ve  $C(3)= 2$  olur. (3.3-3) de,  $V(2)= B$  ve  $C(2)= 1$  olur.

Eğer  $b= 1$  ise, o zaman

$$\hat{m} = \frac{1}{d} n^{T(1)} \quad (3.3-6)$$

olur. Örneğin  $n_{ijkl}^{ABCD} = n_{ijkl}$  olduğu doymuş modelde,  $d= 1$  ve

$$\hat{m}_{ijkl} = n_{ijkl}$$

olur. Eğer  $b=2$  ise, o zaman

$$\hat{m} = \frac{1}{d} n^{T(1)} \times n^{T(2)} / n^{V(2)} \quad (3.3-7)$$

olur. Eğer  $V(2)$  boş ise, o zaman  $n^{V(2)} = N$  dir. (3.3-3) de,  $d=1, T(1)=AB, T(2)=BCD$  ve  $V(2)=B$  olur. Böylece

$$\hat{m}_{ijkl} = n_{ij}^{AB} n_{jkl}^{BCD} / n_j^B$$

olur. Eğer  $b \geq 3$  ve  $2 \leq a \leq b$  için  $c(a) < a$  ise, o zaman

$$\hat{m} = \frac{1}{d} n^{T(1)} \times \left( \frac{n^{T(2)}}{n^{V(2)}} \times \dots \times \left( \frac{n^{T(b)}}{n^{V(b)}} \right) \right) \quad (3.3-8)$$

olur.

(3.3-2)'de,  $d=1, T(1)=AB, T(2)=BC, T(3)=CD, V(1)=B, V(2)=C, C(2)=1$  ve  $C(3)=2$  olur, böylece

$$\hat{m} = n^{AB} (n^{BC} / n^B) (n^{CD} / n^C)$$

veya

$$\hat{m}_{ijkl} = n_{ij}^{AB} n_{jk}^{BC} n_{kl}^{CD} / n_j^B n_k^C$$

dır.

Daha genel olarak anlatılacak olursa, bağımsız ve benzer dağılımlı  $(A_h, B_h, C_h, D_h)$  gözlemlerinin,  $1 \leq h \leq N$ , karşılıklı sınıflandığı,  $n_{ijkl}$  değerlerine sahip  $r \times s \times t \times v$  boyutlu bir tablo dikkate alalım.  $n_{ijkl}, A_h = i, B_h = j, C_h = k$  ve  $D_h = l$  olan birey sayısıdır.  $p_{ijkl} = n_{ijkl} / n$ ,  $A_h = i, B_h = j, C_h = k, D_h = l$  olması olasılığıdır,  $1 \leq A_h \leq r, 1 \leq B_h \leq s, 1 \leq C_h \leq t$  ve

$1 \leq D_h \leq v$  olduğu varsayılır. Aşağıdaki hiyerarşik modeller için  $n_{ijkl}$ 'nin beklenen değeri olan  $m_{ijkl} = Np_{ijkl}$  için enbüyük-olabilirlik tahminlerini bulalım:

$$(a) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B$$

$$(b) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D$$

$$(c) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{il}^{AD}$$

$$(d) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{kl}^{CD}$$

(a) için, model tanımlayan sınıf  $T(1) = A$ ,  $T(2) = B$  den oluşur. Böylece  $V(2)$  boştur.  $U = AB$  ve  $d = tu$  dur. (3.3-7) izlenerek,

$$m_{ijkl} = \frac{n_i^A n_j^B}{tu N}$$

olur.

(b) için, burada model tanımlayan sınıf  $T(1) = A$ ,  $T(2) = B$ ,  $T(3) = C$ ,  $T(4) = D$  den oluşur. Böylece  $V(2)$ ,  $V(3)$  ve  $V(4)$  boştur.  $U = ABCD$ ,  $d = 1$  ve  $C(2) = C(3) = C(4) = 1$  dir. (3.3-8) izlenerek,

$$\hat{m}_{ijkl} = n_i^A (n_j^B/N)(n_k^C/N)(n_l^D/N) = \frac{n_i^A n_j^B n_k^C n_l^D}{N^3}$$

olur.

(c) için, model tanımlayan sınıf  $T(1) = AB$ ,  $T(2) = AC$ ,  $T(3) = AD$  den oluşur. Böylece  $V(2) = V(3) = A$ ,  $U = ABCD$ ,  $d = 1$ , ve  $C(2) = C(3) = 1$  olur. (3.3-8) izlenerek

$$\hat{m}_{ijkl} = n_{ij}^{AB} (n_{ik}^{AC} / n_i^A) (n_{il}^{AD} / n_i^A) = \frac{n_{ij}^{AB} n_{ik}^{AC} n_{il}^{AD}}{(n_i^A)^2}$$

olur.

(d) için, model tanımlayan sınıf  $T(1) = AB$  ve  $T(2) = CD$  den oluşur. Böylece  $V(2) = \text{boş}$ 'tur,  $U = ABCD$  dir,  $d = 1$  dir.

(3.3-7) izlenerek

$$\hat{m}_{ijkl} = n_{ij}^{AB} n_{kl}^{CD} / N$$

olur.

Hiyerarşik modellerin kapalı-formda ifadelerle sahip olmadığını belirlemek için aşağıdaki modelleri yorumlayalım;  $n_{ijkl}$ , yine  $A_h = i$ ,  $B_h = j$ ,  $C_h = k$  ve  $D_h = l$  olan birey sayısıdır.

$$(a) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC},$$

$$(b) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{kl}^{CD}$$

$$(c) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jl}^{BC} + \lambda_{ijk}^{ABC},$$

$$(d) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{il}^{AD} + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD} + \lambda_{ijk}^{ABC} + \lambda_{ijl}^{ABD} + \lambda_{jkl}^{BCD}$$

(a) için, model tanımlayan sınıf AB, AC ve BC'den oluşur. Bu üç harf kombinasyonunun nasıl düzenlendiği önemli değildir.  $C(3) = 3$  olduğundan  $\hat{m}_{ijkl}$ , kapalı-formda ifade edilemez.

(b) için, model tanımlayan sınıf  $T(1) = A$ ,  $T(2) = B$  ve  $T(3) = CD$ 'den oluşur. Böylece  $V(1)$  ve  $V(2)$  boştur.  $C(2) = C(3) = 1$  dir. Böylece  $\hat{m}_{ijkl}$ , kapalı formda ifade edilebilir.

(c) için, model tanımlayan sınıf ABC ve D den oluşur. Sadece iki eleman olduğundan kapalı-formda ifadeler uygun olur.

(d) için, model tanımlayan sınıf ABC, ABD ve BCD'den oluşur. (a) daki gibi  $T(1)$ ,  $T(2)$  ve  $T(3)$ 'ün herhangi tanımı altında  $C(3) = 3$  dür. Kapalı formda ifadeler uygun değildir.

### 3.3.1.3. Uyum Testi

$$L^2 = 2 \sum_i \sum_j \sum_k \sum_l n_{ijkl} \log (n_{ijkl} / \hat{m}_{ijkl})$$

ve

$$\chi^2 = \sum_i \sum_j \sum_k \sum_l (n_{ijkl} - \hat{m}_{ijkl})^2 / \hat{m}_{ijkl}$$

Khi-kare istatistikleri (3.3-2) veya (3.3-3) ile ifade edilen modeller için kolayca hesaplanır. Serbestlik derecesini bulmak için Bölüm 3.2'deki yöntem kullanılır. (3.3-2) de  $\lambda_{ik}^{AC}$ ,  $\lambda_{il}^{AD}$ ,  $\lambda_{jl}^{BD}$ ,  $\lambda_{ijk}^{ABC}$ ,  $\lambda_{ijl}^{ABD}$ ,  $\lambda_{ikl}^{ACD}$ ,  $\lambda_{jkl}^{BCD}$  ve  $\lambda_{ijkl}^{ABCD}$  parametrelerinin hepsi sıfır kabul edilir. Serbestlik derecesi,  $b(AC) + b(AD) + b(BD) + b(ABC) + b(ABD) + b(ACD) + b(BCD) + b(ABCD)$  dir.  $A_h, 2$ ,  $B_h, 2$ ,  $C_h, 3$  ve  $D_h, 2$  kategoriye sahip olduğundan,

$$b(AC) = (2-1)(3-1) = 2$$



$$b(AD) = (2-1)(2-1) = 1$$

$$b(BD) = (2-1)(2-1) = 1$$

$$b(ABC) = (2-1)(2-1)(3-1) = 2$$

$$b(ABD) = (2-1)(2-1)(2-1) = 1$$

$$b(ACD) = (2-1)(3-1)(2-1) = 2$$

$$b(BCD) = (2-1)(3-1)(2-1) = 2$$

$$b(ABCD) = (2-1)(2-1)(3-1)(2-1) = 2$$

olur. Böylece 13 serbestlik derecesi olur. Modelin verilere uyup uymadığı belirlenebilir. (3.3-3), (3.3-2) den daha az kısıtlı bir modeli belirlediğinden, (3.3-3) ile belirtilen modelinde verilere uyduğu varsayılabilir.

İki model için olabilirlik-oran istatistikleri arasındaki fark da test edilir. Bu fark,

$$\begin{aligned} & 2 \sum_i \sum_j \sum_k \sum_l n_{ijkl} \log \left( \frac{n_{ijkl} n_j^B n_k^C}{n_{ij}^{AB} n_{jk}^{BC} n_{kl}^{CD}} \right) \\ & - 2 \sum_i \sum_j \sum_k \sum_l n_{ijkl} \log \left( \frac{n_{ijkl} n_j^B}{n_{ij}^{AB} n_{jkl}^{BCD}} \right) \\ & = 2 \sum_i \sum_j \sum_k \sum_l n_{ijkl} \log \left( \frac{n_{jkl}^{BCD} n_k^C}{n_{jk}^{BC} n_{kl}^{CD}} \right), \\ & = 2 \sum_j \sum_k \sum_l n_{jkl}^{BCD} \log \left( \frac{n_{jkl}^{BCD} n_k^C}{n_{jk}^{BC} n_{kl}^{CD}} \right) \end{aligned}$$

dir. Bu, C verildiğinde B ve D nin şartlı bağımsızlığı için bir test istatistiği oluşturur. Bu fark  $\hat{\lambda}_{ik}^{AD}$  ve  $\hat{\lambda}_{ijl}^{ACD}$

nin (0) olup olmadığını belirler.

### 3.3.1.4. Modellerin Yorumu

GOODMAN (1970), kapalı-formda enbüyük-olabilirlik tahminine sahip her hiyerarşik modelin bağımsızlık, şartlı bağımsızlık ve eşit olasılık terimleri cinsinden yorumlanabileceğini bildirir. Diğer durumda yorum daha karmaşıklık arzeder. Aşağıdaki gibi bir gözlem genelde faydalı olur. Her harf  $U(g)$ ,  $1 \leq g \leq h$ ,  $h \geq 2$  veya  $W$  deki bir değişkeni gösterir. Hiçbir harf hem  $U(g)$  hem de  $U(f)$ ,  $g \neq f$ , veya hem  $U(g)$  ve hem de  $W$ ,  $1 \leq g \leq h$ , de olmaz.  $\lambda^T$  parametresi sadece  $g \neq f$  için  $T$  deki bir harf  $U(g)$  de de olursa ve  $T$ 'deki bir harf  $U(f)$  de de olursa sıfır varsayılır.

Hiçbir  $U(g)$  boş değildir. Eğer  $W$  boş ise, o zaman hiyerarşik model, farklı  $U(g)$  ile gösterilen değişkenlerin karşılıklı bağımsız olduğu modele eşdeğerdir. Diğer durumda,  $W$  ile gösterilen değişkenler verildiğinde, model, farklı  $U(g)$ 'ye tekabül eden değişkenlerin şartlı olarak bağımsız olduğu modele eşdeğerdir. Örnek olarak (3.3-3)'ü dikkate alalım.  $U(1) = A$ ,  $U(2) = CD$  ve  $W = B$  olsun. Eğer  $T$ ,  $A$  ile birlikte  $C$  yada  $D$  yi bulundurursa her  $\lambda^T$  parametresinin sıfır olduğu belirtilir. Böylece (3.3-3), aşağıdaki iddialara eşdeğer olur;

(a)  $B_h = j$  olduğu verildiğinde,  $C_h$  ve  $D_h$ ,  $A_h$  den bağımsızdır. (3.3-2) deki model ancak (3.3-3)'deki model sağlanırsa ve

$$\log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{kl}^{CD} + \lambda_{ik}^{AC} + \lambda_{ijk}^{ABC} \quad (3.3-9)$$

olursa geçerlidir.

(3.3-3),  $\lambda_{ijk}^{ABC}$  ve  $\lambda_{ik}^{AC}$  nin sıfır olmasını gerektirir. Bu kısıtlamalar verildiğinde (3.3-9), (3.3-2) ye indirgenir. (3.3-9) da  $U(1)=AB$ ,  $U(2)=D$  ve  $W=C$  olur. O zaman bir  $\lambda^T$  -parametresi, ancak  $C$ ,  $T$ 'de olduğunda ve  $A$  veya  $B$ ,  $T$  de olduğunda sıfır varsayılır, ve

(b)  $C_h = k$  olduğu verildiğinde  $D_h$ ,  $A_h$  ve  $B_h$ 'dan bağımsızdır. (3.3-2) ve (3.3-3) ile ifade edilen modeller için olasılık-oran khi-karelerinin analizi (a) iddiasının verilere uygun olduğu fakat (a) ve (b) iddialarının kombinasyonunun uygun olmayabileceği ortaya çıkar. Böylece (b) ye karşı delil vardır. Olabilirlik-oran khi-karelerinin karşılaştırılması,  $C_h$  verildiğinde  $B_h$  ve  $D_h$  arasında ilişki olup olmadığını ortaya koyar.

$$\log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD} \quad (3.3-10)$$

modelinin, eğer (3.3-3) sağlanırsa ve  $\lambda_{jkl}^{ACD} = 0$  ise sağlandığı ortaya çıkar. (3.3-3)'deki model altında,

$$\begin{aligned} \tau_{(kk')(12).ij}^{CD.AB} &= \log \left( \frac{m_{ijkl} m_{ijk'2}}{m_{ijk2} m_{ijk'2}} \right) \\ &= (\lambda_{k1}^{CD} - \lambda_{k'1}^{CD} - \lambda_{k2}^{CD} + \lambda_{k'2}^{CD}) \\ &\quad + (\lambda_{jkl}^{BCD} - \lambda_{jk'1}^{BCD} - \lambda_{jk2}^{BCD} \\ &\quad \quad \quad + \lambda_{jk'2}^{BCD}) \end{aligned}$$

olduğundan,  $1 \leq k < k' \leq 3$ , (3.3-9), (a) iddiası sağlandığında ve  $\frac{CD \cdot AB}{(kk')(12) \cdot ij}$  çapraz-çarpım oranları,  $A_h = i$  ve  $B_h = j$  bağımsız olduğunda sağlanır. Böylece (3.3-9), C ve D arasındaki interaksiyonun,  $A_h$  ve  $B_h$ 'dan bağımsız olmasını gerektirir. Eşdeğer olarak, (3.3-9), ancak (a) sağlanırsa ve

$$\frac{CD \cdot AB}{(12)(12) \cdot ik} = \log \left( \frac{m_{ilk1} m_{i2k2}}{m_{ilk2} m_{i2k1}} \right)$$

çapraz-çarpım oranları,  $A_h$  ve  $C_h$  dan bağımsız ise sağlanır.

Aşağıdaki modelleri bağımsızlık, şartlı bağımsızlık ve eşit olasılık terimleri cinsinden yorumlayalım;

$$(a) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B$$

Burada  $U(1) = A$ ,  $U(2) = B$  ve  $U(3) = CD$  ve  $W$ , boştur. Tekabül eden model,

$$\log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{kl}^{CD}$$

dir. Modelde sadece  $p_{ijkl}$ ,  $k$  ve  $l$  den bağımsız ise, yani

$$\frac{CD \cdot AB}{p_{kl \cdot ij}} = p_{ijkl} / p_{kl}^{CD} = 1/(tu),$$

$\lambda_k^C = 0$ ,  $\lambda_l^D = 0$ , ve  $\lambda_{kl}^{CD} = 0$  olması gerekir. Böylece  $A_h$  ve  $B_h$  bağımsızdır.  $A_h$  ve  $B_h$  verildiğinde tüm  $C_h$  ve  $D_h$  kombinasyonları eşit olarak olasıdır.

$$(b) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D,$$

modelinde  $U(1)=A$ ,  $U(2)=B$ ,  $U(3)=C$  ve  $U(4)=D$  ve  $W$  boştur. O zaman, bu modele,  $A_h$ ,  $B_h$ ,  $C_h$  ve  $D_h$ 'in karşılıklı olarak bağımsızlığı modeli tekabül eder.

$$(c) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{il}^{AD}$$

modelinde  $U(1)=B$ ,  $U(2)=C$ ,  $U(3)=D$  ve  $W=A$  dır. O zaman, bu modele  $A_h$  verildiğinde  $B_h$ ,  $C_h$  ve  $D_h$ 'in şartlı olarak bağımsızlığı modeli tekabül eder.

$$(d) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{kl}^{CD}$$

modelinde  $U(1)=AB$ ,  $U(2)=CD$  ve  $W$  boştur. O zaman model,  $(A_h, B_h)$  ve  $(C_h, D_h)$  çiftlerinin bağımsız olduğu bir modeldir.

(3.3-3) ile ifade edilen model altında parametre tahmini yapılmak istendiğinde, ilgilenilen esas parametreler  $\zeta_{(12)(12).ik}^{BD.AC}$ ,  $\zeta_{(12)(12).ij}^{CD.AB}$ ,  $\zeta_{(13)(12).ij}^{CD.AB}$  ve  $\zeta_{(23)(12).ij}^{CD.AB}$  çapraz-çarpım oranlarıdır. Burada,

$$\begin{aligned} \hat{\zeta}_{(12)(12).ik}^{BD.AC} &= \log \left( \frac{\hat{m}_{ilk1} \hat{m}_{i2k2}}{\hat{m}_{ilk2} \hat{m}_{i2k1}} \right) \\ &= \log \left[ \frac{(n_{i1}^{AB} n_{1kl}^{BCD} / n_1^B) (n_{i2}^{AB} n_{2k2}^{BCD} / n_2^B)}{(n_{i1}^{AB} n_{1k2}^{BCD} / n_1^B) (n_{i2}^{AB} n_{2k1}^{BCD} / n_2^B)} \right] \end{aligned}$$

$$= \log \left[ \frac{n_{1k1}^{BCD} n_{2k2}^{BCD}}{n_{1k2}^{BCD} n_{2k1}^{BCD}} \right]$$

$$= \hat{\tau}_{(12)(12).k}^{BD.C}$$

olur. Böylece

$$s^2(\hat{\tau}_{(12)(12).ik}^{BD.AC}) = s^2(\hat{\tau}_{(12)(12).k}^{BD.C}) = \frac{1}{n_{1k1}^{BCD}}$$

$$+ \frac{1}{n_{2k1}^{BCD}} + \frac{1}{n_{2k2}^{BCD}} + \frac{1}{n_{1k2}^{BCD}}$$

olur. Benzer şekilde,

$$\tau_{(kk')(12).ij}^{CD.AB} = \log \left[ \frac{n_{jk1}^{BCD} n_{jk'2}^{BCD}}{n_{jk2}^{BCD} n_{jk'1}^{BCD}} \right] = \hat{\tau}_{(kk')(12).j}^{CD.B}$$

ve

$$s^2(\tau_{(kk')(12).ij}^{CD.AB}) = s^2(\hat{\tau}_{(kk')(12).j}^{CD.B}) = \frac{1}{n_{jkl}^{BCD}}$$

$$+ \frac{1}{n_{jk2}^{BCD}} + \frac{1}{n_{jk'1}^{BCD}} + \frac{1}{n_{jk'2}^{BCD}}$$

olur. Tahminler  $A_h = i$  den bağımsızdır.

### 3.3.2. 3x3x3x3 Boyutlu Tablolar

Kısım 3.3.1. de tanımlanan temel işlemler, üzerinde durulan değişkenlerden hiçbirisinin iki şıklı olmadığı

dört yönlü tablolara kolaylıkla uygulanır. Çapraz-çarpım oranlarının kontrolü daha karmaşık olur. Rezidüyal analizi yapılmalıdır. Bununla beraber, temel prensibler değişmez.

Böyle bir tabloda,  $n_{ijkl}$ ,  $A_h=i$ ,  $B_h=j$ ,  $C_h=k$  ve  $D_h=1$  olan bireylerin sayısıdır.  $m_{ijkl}$  beklenen değerleri veya  $n_{ijkl}$ 'ler  $A_h$ ,  $B_h$ ,  $C_h$  ve  $D_h$ 'ın tüm olası kombinasyonları için pozitiftir.  $n = \{n_{ijkl}\}$ , tam faktöriyel bir tablodur.  $A_h=i$  olduğu verildiğinde,  $B_h=j$ ,  $C_h=k$  ve  $D_h=1$  olmasının şartlı olasılığı  $p_{jkl.i}^{BCD.A} > 0$  dır. Tüm  $A_h$ 'lar verildiğinde,  $1 \leq h \leq N$ ,  $(B_h, C_h, D_h)$  üçlülerinin bağımsız olduğu varsayılır.  $n_i^A$  toplamının sabit olduğu verildiğinde  $m_{ijkl} = n_i^A p_{jkl.i}^{BCD.A}$  olur ve her  $n_{ijkl}$  değerler tablosu, bağımsız multinomial dağılımı gösterir,  $1 \leq j \leq 3$ ,  $1 \leq k \leq 3$ ,  $1 \leq l \leq 3$ .

### 3.3.2.1. Doymuş Model

Kısım 3.3.1.'de olduğu gibi analize, doymuş model

$$\begin{aligned} \log m_{ijkl} = & \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} \\ & + \lambda_{il}^{AD} + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD} + \lambda_{ijk}^{ABC} \\ & + \lambda_{ijl}^{ABD} + \lambda_{ikl}^{ACD} + \lambda_{jkl}^{BCD} + \lambda_{ijkl}^{ABCD} \end{aligned}$$

kontrolüyle başlamak uygun olur. Burada,  $\lambda$ - parametreleri genel kısıtlamalara tabidir.  $\hat{\lambda}$ -parametrelerinin tahminleri kısım 3.3.1 deki gibi bulunabilir.  $n_{ijkl}$  değerleri küçük olduğunda  $n_{ijkl}$  den ziyade  $n_{ijkl} + 1/2$  kullanılır. Böylece

$\hat{m}_{ijkl} = \log(n_{ijkl} + 1/2)$  olur.  $3 \times 3 \times 3 \times 3$  boyutlu bir tablo ile

ilgilenildiğinde,

$$\hat{\lambda} = \frac{1}{3.3.3.3} \sum_i \sum_j \sum_k \sum_l \hat{m}_{ijkl},$$

$$\hat{\lambda}_{1, D} = \frac{1}{3.3.3.3} \sum_i \sum_j \sum_k \sum_l (3 f_{11, -1}) \hat{m}_{ijkl},$$

$$\hat{\lambda}_{k, 1}^{CD} = \frac{1}{3.3.3.3} \sum_i \sum_j \sum_k \sum_l (3 f_{kk, -1})(3 f_{11, -1}) \hat{m}_{ijkl},$$

$$\hat{\lambda}_{j'k'1'}^{BCD} = \frac{1}{3.3.3.3} \sum_i \sum_j \sum_k \sum_l (3 f_{kk, -1})(3 f_{jj, -1})(3 f_{11, -1}) \hat{m}_{ijkl}$$

olur.

$\hat{\lambda}$ -parametrelerinin sıfırdan farklı olup olmadığı standardize değerlerin test edilmesiyle bulunur.

Bu tip tablolarda kontrol edilebilecek bazı hiyerarşik modeller aşağıdaki şekilde sıralanabilir;

$$(1) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{jl}^{BD} + \lambda_{kl}^{CD}$$

$$(2) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD},$$

$$(3) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD} + \lambda_{jkl}^{BCD}$$



$$(4) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{il}^{AD} \\ + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD} + \lambda_{jkl}^{BCD}$$

$$(5) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} \\ + \lambda_{il}^{AD} + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD} \\ + \lambda_{jkl}^{BCD}$$

$$(6) \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{ij}^{AB} \\ + \lambda_{il}^{AD} + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD} \\ + \lambda_{ijl}^{ABD} + \lambda_{jkl}^{BCD}$$

Model 1, 3x3x3x3 tipi tablolarda mutlaka kullanılır. Model 3, model 1'in tüm parametreleri ile birlikte  $\lambda_{jkl}^{BCD}$  ve  $\lambda_{jkl}^{BC}$ 'yi ihtiva ediyor.  $\lambda_{jk}^{BC}$  sadece hiyerarşi esasını korumak için konulmuştur. Model 2'de sadece  $\lambda_{jk}^{BC}$  ilave edilmiştir. Bu da,  $\lambda_{jk}^{BC}$  ve  $\lambda_{jkl}^{BCD}$ 'nin nisbi önemini göstermek için gereklidir. Model 4'de de model 3'e  $\lambda_{il}^{AD}$  eklenmiştir. Model 6'da model 4'e  $\lambda_{ijl}^{ABD}$  ve  $\lambda_{ij}^{AB}$  eklenmiştir, bu son terim yine hiyerarşiyi sağlamak için eklenmiştir. Model 5'de ise model 4'e  $\lambda_{ij}^{AB}$  eklenmiştir.

Bu modeller artan karmaşıklık sırasına bağlı olarak seçilirler. Model 5 ve model 2 için iterative oransal uydurma algoritması kullanılır, çünkü model 1, 3, 4 ve 6 enbüyük-olabilirlik tahminleri için kesin ifadelerle sahip-

tirler. Böylece üzerinde durulan 6 modelden 4'ü için, bu algoritma tek devre sonunda enbüyük-olabilirlik tahminine yaklaşır.

Çizelge 12- Model 1 - Model 6 İçin Beklenen Değerler İçin Kesin İfadeler.

Model	Model tanımlayan sınıf	Başlangıç iterasyonları <sup>a</sup>	$m_{ijkl}$
1	A, BD, CD	$m_{ijkl1} = m_{ijkl0} n_i^A / m_{i0}^A$ $m_{ijkl2} = m_{ijkl1} n_{jl}^{BD} / m_{j11}^{BD}$ $m_{ijkl3} = m_{ijkl2} n_{kl}^{CD} / m_{kl2}^{CD}$	$\frac{n_i^A n_{jl}^{BD} n_{kl}^{CD}}{N n_1^D}$
2	A, BC, BD, CD	$m_{ijkl} = m_{ijkl0} n_i^A / m_{i0}^A$ $m_{ijkl2} = m_{ijkl1} n_{jkl}^{BC} / m_{jkl}^{BC}$ $m_{ijkl3} = m_{ijkl2} n_{jl}^{BD} / m_{jl2}^{BD}$ $m_{ijkl4} = m_{ijkl3} n_{kl}^{CD} / m_{kl3}^{CD}$	Hiç bir keskin eşitlik yok
3	A, BCD	$m_{ijkl1} = m_{ijkl0} n_i^A / m_{i0}^A$ $m_{ijkl2} = m_{ijkl1} n_{jkl}^{BCD} / m_{jkl1}^{BCD}$	$\frac{n_i^A n_{jkl}^{BCD}}{N}$
4	AD, BCD	$m_{ijkl1} = m_{ijkl0} n_{il}^{AD} / m_{i10}^{AD}$ $m_{ijkl2} = m_{ijkl1} n_{jkl}^{BCD} / m_{jkl1}^{BCD}$	$\frac{n_{il}^{AD} n_{jkl}^{BCD}}{n_1^D}$

## Çizelge 12 - Devam

5	AD, AB, BCD	$m_{ijkl1} = m_{ijk10} n_{il}^{AD} / m_{i10}^{AD}$ $m_{ijkl2} = m_{ijkl1} n_{ij}^{AB} / m_{ij1}^{AB}$ $m_{ijkl3} = m_{ijkl2} n_{jkl}^{BCD} / m_{jkl2}^{BCD}$	Hiç bir kesin eşitlik yok
6	ABD, BCD	$m_{ijkl1} = m_{ijk10} n_{ijl}^{ABD} / m_{ij10}^{ABD}$ $m_{ijkl2} = m_{ijk10} n_{jkl}^{BCD} / m_{jkl10}^{BCD}$	$\frac{n_{ijl}^{ABD} n_{jkl}^{BCD}}{n_{j1}^{BD}}$

<sup>a</sup>Her durumda  $m_{ijk0} = 1$  kabul edilir. Model 1 de  $m_{ijkl3} = \hat{m}_{ijkl}$ , model 3,4 ve 6'da  $m_{ijkl2} = \hat{m}_{ijkl}$  dir.

## 3.3.2.2. Khi-Kare İstatistikleri

Tüm olası modeller için  $\chi^2$  ve  $L^2$  istatistikleri ayrı ayrı hesaplanır.  $u$ ,  $L^2$ 'ye ait serbestlik derecesi ise  $L^2/u$ 'den F istatistikleri bulunur. F değeri 1'den küçük olan modellerin hepsi iyi uyum sağlayan modellerdir.  $L^2$  istatistikleri arasındaki fark test edilerek hangi modelin verilere daha uygun olduğu tayin edilir. Serbestlik derecesi, yine sıfır varsayılan  $\lambda$ -parametrelerine ait serbestlik dereceleri toplamıdır, ve her model için ayrıca hesaplanır.

Model 4, 5 ve 6'nın F'leri 1'den küçük olsun ve hepsi iyi uyum sağlasın. Bunlar arasındaki  $L^2$  farkları önemli değilse model 4 tercih edilir. Model 4'ün sağlandığı verildiğinde, en azından yaklaşık olarak model 4 ve model

3 için  $L^2$  istatistikleri arasındaki fark büyük ise, bu  $\lambda_{il}^{AD}$  interaksyonunun sıfır olmadığını göstermektedir.  $\lambda_{jkl}^{ABCD}$ 'nin varlığını test etmek için model 2' gibi bir model'in  $L^2$  si gerekir.

$$(2') \log m_{ijkl} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_l^D + \lambda_{il}^{AD} + \lambda_{jk}^{BC} + \lambda_{jl}^{BD} + \lambda_{kl}^{CD}$$

Bu model, model 2 ile enbüyük-olabilirlik tahminleri bakımından yakın bir ilişkiye sahiptir, ancak  $\lambda_{il}^{AD}$  farklıdır.  $\lambda_{jkl}^{BCD}$  ilişkisi, model 3 ve model 2 nin karşılaştırılmasında uygun  $L^2$  istatistiklerinin kullanılması ile test edilebilir. Model 2 ile model 3'ün  $L^2$ 'leri arasındaki fark,

$$\begin{aligned} & 2 \sum_i \sum_j \sum_k \sum_l n_{ijkl} \log \left[ \frac{n_i^A n_{jkl}^{BCD}/n}{n_i^A \hat{m}_{jkl}^{BCD}/n} \right] \\ & = 2 \sum_j \sum_k \sum_l n_{jkl}^{BCD} \log \left[ \frac{n_{jkl}^{BCD}}{\hat{m}_{jkl}^{BCD}} \right] \end{aligned}$$

olur.

Aynı şekilde model 4 ile model 2' 'nün  $L^2$ 'leri arasındaki fark da

$$\begin{aligned} & 2 \sum_i \sum_j \sum_k \sum_l n_{ijkl} \log \left[ \frac{n_{il}^{AD} n_{jkl}^{BCD}/n_l^D}{n_{il}^{AD} \hat{m}_{jkl}^{BCD}/n_l^D} \right] \\ & = 2 \sum_j \sum_k \sum_l n_{jkl}^{BCD} \log \left[ \frac{n_{jkl}^{BCD}}{\hat{m}_{jkl}^{BCD}} \right] \end{aligned}$$

olduğundan, test neticeleri aynı sonuca götürür.

Model 4'ün verilerle uyum sağladığı verildiğinde,  $\lambda_{jkl}^{BCD}$  nin sıfır olmadığı ortaya çıkarsa,  $\lambda_{jkl}^{BCD}$  nin sıfır olmadığı hiyerarşik modelde de mutlaka  $\lambda_{jk}^{BC}$  de bulunmalıdır.

Model seçimi ile ilgili olarak burada kullanılan işlemler tek araç değildir. Hiyerarşik modellerin seçimiyle ilgili detaylı bilgi GOODMAN (1971a, 1973c) ve BISHOP, FIENBERG ve HOLLAND (1975) de bulunmaktadır.

### 3.3.2.3. Rezidüyal Analizi

Genelde, düzeltilmiş rezidüyaller  $\hat{m}$ , kapalı-formda bir ifadeye sahip olduğunda, en kolay şekilde hesaplanabilir. Eğer

$$\hat{m} = n^{T(1)} \times \left[ \frac{n^{T(2)}}{n^{V(2)}} \right] \times \dots \times \left[ \frac{n^{T(b)}}{n^{V(b)}} \right]$$

ise, o zaman

$$r_{ijkl} = (n_{ijkl} - \hat{m}_{ijkl}) / \hat{c}_{ijkl}^{1/2}$$

olur, ve sembolik olarak

$$\hat{c} = \hat{m} \left\{ 1 - \hat{m} \sum_{a=1}^b \left[ 1/n^{T(a)} \right] - \hat{m} \sum_{a=2}^b \left[ 1/n^{V(a)} \right] \right\}$$

yazılabilir. Bu formül HABERMAN (1974 a) daki formüller kullanılarak ispat edilebilir. Model 4'de bu formül,

$$n_{jkl}^{BCD} n_{il}^{AD} (n_1^D - n_{jkl}^{BCD}) (n_1^D - n_{il}^{AD}) / (n_1^D)^3$$

$$= \hat{m}_{ijkl} \left[ 1 - \hat{m}_{ijkl} \left( \frac{1}{n_{jkl}^{BCD}} + \frac{1}{n_{il}^{AD}} \right) + \frac{\hat{m}_{ijkl}}{n_1^D} \right]$$

olur.

$$\hat{m}_{ijkl} = \frac{n_{il}^{AD} n_{jkl}^{BCD}}{n_1^D},$$

yerine konursa ,

$$r_{ijkl} = \frac{n_{ijkl} - \hat{m}_{ijkl}}{\left[ (n_{jkl}^{BCD} n_{il}^{AD} (n_1^D - n_{jkl}^{BCD}) (n_1^D - n_{il}^{AD})) / (n_1^D)^3 \right]^{1/2}}$$

elde edilebilir.

Standardize rezidüyaller ,

$$S_{ijkl} = (n_{ijkl} - \hat{m}_{ijkl}) / \hat{m}_{ijkl}^{1/2}$$

ile hesaplanabilir. Her uydurulan model için  $\hat{m}_{ijkl}$ 'ler değişecektir. Bu rezidüyaller  $\hat{m}_{ijkl}$ 'ler kapalı-formda ifadelere sahip olsada olmasada ve Newton-Raphson algoritmi kullanılsada kullanılmasada kolaylıkla hesaplanabilir.

$$\chi^2 = \sum_i \sum_j \sum_k \sum_l S_{ijkl}^2$$

elde edilir. Eğer üzerinde durulan model sağlanıyorsa  $S_{ijkl}$ ,  $u_{ijkl} = c_{ijkl} / m_{ijkl} < 1$  olan yaklaşık  $N(0, u_{ijkl})$  dağılışı gösterir. Eğer  $u_{ijkl}$  1'e yaklaşıyorsa, o zaman  $S_{ijkl}$ ,  $r_{ijkl}$ 'ye benzer.

Model 4'ün seçildiği kabul edilsin. Bu durumda böyle bir modelin yorumu aşağıdaki şekilde yapılabilir.

Model,  $D_h$  verildiğinde  $B_h$  ve  $C_h$ 'ın  $A_h$  dan bağımsız olduğunu ifade eder. Bu netice  $B_h$  ve  $C_h$ 'ın  $A_h$  ve  $D_h$  arasındaki ilişkiyi etkilemediğini ortaya koyar. Benzer şekilde  $A_h$ ,  $B_h$ ,  $C_h$  ve  $D_h$  arasındaki ilişkiyi etkilemez.

$A_h$  ve  $D_h$  arasındaki ilişkiyi analiz etmek için

$$\chi^2_{(ii')(11')}^{AD} = \log \left[ \frac{m_{i1}^{AD} m_{i'1'}^{AD}}{m_{i1'}^{AD} m_{i'1}^{AD}} \right]$$

çapraz-çarpım oranları dikkate alınır.

$m_{i1}^{AD} = n_i^A p_{1.i}^{D.A}$  olduğundan,

$$\begin{aligned} \chi^2_{(ii')(11')}^{AD} &= \log \left( \frac{n_i^A p_{1.i}^{D.A} n_{i'1'}^{A.D.A}}{n_i^A p_{1'.i}^{D.A} n_{i'1}^{A.D.A}} \right) \\ &= \log \left( \frac{p_{1.i}^{D.A} p_{1'.i'}^{D.A}}{p_{1'.i}^{D.A} p_{1.i}^{D.A}} \right) \end{aligned}$$

olduğu izlenir. Burada  $p_{1.i}^{D.A}$ ,  $A_h = i$  olduğu verildiğinde  $D_h = 1$  olması olasılığıdır.

Model 4 altında,  $A_h = i$  ve  $D_h = 1$  olduğu verildiğinde  $B_h = j$  ve  $C_h = k$  olmasının şartlı olasılığı, yani

$$p_{jk.il}^{BC.AD} = p_{jkl.i}^{BCD.A} / p_{1.i}^{D.A}$$

$i$ 'den bağımsız olur. Böylece

$$m_{ijkl} = n_i^A p_{jkl.i}^{BCD.A} = n_i^A p_{1.i}^{D.A} p_{jk.il}^{BC.AD} = n_i^A p_{1.i}^{D.A} p_{jk.il}^{BC.AD}$$

olur.  $B_h$  ve  $C_h$  verildiğinde  $A_h$  ve  $D_h$  için şartlı logaritmik çapraz-çarpım oranı

$$\begin{aligned}
\tau_{(ii')(11')}.jk^{AD.BC} &= \log \left( \frac{m_{ijkl} m_{i'jkl'}}{m_{ijkl'} m_{i'jkl}} \right) \\
&= \log \left( \frac{\frac{n_i^A p_{1.i}^{D.A} p_{jk.il}^{BC.AD}}{n_i^A p_{1'.i}^{D.A} p_{jk.il'}^{BC.AD}} - \frac{n_{i'}^A p_{1'.i'}^{D.A} p_{jk.il'}^{BC.AD}}{n_{i'}^A p_{1'.i'}^{D.A} p_{jk.il}^{BC.AD}}}{\frac{p_{1.i}^{D.A} p_{1'.i'}^{D.A}}{p_{1'.i}^{D.A} p_{1.i}^{D.A}}} \right) \\
&= \log \left( \frac{p_{1.i}^{D.A} p_{1'.i'}^{D.A}}{p_{1'.i}^{D.A} p_{1.i}^{D.A}} \right) \\
&= \tau_{(ii')(11')}^{AD}
\end{aligned}$$

Neticesi olarak,  $\tau_{(ii')(11')}^{AD.BC}$  şartlı logaritmik çapraz-çarpım oranı j ve k dan bağımsızdır, yani  $A_h$  ve  $D_h$  arasındaki interaksiyon  $B_h$  ve  $C_h$  dan bağımsızdır. Bu netice  $A_h=i$ ,  $B_h=j$  ve  $C_h=k$  olduğu verildiğinde  $D_h=1$ 'den ziyade  $D_h=1$  olmasının

$$\tau_{11'.ijk}^{D.ABC} = \log \left( \frac{\frac{p_{1.ijk}^{D.ABC}}{D}}{p_{1'.ijk}} \right)$$

logaritmik risk terimlerinde tanımlanabilir.

$$p_{jkl.i}^{BCD.A} = p_{1.ijk}^{D.ABC} p_{jk.i}^{BC.A}$$

olduğundan,

$$\begin{aligned}
\tau_{(11')}.ijk^{D.ABC} &= \log \left( \frac{\frac{n_i^A p_{1.ijk}^{D.ABC} p_{jk.i}^{BC.A}}{n_i^A p_{1'.ijk}^{D.ABC} p_{jk.i}^{BC.A}}}{\frac{m_{ijkl}}{m_{ijkl'}}} \right) \\
&= \log \left( \frac{m_{ijkl}}{m_{ijkl'}} \right)
\end{aligned}$$

dür. Burada  $p_{jk.i}^{BC.A}$ ,  $A_h=i$  olduğu verildiğinde,  $B_h=j$  ve



$C_h = k$  olmasının şartlı olasılığıdır. Bundan dolayı,

$\hat{\tau}_{(ii')(11')}.jk^{AD.BC}$ ,  $B_h = j$  ve  $C_h = k$  olduğunda  $A_h = i$ 'nin  $A_h = i$  ye değişimi ile neticelenen  $D_h = 1$ ' den ziyade  $D_h = 1$  olmasının logaritmik risk oranları arasındaki farktır. Bu fark  $\hat{\tau}_{11'.ijk}^{D.ABC} - \hat{\tau}_{11'.i'jk}^{D.ABC}$  dir.

$$\text{Genelde } \hat{\tau}_{(ii')(11')} = \log \left( \frac{\frac{n_{il}^{AD} n_{i'l'}^{AD}}{AD}}{n_{il'}^{AD} n_{i'l}^{AD}} \right)$$

ve

$$S^2(\hat{\tau}_{(ii')(11')}^{AD}) = \frac{1}{n_{il}^{AD}} + \frac{1}{n_{il'}^{AD}} + \frac{1}{n_{i'l}^{AD}} + \frac{1}{n_{i'l'}^{AD}}$$

olur.  $\hat{q}_{(ii')(11')}^{AD}$  tahmin edilen çapraz çarpım oranı,

$$\frac{\frac{n_{il}^{AD} n_{i'l'}^{AD}}{AD}}{n_{il'}^{AD} n_{i'l}^{AD}}$$

olur.  $A_h$  ve  $C_h$  verildiğinde,  $B_h$  ve  $D_h$  için

$$\begin{aligned} \hat{\tau}_{(jj')(11')}.ik^{BD.AC} &= \log \left( \frac{m_{ijkl} m_{ij'kl'}}{m_{ijkl'} m_{ij'kl}} \right) \\ &= \log \left( \frac{\frac{P_{1.ijk}^{D.ABC} P_{1'.ij'k}^{D.ABC}}{D.ABC}}{\frac{P_{1'.ijk}^{D.ABC} P_{1.ij'k}^{D.ABC}}{D.ABC}} \right) \end{aligned}$$

logaritmik çapraz çarpım oranının  $i$ 'den bağımsız olduğu bulunur.

$$\hat{\tau}_{(jj')(11')}.k^{BD.C} = \log \left( \frac{\frac{m_{jkl}^{BCD} m_{j'kl'}^{BCD}}{BCD}}{m_{jkl'}^{BCD} m_{j'kl}^{BCD}} \right)$$

dir. Böylece

$$\hat{\zeta}_{(jj')(11'),ik}^{BD.AC} = \hat{\zeta}_{(jj')(11'),k}^{BD.C} = \log \left( \frac{n_{jkl}^{BCD} n_{j'kl'}^{BCD}}{n_{jkl'}^{BCD} n_{j'kl}^{BCD}} \right),$$

$$\hat{q}_{(jj')(11'),ik}^{BD.AC} = \hat{q}_{(jj')(11'),k}^{BD.C} = \frac{n_{jkl}^{BCD} n_{j'kl'}^{BCD}}{n_{jkl'}^{BCD} n_{j'kl}^{BCD}}$$

ve

$$S^2(\hat{\zeta}_{(jj')(11'),ik}^{BD.AC}) = S^2(\hat{\zeta}_{(jj')(11'),k}^{BD.C})$$

$$= \frac{1}{n_{jkl}^{BCD}} + \frac{1}{n_{jkl'}^{BCD}} + \frac{1}{n_{j'kl}^{BCD}} + \frac{1}{n_{j'kl'}^{BCD}}$$

olur.

$$\hat{\zeta}_{(kk')(11'),ij}^{CD.AB} = \hat{\zeta}_{(kk')(11'),j}^{CD.B} = \log \left( \frac{n_{jk'l'}^{BCD} n_{jkl}^{BCD}}{n_{jkl'}^{BCD} n_{jk'l}^{BCD}} \right),$$

$$\hat{q}_{(kk')(11'),ij}^{CD.AB} = \hat{q}_{(kk')(11'),j}^{CD.AB} = \frac{n_{jkl}^{BCD} n_{jk'l'}^{BCD}}{n_{jkl'}^{BCD} n_{jk'l}^{BCD}}$$

ve

$$S^2(\hat{\zeta}_{(kk')(11'),ij}^{CD.AB}) = S^2(\hat{\zeta}_{(kk')(11'),j}^{ACD.B})$$

$$= \frac{1}{n_{jkl}^{BCD}} + \frac{1}{n_{jk'l'}^{BCD}} + \frac{1}{n_{jkl'}^{BCD}} + \frac{1}{n_{jk'l}^{BCD}}$$

tahminleri verilir.

$C_h$ 'ın  $B_h$  ve  $D_h$  kısmi ilişkisi üzerine etkisini kontrol etmek için,

$$\begin{aligned}\hat{\zeta}_{(jj')(kk')(ll')}^{BCD} &= \hat{\zeta}_{(jj')(ll')}^{BD.C} \cdot k - \hat{\zeta}_{(jj')(ll')}^{BD.C} \cdot k' \\ &= \log \left( \frac{n_{jkl}^{BCD} \cdot n_{j'kl'}^{BCD} \cdot n_{jk'l'}^{BCD} \cdot n_{j'k'l'}^{BCD}}{n_{jkl'}^{BCD} \cdot n_{j'kl}^{BCD} \cdot n_{jk'l}^{BCD} \cdot n_{j'k'l'}^{BCD}} \right)\end{aligned}$$

üçlü-faktör interaksiyonları dikkate alınır. Burada,

$$\begin{aligned}S(\hat{\zeta}_{(jj')(kk')(ll')}^{BCD}) &= \left[ S^2(\hat{\zeta}_{(jj')(ll')}^{BD.C} \cdot k) \right. \\ &\quad \left. + S^2(\hat{\zeta}_{(jj')(ll')}^{BD.C} \cdot k') \right]^{1/2} \\ &= \left( \frac{1}{n_{jkl}^{BCD}} + \frac{1}{n_{j'kl'}^{BCD}} + \frac{1}{n_{jk'l}^{BCD}} + \frac{1}{n_{j'k'l'}^{BCD}} + \frac{1}{n_{jkl'}^{BCD}} + \frac{1}{n_{j'kl}^{BCD}} \right. \\ &\quad \left. + \frac{1}{n_{jk'l'}^{BCD}} + \frac{1}{n_{j'k'l'}^{BCD}} \right)\end{aligned}$$

ve standardize değerler

$$\frac{\hat{\zeta}_{(jj')(kk')(ll')}^{BCD}}{S(\hat{\zeta}_{(jj')(kk')(ll')}^{BCD})}$$

olur.

### 3.4. Tam Tablolar İçin Enbüyük-Olabilirlik Tahminleri

#### 3.4.1. Üç ve Dört Boyutlu Tablolarda Direkt Tahminler

Örnek olarak  $\lambda_{ij}^{AB} = \lambda_{ijk}^{ABC} = 0$  kabul edilen model için, direkt tahminler elde edelim. Basit hücreler için

model,

$$m_{ijk} = \exp(\lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}) \quad (3.4-1)$$

olur. AC, BC den oluşan model tanımlayan sınıfın  $n^{AC}$  ve  $n^{BC}$  elemanları

$$\begin{aligned} n^{AC} &= \exp(\lambda + \lambda_i^A + \lambda_k^C + \lambda_{ik}^{AC}) \sum_j \exp(\lambda_j^B + \lambda_{jk}^{BC}) \\ n^{BC} &= \exp(\lambda + \lambda_i^A + \lambda_k^C + \lambda_{jk}^{BC}) \sum_j \exp(\lambda_i^A + \lambda_{ik}^{AC}) \end{aligned} \quad (3.4-2)$$

nın tahminleridir ve  $n^C$  hücrelerini ortak olarak bulunduran C vektörüne sahiptirler, bunlar

$$n^C = \exp(\lambda + \lambda_k^C) \sum_{i,j} \exp(\lambda_i^A + \lambda_j^B + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}) \quad (3.4-3)$$

nın tahminleridir. Burada toplamı kısım direkt tahmin için gerekli olan unsurlarına ayrılabilir. O zaman,

$$\begin{aligned} n^C &= \exp(\lambda + \lambda_k^C) \sum_j \exp(\lambda_j^B + \lambda_{jk}^{BC}) \\ &\quad \sum_i \exp(\lambda_i^A + \lambda_{ik}^{AC}) \end{aligned} \quad (3.4-4)$$

olur.  $n^{AC} n^{BC}$  çarpımını  $n^C$  ile böldüğümüzde, toplanan terimler kısalır ve üstteki  $m_{ijk}$  eşitliğine denk bir eşitlik elde edilir. Dolayısıyla da

$$\hat{m}_{ijk} = \frac{n^{AC} n^{BC}}{n^C}$$

ilişkisi elde edilir. Böylece her basit hücre için kulla-

nılan modele göre tahminler direkt olarak elde edilebilir.

Aslında, 1. ve 2. değişkenle ilgili iki yönlü k tane tablodan herbiri alınır ve her tablo için hücre tahminlerini tablo için tekboyutlu toplamlardan iki-boyutlu tablolar için bilinen tarzda

$$\text{Beklenen Değer} = \frac{(\text{satır toplamı})(\text{sütun toplamı})}{\text{Genel tablo toplamı}}$$

elde ederiz. Üzerinde durulan model gibi bir tane ikili faktör interaksyonu ve çoklu-faktör interaksyonu bulundurmayan modellere benzer modeller genellikle üç ve daha çok boyutlu tablolarda ortaya çıkar ve hücre tahminleri ayrı yolla elde edilebilir. Çizelge 13' de, dört boyutlu tablolar için direkt tahminleri veren her tip model için  $m_{ijk}$ 'ların direkt hücre tahmin şekli verilmektedir. Bu ifadeler (3.4-1)-(3.4-4) eşitlikleri izlenerek elde edilebilir:

- (i) model tanımlayan sınıfın her elemanının hücreleri için ifade yazmak;
- (ii) çakışan model tanımlayan sınıf elemanı varsa onlar için ifade yazmak ve mümkünse çarpanlarına ayırmak;
- (iii) direkt tahminin bulunup bulunmadığını ifade etmek için, (i) ve (ii) deki ifadeleri kontrol etmek. Direkt tahminin genel şekli tahmin edilebilir. Pay model tanımlayan sınıfın her elemanından değerlere sahiptir. Payda çakışmaya neden olan gerekenden fazla elemanlardan değerlere sahiptir; N etkisindeki terimler doğru büyüklük seviyesini sağlamak için ya pay yada paydada gözükürler.

Çizelge 13- Dört Boyutlu Tablolarda Mümkün  
Direkt Tahmin Tipleri.

Tanıtım numarası <sup>a</sup>	Model tanımlayan sınıf	Direkt hücre tahminleri
4	ABC, ABD	$\frac{n^{ABC} n^{ABD}}{n^{AB}}$
7	ABC, AD	$\frac{n^{ABC} n^{AD}}{n^A}$
8	ABC, D	$\frac{n^{ABC} n^D}{N}$
12a	AB, AC, AD	$\frac{n^{AB} n^{AC} n^{AD}}{(n^A)^2}$
12b	AB, AC, BD	$\frac{n^{AB} n^{AC} n^{BD}}{n^A n^B}$
13a	AB, CD	$\frac{n^{AB} n^{CD}}{N}$
13b	AB, AC, D	$\frac{n^{AB} n^{AC} n^D}{n^A N}$
14	AB, C, D	$\frac{n^{AB} n^C n^D}{N^2}$
15	A, B, C, D	$\frac{n^A n^B n^C n^D}{N^3}$

<sup>a</sup>Çizelge 14 ' e bakınız.

Bazen model tanımlayan sınıf elemanlarının hiç bir

kombinasyonu direkt bir tahmin vermez.  $n^D$  için (3.4-4) ifadesinde 1. deęişken üzerinden toplanan terimler 2. deęişken üzerinden toplananlardan ayrılabilir ve bu ayrılabilirlik  $n^D$  nin toplanan terimlerini  $n^{AD}$  ve  $n^{CD}$  için toplanan terimleri ile çıkartmamızda yardımcı olur. Gereken-den fazla model tanımlayan sınıf elemanına sahip ve bunların ayrılamadığı modellerde bir  $m_{ijk}$  tahmini veren direkt bir fonksiyon bulamayız. Böylece  $\lambda_{ijk}^{ABC} = 0$  olan üç boyutlu modeller direkt tahminlere sahip deęildirler.

Çakışan iki model tanımlayan sınıf elemanını bir dięerine birleşmiş şekilde düşünebiliriz.  $\lambda_{ijk}^{ABC} = 0$  olduđu modelde, model tanımlayan sınıf AB, BC ve AC elemanlarından oluşur ve eđer her çakışan çift birleştirilirse, kapalı bir halka elde edilir. Her boyutta böyle kapalı halkalar bulunduğunda, model tanımlayan sınıf elemanlarının hepsi yada sadece bir kaçı halkada bulunsa iterasyon gerekli olmaktadır. Örneğin, beş boyutluda model tanımlayan sınıfı AB, BC, CD ve DE elemanlarından oluşan bir modeli direkt olarak uydurabiliriz. Bu dördüne iki boyutlu başka bir eleman eklersek iterative metodlar kullanmak zorunda kalınır. Bununla beraber, bu model altı tane ikili-faktör etkisine sahip olsada ABC ve CDE elemanlarından direkt tahminler elde edebiliriz. Çünkü model tanımlayan sınıf elemanları kendi kendilerine bir halka oluşturmazlar. Bir sonraki kısımda verilen kurallar böyle halkaları otomatik olarak ortaya çıkarmamıza olanak sağlar.

### 3.4.2. Direkt Tahminlerin Varlığını Tayin Etmek İçin Kurallar

Model tanımlayan sınıf elemanlarını kontrol etmek için aşağıdaki kurallar, bir modelin direkt olarak uydurulup uydurulamayacağını tayin etmede kullanılırlar. Her sa-

yıda boyuta uygulanabilirler. Sınıflama kuralları olarak kabul edilirler.

1. basamak: Daima birlikte gözüken değişken grubunu tek bir değişken olarak yeniden isimlendirmek.
2. basamak: Her gruplaşmada gözüken herhangi değişkeni çıkarmak.
3. basamak: Sadece bir gruplaşmada gözüken herhangi değişkeni çıkarmak.
4. basamak: Gerekenden fazla gruplaşmaları uzaklaştırmak.
5. basamak: 1-4 basamakları,
  - (i) bir kapalı-form tahmininin kullanışlı olduğunun göstergesi olarak ikiden fazla gruplaşma kalmayınca veya
  - (ii) hiç bir kapalı form tahmininin mevcut olmadığına göstergesi olarak, başka hiç bir basamak uygulanamayınca kadar tekrarlamak.

1. basamak daha sonraki basamaklara geçişi kolaylaştırmada bir aşamadır. 2. basamak, eğer  $s$  hücreli model tanımlayan sınıf elemanlarını birlikte sabitleştirmenin,  $s$  bağımsız dağılıma götürdüğü hatırlanırsa açıklık kazanır. Böylece, eğer model, tekrarlanan değişkenin bir tek kategorisi için uydurulabiliyorsa, her kategori için uygulanabilir ve bundan dolayı böyle değişkenleri dikkate almaz. 3. basamakta tek değişkenler uzaklaştırılır. Böylece çakışan elemanlarla oluşan alt gruplaşmalarda görülmezler ve direkt tahminler için gerek duyulan çarpanlarına ayırma işlemi bunlarda yapılamaz. 4. basamak gereklidir. Çünkü 1-3. basamaklar bazı gerekenden fazla elemanı uzaklaştırır. Örneğin çizelge 13 'deki 4 boyutlu  $12b$  modelinde orjinal grup AB, AC, BD dir. 3. basamağı uygulayarak AB, AC, B elde edilir. Şimdi B gerekenden fazladır. Bu nedenle 4. basamak uygulanarak AB, AC kendi haline kalır. Eğer 1-3. basa-



maklar arasındaki devreye devam edilseydi 2. basamağı uygulamak gerekecek ve B, C elde edilecekti ve 3. basamağı yeniden uygulama bu terimleri uzaklaştıracaktı. İki gruplaşma tamamen uzaklaştırılabildiğinden sayı ikiye indirildiğinde hiç bir ilave işleme gerek duyulmaz.

Eğer üçden daha az gruplaşma elde edemezsek, iterasyon yapmalıyız.  $\lambda_{ijk}^{ABC} = 0$  olduğu üç boyutlu model dik-kate alınır model tanımlayan sınıf AB, AC ve BC den oluşur. 1-4. basamaktan hiçbiri uygulanamaz, ve hücre tahminleri sadece iterative yöntemlerle elde edilebilir. Aynı içinden çıkılmaz durum ABC, BCD, AD model tanımlayan sınıfı ile de ortaya çıkar. Bu sınıf sadece 1. basamakta BC nin B' olarak yeniden isimlendirilmesiyle AB', B'D, AD olarak basitleştirilebilir. Bu iterasyon isteyen başka bir kapalı halkadır.

Çizelge 14- Dört Boyutlu Tablolar İçin Hiyerarşik Modeller.

Model	Çıkarılan etkilerin numarası <sup>a</sup>			Serbestlik derecesi <sup>b</sup>	Her boyutta model tanımlayan sınıf elemanları <sup>c</sup>			Yorumlar	İterative (I) veya direkt çarpımlı (m)
	4f	3f	2f		3D	2D	1D		
1	1	-	-	1	4	-	-	1	I
2	1	1	-	2	3	-	-	4	I
3	1	2	-	3	2	1	-	6	I
4	1	2	1	4	2	-	-	6	M
5	1	3	-	4	1	3	-	4	I
6	1	3	1	5	1	2	-	12	I
7	1	3	2	6	1	1	-	12	M
8	1	3	3	7	1	-	1	4	M
9	1	4	-	5	-	6	-	1	I
10	1	4	1	6	-	5	-	6	I
11	1	4	2	7	-	4	-	15	I

## Çizelge 14- Devam

12 a	1	4	3	8	-	3	-	4	M
b	-	-	-	-	-	3	-	12	M
c	-	-	-	-	-	3	1	4	I
13 a	1	4	4	9	-	2	-	3	M
b	-	-	-	-	-	2	1	12	M
14	1	4	5	10	-	1	2	6	M
15	1	4	6	11	-	-	4	1	M
Toplam									113

<sup>a</sup>4f, 3f, 2f, sırasıyla dört, üç ve ikili faktör etkilerini gösterir; 3D, 2D ve 1D, sırasıyla üç, iki ve tek boyutlu model tanımlayan sınıf elemanlarını gösterir.

<sup>b</sup>Serbestlik dereceleri her değişken iki şıklı varsayılarak hesaplanmıştır.

<sup>c</sup>Her değişkeni içermeyen model bulunmaz.

Çizelge 14 'deki dört boyutlu hiyerarşik modellerin sınıflanması çoğu model tipleri için, her yorumun ya direkt çarpım (M) yada iterative bir yöntem (I) ile elde edilen tahminler verdiğini gösterir. 12. tip model, istisna oluşturur. Burada sıfıra eşit olan terimler grubu dört-faktörlü, bütün üç-faktörlü ve üç tane iki-faktörlü  $\lambda$  terimleridir. Sadece bazı yorumlar Çizelge 13 'de gösterilen formda direkt tahminler verir; diğerleri, AB, AC, BC, D model tanımlayan sınıf elemanlarına sahip bir yorum gibi iterative hesaplamalar gerektirir.

Çok boyutlu tabloların analiziyle ilgili parametre tahminleri genellikle sayısal yöntemler kullanılarak elde edilir. Bu nedenle model uydurma ve model seçmede bilgisayardan yararlanmak gerekir. Bu amaçla ekte verilen bilgisayar programı kullanılabilir.

Ele alınan uygulamalardan ilkinden Çizelge 15 'de verilen Ç.Ü.Z.F. Araştırma Çiftliğinde son üç yılda doğan

kuzuların ana yaşı, doğum şekli ve cinsiyet'e göre karşılıklı sınıflanmalarıyla elde edilen  $3 \times 2 \times 2$  boyutlu bir olabilirlik tablosu, ikincisinde ise Çizelge 17 'de verilen ve Urfa Ceylanpınar Tarım İşletmesinde doğan buzağuların ana ırkı, doğum mevsimi ve cinsiyete göre karşılıklı sınıflanmasıyla elde edilen  $4 \times 4 \times 2$  boyutlu bir olabilirlik tablosu analiz edilmektedir. Her iki tablo analizinde de amaç tablonun boyutlarını oluşturan değişkenlerin bağımlılık yapısının ortaya konulması ve verilerin durumunu en iyi izah eden log-linear modelin tesbit edilmesidir.



## 4. ARAŞTIRMA BULGULARI

## Uygulama 1.

Çukurova Üniversitesi Ziraat Fakültesi Araştırma Çiftliğinde 1982,1983,1984 yılında doğan kuzular ana yaşı, doğum şekli ve cinsiyetlerine göre karşılıklı sınıflanarak çizelge 15'deki şekilde 3x2x2 boyutlu bir olabilirlik tablosu oluşturulmuştur. Amacımız, dikkate alınan bu özelliklerin bağımlılık yapısını araştırmak ve bu yapıyı en iyi ortaya koyan log-linear modeli saptamaktadır.

Çizelge 15- Ç.Ü.Z.F. Araştırma Çiftliğinde 3 Yılda Doğan Kuzuların Ana Yaşı (2, 3, 4), Cinsiyet (Erkek, Dişi), Doğum şekli (Tek, İkiz) ne Göre Karşılıklı Sınıflanmasıyla Oluşturulan Olabilirlik Tablosu.

(A) Cinsiyet	(B) Doğum şekli	(C) Ana yaşı			Toplam
		2	3	4	
Erkek	Tek	75	81	30	186
	İkiz	39	48	21	108
Dişi	Tek	52	63	50	165
	İkiz	35	37	29	101
Toplam		201	229	130	560

Verilere en uygun modeli seçmek amacıyla sadece ana etkileri bulunduran modelden başlamak üzere, interaksyon terimleri birer birer modele dahil edilerek oluşturulan 5 model, verilere uydurularak  $L^2$ 'de meydana gelen değişimler incelenmiştir. Her model için elde edilen  $x^2$  ve

$L^2$  değerleri çizelge 16'da özetlenmektedir.

Çizelge 16- Çizelge 15'deki Veriler İçin Uydurulan Hiyerarşik Modeller ve Bunlara Ait Khi-Kare İstatistikleri.

Model	Model Tanımlayan sınıf	$x^2$	$L^2$	SD	$L^2$ 'deki indirgenme	SD'ki indirgenme	F
1	A,B,C	13.019	13.124	7			1.87
2	B,AC	1.128	1.126	5	11.99 <sup>a</sup>	2	0.23
3	AC,AB	1.039	1.035	4	0.09 <sup>b</sup>	1	0.25
4	AC,BC	1.030	1.028	3	0.10 <sup>c</sup>	2	0.34
5	AB,AC,BC	0.963	0.961	2	0.07,0.067 <sup>d</sup>	2,1	0.48

$$a_{L_1^2-L_2^2}; \quad b_{L_2^2-L_3^2}; \quad c_{L_2^2-L_4^2}$$

$$d_{L_3^2-L_5^2}; \quad L_4^2-L_5^2$$

Doymuş modele ait parametre tahminleri ve standart hatalar sırasıyla şöyledir:

<u>A</u>	<u>B</u>	<u>C1</u>	<u>C2</u>	<u>AB</u>	
0.01270	0.25056	0.10254	0.23556	0.00508	
0.04551	0.04551	0.06228	0.06047	0.04551	
<u>AC1</u>	<u>AC2</u>	<u>BC1</u>	<u>BC2</u>	<u>ABC1</u>	<u>ABC2</u>
0.10591	0.11519	0.01190	0.01331	0.05942	-0.0723
0.06228	0.06047	0.06228	0.06047	0.06228	0.06047

Çok boyutlu tablo analizinin model belirlemeden sonraki aşaması, seçilen model üzerinde yorum yapabilmek için modelin verilere uydurularak parametre ve beklenen değerlerin tahminlerinin elde edilmesini gerektirir. Üzerinde durulan yani verilerimize en iyi uyum sağlayan modele ait para-

metre tahminleri ve standart hatalar şöyle bulunmuştur:

<u>A</u>	<u>B</u>	<u>C1</u>	<u>C2</u>	<u>AC1</u>	<u>AC2</u>
0.01455	0.25923	0.10633	0.23776	0.12059	0.11277
0.04415	0.04369	0.06034	0.05858	0.06034	0.05858

Bu model için gözlenen ve beklenen değerler ile düzeltilmiş rezidüaller ise şöyledir:

<u>Hücre</u>	<u>Gözlenen değer</u>	<u>Beklenen değer</u>	<u>Düzeltilmiş rezidüyal</u>
111	75.000	71.454	0.770
211	52.000	54.530	-0.610
121	39.000	42.546	-0.770
221	35.000	32.470	0.610
111	81.000	80.855	0.030
211	63.000	62.679	0.073
122	48.000	48.145	-0.030
222	37.000	37.321	-0.073
112	30.000	31.966	-0.597
212	50.000	45.516	-0.121
122	21.000	19.034	0.597
222	29.000	29.484	-0.121

## Uygulama 2

Urfa Ceylanpınar Tarım İşletmesinde doğan buzağılar, doğum mevsimi (ilkbahar, yaz, sonbahar, kış), ana ırkı (Alman, Danimarka, Hollanda, Çiftlik) ve cinsiyete (erkek, dişi) göre sınıflanarak üç yönlü 4x4x2 boyutlu bir olabilirlik tablosu elde edilmiştir. Amaç tablonun boyutlarını oluşturan değişkenler arasındaki bağımlılık yapısını ortaya koymak ve bu yapıyı en iyi izah eden log-linear modeli saptamaktadır.

Çizelge 17 - Ceylanpınar Tarım İşletmesinde Doğan Buzagıların Doğum Mevsimi, Ana Irkı, Cinsiyete Göre Karşılıklı Sınıflanmasıyla Elde Edilen Olabilirlik Tablosu.

(A) Doğum mevsimi	(B) Ana ırkı	(C) Cinsiyet		Toplam
		Erkek	Dişi	
İlkbahar	Alman	126	104	230
	Danimarka	30	23	53
	Hollanda	35	24	59
	Çiftlik	311	253	564
Yaz	Alman	112	108	220
	Danimarka	52	43	95
	Hollanda	27	23	50
	Çiftlik	509	467	976
Sonbahar	Alman	100	103	203
	Danimarka	79	68	147
	Hollanda	21	18	39
	Çiftlik	393	355	748
Kış	Alman	110	92	202
	Danimarka	56	49	105
	Hollanda	33	35	68
	Çiftlik	196	217	413
Toplam		2190	1982	4172

Bu tablo için seçilen model grubu verilere uydurularak elde edilen  $x^2$  ve  $L^2$  değerleri, Çizelge 18'de özetlenmektedir.

Çizelge 18 - Çizelge 17 'deki Veriler İçin Uyduru-  
lan Hiyerarşik Modeller Ve Bunlara  
Ait Khi-Kare İstatistikleri.

Model	Model tanım- layan sınıf	$x^2$	$L^2$	$L^2$ 'deki SD indir- genme	SD'de- ki in- dirgen- me	F
1	A,B,C	153.897	153.825	24		6.409
2	C,AB	10.285	10.297	15	143.53 <sup>a</sup>	9 0.686
3	AB,AC	5.263	5.268	12	5.03 <sup>b</sup>	3 0.439
4	AB,BC	9.676	9.690	12	0.607 <sup>c</sup>	3 0.807
5	AB,AC,BC	4.347	4.352	9	0.915,5.338	3,3 0.483

$$a_{L_1^2 - L_2^2}; \quad b_{L_2^2 - L_3^2}; \quad c_{L_2^2 - L_4^2};$$

$$d_{L_3^2 - L_5^2}; \quad L_4^2 - L_5^2$$

Çizelge 17'deki veriler için uydurulan doymuş modele

$$\log m_{ijk} = \mu + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

ait parametre tahminleri ve standart hatalar şöyledir:

<u>A1</u>	<u>A2</u>	<u>A3</u>	<u>B1</u>	<u>B2</u>
-0.14260	0.12641	0.06623	0.27234	-0.55118
0.04514	0.05105	0.04445	0.03440	0.04524
<u>B3</u>	<u>C</u>	<u>A1B1</u>	<u>A1B2</u>	<u>A1B3</u>
-1.09506	0.05591	0.21500	-0.43346	0.20876
0.05609	0.02439	0.06069	0.09004	0.09526
<u>A2B1</u>	<u>A2B2</u>	<u>A2B3</u>	<u>A3B1</u>	<u>A3B2</u>
-0.09403	-0.11459	-0.08487	-0.07176	0.38384
0.06272	0.07590	0.11320	0.07206	0.07169



<u>A3B3</u>	<u>A1C</u>	<u>A2C</u>	<u>A3C</u>	<u>B1C</u>
-0.39932	0.07425	-0.02840	-0.01950	-0.00874
0.10751	0.04418	0.04329	0.03999	0.03443
<u>B2C</u>	<u>B3C</u>	<u>A1B1C</u>	<u>A1B2C</u>	<u>A1B3C</u>
0.03649	-0.00840	-0.02548	-0.03380	0.06688
0.04518	0.05588	0.06030	0.09042	0.09436
<u>A2B1C</u>	<u>A2B2C</u>	<u>A2B3C</u>	<u>A3B1C</u>	<u>A3B2C</u>
-0.00059	0.03102	-0.06535	-0.08491	0.00207
0.05990	0.07669	0.10126	0.09653	0.07337
<u>A3B3C</u>				
0.04906				
0.11142				

Verilere uydurulan modeller içerisinde durulan modele ait parametre tahminleri ve standart hatalar sırasıyla

<u>A1</u>	<u>A2</u>	<u>A3</u>	<u>B1</u>	<u>B2</u>
-0.13262	0.11188	0.05805	0.27241	-0.54854
0.04401	0.04667	0.04282	0.03423	0.04491
<u>B3</u>	<u>C</u>	<u>A1B1</u>	<u>A1B2</u>	<u>A1B3</u>
1.09796	0.04859	0.20864	-0.43396	0.22215
0.05544	0.01552	0.05997	0.08911	0.09336
<u>A2B1</u>	<u>A2B2</u>	<u>A2B3</u>	<u>A3B1</u>	<u>A3B2</u>
-0.08560	-0.10971	-0.11035	-0.10674	0.39198
0.06123	0.07581	0.10733	0.06027	0.07025
<u>A3B3</u>				
-0.37996				
0.10443				

şeklinde bulunmuştur. Gözlenen ve beklenen değerler ile düzeltilmiş rezidüaller ise şöyledir:

<u>Hücre</u>	<u>Gözlenen değer</u>	<u>Beklenen değer</u>	<u>Düzeltilmiş rezidüyal</u>
111	126.000	120.738	0.715
211	112.000	114.879	-0.397
311	100.000	106.582	-0.948
411	110.000	106.058	0.569
121	30.000	27.941	0.568
221	52.000	49.345	0.546
321	79.000	77.223	0.299
421	56.000	55.203	0.157
131	35.000	31.086	1.024
231	27.000	28.467	-0.607
331	21.000	20.601	0.128
431	33.000	35.805	-0.684
141	311.000	295.846	1.375
241	509.000	511.233	-0.163
341	393.000	392.313	0.056
441	196.000	216.680	-2.148
112	104.000	109.556	-0.753
212	108.000	104.240	0.521
312	103.000	96.711	0.904
412	92.000	96.236	-0.611
122	23.000	25.353	-0.648
222	43.000	44.775	-0.368
322	68.000	70.071	-0.347
422	49.000	50.090	-0.215
132	24.000	28.207	-1.099
232	23.000	20.652	0.607
332	18.000	18.693	-0.222
432	35.000	32.489	0.612
142	253.000	268.447	-1.399
242	467.000	463.886	0.229
342	355.000	355.980	-0.079
442	217.000	196.613	2.114

## 5. TARTIŞMA VE SONUÇLAR

Üç boyutlu bir tabloya ilişkin doymuş model;

$$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

şeklindedir. Parametre tahminlerinin yapılabilmesi için ana etki ve interaksiyonlarla ilgili uygun kısıtlamaların konması gerekir.

Eğer hiyerarşik modeller dikkate alınırsa, verilere uyabilen en basit indirgenmiş model, üçlü interaksiyonu bulunmayan

$\log m_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$  modelidir.  $\lambda_{ij}^{AB}$ ,  $\lambda_{ik}^{AC}$  ve  $\lambda_{jk}^{BC}$  ler sıfır kabul edilmediğinden hiç bir hiyerarşik model daha alt derecedeki  $\lambda$ ,  $\lambda_i^A$ ,  $\lambda_j^B$  veya  $\lambda_k^C$  terimlerini sıfır kabul edememektedir.

Hiyerarşik modellerle ilgili hesaplamaların sağladığı faydalar yanında hiyerarşik modellerle sınırlamayla elde edilen esas kazanç yorumunun kolaylığıdır. Çünkü hiyerarşik modelde  $\lambda_i^A$  terimi sıfır olduğunda, A'yı ihtiva eden diğer  $\lambda$ -terimleri de sıfır olacaktır. Örnek olarak

$$\lambda_1^A = \lambda_2^A = 0 \text{ varsayalım. } \tau_{12 \cdot jk}^{A \cdot BC} = \log(m_{1jk}/m_{2jk}) = \mu_{1jk}$$

$-\mu_{2jk}$  ifadesi C değişkeninin k. kategorisinde ve B değişkeninin j. kategorisindeki bir bireyin A değişkeninin 2. kategorisinden ziyade 1. kategorisinden olmasının logaritmik

risklerini göstermektedir.

$$\begin{aligned}\lambda_1^A - \lambda_2^A &= 2\lambda_1^A = \frac{1}{4} (m_{111} - m_{211} + m_{121} - m_{221} \\ &+ m_{112} - m_{212} + m_{122} - m_{222}) \\ &= \frac{1}{4} (\tau_{12.11}^{A.BC} + \tau_{12.21}^{A.BC} + \tau_{12.12}^{A.BC} \\ &+ \tau_{12.22}^{A.BC})\end{aligned}$$

olduğundan,  $\lambda_1^A = \lambda_2^A = 0$  olması,  $\tau_{12.ij}^{A.BC}$  logaritmik risklerinin,  $1 \leq j \leq 2$ ,  $1 \leq k \leq 2$ , ortalamalarının sıfır olduğunu ifade eder.  $\tau_{12.jk}^{A.BC}$  'ların hepsi sıfır olmadıkça ortalamasının sıfır olması durumu fazla birşey ifade etmez. Tüm  $\tau_{12.jk}^{A.BC}$  'ların eşit olması ancak,

$$\begin{aligned}8 \lambda_{11}^{AB} &= \tau_{(12)(12).1}^{AB.C} + \tau_{(12)(12).2}^{AB.C} \\ &= \tau_{12.11}^{A.BC} - \tau_{12.21}^{A.BC} + \tau_{12.12}^{A.BC} - \tau_{12.22}^{A.BC}\end{aligned}$$

şartının sağlanması ile olur. Böylece  $\lambda_i^A$  'ların sıfır olması hipotezinin basit bir yorumu,  $\lambda_{ij}^{AB}$  'ler sıfır ise mümkündür.  $\tau_{12.jk}^{A.BC}$  'ların eşitliği ancak,

$$\lambda_{11}^{AC} = \tau_{12.11}^{A.BC} + \tau_{12.21}^{A.BC} - \tau_{12.12}^{A.BC} - \tau_{12.22}^{A.BC} = 0$$

ve

$$\lambda_{111}^{ABC} = \tau_{12.11}^{A.BC} - \tau_{12.21}^{A.BC} - \tau_{12.12}^{A.BC} + \tau_{12.22}^{A.BC} = 0$$

olmasını gerektirir. Böylece  $\lambda_i^A$  'nın sıfır kabul edildiği

fakat  $\lambda_{ij}^{AB}$ ,  $\lambda_{ik}^{AC}$  veya  $\lambda_{ijk}^{ABC}$ 'lerin sıfır kabul edilmediği bir modeli yorumlamak zordur. Dolayısıyla da hiyerarşik modeller tercih edilmektedir.

Çizelge 15'deki veriler için seçilen model grubu Çizelge 16'dan izlenebilir. 1. model sadece ana etkileri bulundurur ve değişkenlerin karşılıklı bağımsızlığını ifade eder. 2. modelde değişkenlerden ikisi birbirlerine bağımlı iken üçüncünün bağımsızlığı (kısmi bağımsızlık) söz konusudur. A ile B birbirine bağımlı iken B bunlardan bağımsızdır. 3. modelde ise değişkenlerden ikisi üçüncünün her seviyesinde birbirlerinden bağımsızdır (şartlı bağımsızlık). Burada AC ve AB interaksiyonları mevcutken BC interaksiyonu söz konusu değildir. 4. modelde şartlı bağımsızlık bu sefer A ile B arasındadır. 5. model, üçlü-faktör interaksiyonu olmayan model olup tüm ikili-faktör interaksiyonlarını bulundurmaktadır.

Bu modellerin herbirine ait  $L^2$ 'ler incelenerek  $L^2$  de meydana gelen değişmeler araştırılmış, verilere en iyi uyan modeli saptamada modellere ait  $L^2$  farklarından yararlanılmıştır.  $L_1^2$  ve  $L_2^2$  arasındaki fark, 11.99, 2 serbestlik derecesine göre  $L^2$ 'de oldukça önemli bir değişim olduğunu ifade etmektedir. Bu fark ile AC interaksiyonunun önemli olduğu ortaya çıkar.  $L_2^2-L_3^2$  farkı, 1 serbestlik derecesine göre önemsiz olup AB'nin modele ilavesinin  $L^2$ 'de önemli sayılacak bir değişme sağlamadığını gösterir.  $L_2^2-L_4^2$  farkı ise 2 serbestlik derecesine göre önemsizdir. Böylece BC interaksiyonunun modele ilavesi de  $L^2$  de önemli değişme yaratmamıştır.  $L_3^2-L_5^2$  farkı, yine BC'nin önemini test etmede,  $L_4^2-L_5^2$  farkı, AB interaksiyonunun önemini test etmede kullanılır. Testler neticesi BC ve AB interaksiyonlarının modelde yer almaması gerektiği ortaya çıkar. Böylece verilere en uygun model olarak karşımıza çıkan mo-

del, 2. modeldir. Bu model,

$$\log \hat{m}_{ijk} = \hat{\mu} + \hat{\mu}_i^A + \hat{\mu}_j^B + \hat{\mu}_k^C + \hat{\mu}_{ik}^{AC}$$

şeklindedir ve verimizde bağımlılık yapısını en iyi izah eden modeldir. Bu model, cinsiyet (A) ile ana yaşı (C) birbirine bağımlı iken doğum şeklinin (B) kuzunun cinsiyeti ve ana yaşından bağımsız olduğunu ifade etmektedir.

F değerlerine göre bir değerlendirme yaparak da aynı sonuca varılabilir. 2, 3, 4 ve 5. modellere ait F değerleri 1'den küçüktür. Bu modellere ait  $L^2$ 'ler arasındaki farkların hepsi önemsizdir. O halde en az parametreye sahip 2. modelin geçerli olmaması için hiç bir neden yoktur.

Üzerinde durulan modele ait khi-kare değerleri çok küçük olduğundan rezidüyallere bakılarak bağımlılığın hangi hücrelerden ileri geldiğine dair kesin bir yorum yapmak zordur. Ancak,  $\chi^2$  yi yükselten hücre değerlerinin A=1 ve C=1 'deki yüksek frekans ve A=1 ve C=3'deki düşük frekanstan geldiği görülebilir.

Çok boyutlu tablolarla ilgili modellerin sayısı tablonun boyutlarına bağlı olarak artar. Bu durumda, hangi modellerin iyi uyum sağladığının belirlenmesi gerekir. Ancak çok sayıda parametre içeren modellerin genellikle daha iyi uyum sağlaması, basit ve açıklaması kolay modellerin ortaya çıkartılmasını güçleştirmektedir. Bu konuda tüm parametreleri bulunduran doymuş modelin uydurulmasıyla elde edilen tahminlerin öncelikle incelenmesi faydalı olur. Tahminlerin standardize değerleri hangi parametrelerin gereksiz olabileceği hakkında fikir verir. Model uydurulduktan sonra, interaksiyon parametresi için elde edilen tahmin değerlerinin, standart hataları ile ele alı-

narak bağımlılığın hangi hücrelerden ortaya çıktığı araştırılabilir.

Verilerimize uydurulan doymuş model için bulunan parametre tahminleri ve standart hataları kullanılarak hesaplanan standardize değerler yardımıyla AC interaksyonunun mutlaka modelde bulunması gerektiği anlaşılabilir. AC1 ve AC2 interaksyonları için bulunan standardize değerler sırasıyla 1.7 ve 1.9 dur ve diğer interaksyonlar için bulunanlardan daha büyüktürler.

Çizelge 17'deki verilere uydurulan model grubu Çizelge 18'de yer almaktadır. Model grubunu oluşturan modellerden herbiri için hesaplanan  $x^2$  ve  $L^2$  değerleri yine aynı çizelgede özetlenmiştir.

1. model sadece ana etkileri bulundurur ve değişkenlerin karşılıklı bağımsızlığını ifade eder. 2. modelde değişkenlerden ikisi birbirlerine bağımlı iken üçüncünün bağımsızlığı (kısmi bağımsızlık) söz konusudur. Burada A ve B değişkenleri birbirine bağımlı iken C, A ve B den bağımsızdır. 3. modelde ise değişkenlerden ikisi üçüncünün her seviyesinde birbirlerinden bağımsızdır (şartlı bağımsızlık). Burada AC ve AB interaksyonları mevcutken BC interaksyonu söz konusu değildir. 4. modelde şartlı bağımsızlık bu sefer A ile C arasındadır. 5. model, üçlü-faktör interaksyonu olmayan modeldir ve tüm ikili-faktör interaksyonlarını bulundurmaktadır.

Bu modellere ait  $L^2$  farklarından yararlanılarak verilere en uygun model saptanmaya çalışılmıştır.  $L_1^2-L_2^2$  arasındaki fark, 143.53, 9 serbestlik derecesine göre  $L^2$  de oldukça önemli bir değişim olduğunu göstermektedir. Bu farkın büyüklüğü AB interaksyonunun önemli olduğunu gösterir.  $L_2^2-L_3^2$  ve  $L_4^2-L_5^2$  farkı AC'nin,  $L_2^2-L_4^2$  ve  $L_3^2-L_5^2$  farkı BC'nin önemini test etmede kullanılır. Hep-



side kendi serbestlik derecelerine göre önemsiz sayılan değişimlerdir. Böylece verilere en uygun model 2. modeldir. Model,

$$\log \hat{m}_{ijk} = \hat{\mu} + \hat{\mu}_i^A + \hat{\mu}_j^B + \hat{\mu}_k^C + \hat{\mu}_{ij}^{AB}$$

dir. Bu model, doğum mevsimi (A) ile ana ırkı (B) birbirine bağımlı iken cinsiyetin (C) bunlardan bağımsız olduğunu ifade eder. Kısmi bağımsızlık söz konusudur.

F değerlerine göre değerlendirme yapılarak yine aynı modelin geçerliliği görülebilir. 2, 3, 4 ve 5. modellere ait F değerleri 1'den küçüktür. Bu modellere ait  $L^2$ '-ler arasındaki farkların hepsi de önemsiz olduğundan daha az parametre içeren 2. modelin tercih edilmesi uygun olur.

Üzerinde durulan modele ait rezidüyallerin incelenmesiyle iki rezidüyalın,  $r_{441}$  ve  $r_{442}$ , % 5 önem seviyesinde 1.96 'lık Z değerinden büyük olduğu görülür. Buna göre, A ve B değişkenleri arasındaki bağımlılıkta en büyük role sahip olan hücreler bunlardır.

Verilerimize uydurulan doymuş modele ait parametre tahminleri ve standart hatalar kullanılarak standardize değerler hesaplandığında, AB interaksiyonlarına ait bazı değerlerin % 5 önem seviyesinde 1.96'dan büyük oldukları görülmüştür. Diğer interaksiyonlara ait standardize değerler ise daha küçüktür. Böylece modelde ana etkiler dışında interaksiyon etkisi olarak sadece AB ikili-interaksiyon teriminin bulunması gerektiği baştan, tüm modeller uydurulmadan tahmin edilebilir. Ancak model uydurularak, bağımlılık yapısını ortaya koyan parametre tahminlerinin mutlaka elde edilmesi gerekmektedir.

Uygulamalardan da görüleceği gibi, birden fazla modelin verilere iyi uyum sağlaması olağandır. Bu durumda da-



ha az parametre içeren model tercih edilmelidir. Uyum sağ-layan modellerin eşit parametre içermesi durumunda, hatası en az olan model verilere uygun model olarak seçilir.

Beklenen değerlerin tahminlenmesi için genellikle poisson veya multinom dağılışı varsayımları altında enbüyük-olabilirlik tahminleme yöntemi kullanılır. Yöntemin kullanılması bir çok model için dönüşümlü çözümlere gerektirdiğinden çok boyutlu tablo analizinde bilgisayardan yararlanmak zorunludur. Bu nedenle verilere uydurduğumuz modeller için  $x^2$ ,  $L^2$  değerleri, parametre ve beklenen değer tahminleri, düzeltilmiş rezidüyllerin hesaplanmasında ekte sunulan bilgisayar programından yararlanılmıştır.



## ÖZET

Kalitatif verilerin analizinde esas amaç çok sıklı deęişkenlerin baęımlılık yapılarının kontrol edilmesidir.

İki çok sıklı deęişkenin baęımsızlık modeli bir log-linear modele eşdeęerdir. Bu log-linear model için enbüyük-olabilirlik tahminleri kesin ifadelere sahiptir. Bu nedenle khi-kare testleri ve düzeltilmiş rezidüyaller baęımsızlık modeline kolaylıkla uygulanır. Düzeltilmiş rezidüyallerin kullanımı, baęımsızlıktan ayrılış kaynaklarının, bilinen pearson khi-kare baęımsızlık testinden daha uygun bir analizini sağlar. Baęımsızlıktan ayrılışların büyüklüğünün kantitatif olarak elde edilmesi, en azından mevcut baęımsızlıktan ayrılışların tesbiti kadar önemlidir. Bu tür ilişkiler, normal olarak logaritmik çapraz-çarpım oranları olarak bilinen logaritmik hücre ortalamalarının linear kombinasyonlarının kontrolüyle yapılır.

Logaritmik çapraz-çarpım oranı, ikiye-ikilik olabilirlik tablosunun hücre ortalamalarının linear bir fonksiyonu olduğundan ve iki tane iki sıklı deęişkenin baęımsızlığı, logaritmik çapraz-çarpım oranının sıfır olmasına eşdeęer olduğundan, baęımsızlık modeli, toplanabilir log-linear modele eşdeęerdir. Bu tip tablolar için enbüyük-olabilirlik tahminleri ve khi-kare istatistikleri kolayca hesaplanabilir. Burada, khi-kare testleri sadece bir serbestlik dereceli olduklarından baęımsızlık testi basittir. Çünkü, sadece bir logaritmik çapraz-çarpım oranını dikkate almak gerekmektedir. Burada tüm rezidüyaller  $x^2$ 'ye eşit olduğundan rezidüyal analizi gereksizdir. Bu durum sadece ikiye-ikilik tablolarda ortaya çıkar. Bu tablolarda süreklilik düzeltilmesi yapılması gerekir.

Bağımsızlık, biri iki şıklı diğeri çok şıklı değişkenin oluşturduğu tabloda test edildiğinde, khi-kare testi çarpımsal serbestlik derecesine sahiptir. Burada, rezidüyal analizi önem kazanır ve süreklilik düzeltmesi yapılmaz.

İki çok şıklı değişkenin karşılıklı sınıflanmasında daha çok düzeltilmiş rezidüyal ve çapraz-çarpım oranları dikkate alınmalıdır. Bunun neticesi olarak, mevcut ilişkileri yorumlamada daha çok problemle karşılaşılabilir. Homogenlik hipotezi, bağımsızlık hipotezinden farklı bir olasılık modelinden ortaya çıkmasına rağmen her ikisinde aynı toplanabilir modele eşdeğerdir. Sonuç olarak enbüyük-olabilirlik tahminleri, khi-kare testleri, düzeltilmiş rezidüyal homogenlik modeli için bağımsızlık modelindekilerle aynıdır. Doymuş model, tablonun sahip olduğu hücre sayısı kadar bağımsız parametreye sahip modeldir. İnteraksiyon etkileri ve buna tekabül eden standart sapmaların karşılaştırılması verilerle uyumlu modeli tayin etmede değerli bir araçtır. Bunun için interaksiyon etkilerinin sıfırdan farklılığı test edilmektedir.

Çok boyutlu tabloların analizi özellikle boyut sayısının üç ve daha fazla olduğu durumlarda önemli güçlükler ortaya çıkarır. Kesikli verilerde, log-linear modellerin bu amaçla kullanılır hale gelmesi regresyon analizinin olanaklarının da kullanılmasını sağlamıştır. Log-linear modeller genel olarak k-boyutlu tablolara uydurulabilirler. Yöntemler ve yorumlar temelde aynıdır. Modeli oluşturan interaksiyonların sayısı ve dereceleri de bağımlılık yapısının belirlenmesinde yararlı olmaktadır. Çok boyutlu tabloların bu yöntemle değerlendirilmesi alt tabloların ayrı ayrı analizinde ortaya çıkabilecek hataları da önler.

Log-linear modellerin uygulanmasıyla ortaya çıka-

bilecek en önemli sorunlar model seçimiyle ilgilidir. Üç boyutlu tablolarda hiyerarşik modeller cazip bir model ailesi oluşturur. Bu sistem, daha çok boyutlu tablolarada kolaylıkla genelleştirilebilmektedir. Boyut sayısı arttıkça incelenmesi gerekli modellerin sayısında artmaktadır. En uygun model seçimi konusu, sistematik olarak bütün model gruplarının incelenmesini gerektirir. Üçden fazla boyutlu tablolarda bu tür sistematik yaklaşım fazla yararlı olmamaktadır. Bu konuda, değişkenler hakkındaki ön bilgileri kullanarak bazı modelleri baştan elemek, uğraşılacak model sayısını azaltmak yararlı olabilir. Modelin uyum mükemmelliğini test etmek amacıyla  $F$ ,  $x^2$ ,  $L^2$  istatistiklerinden birisi kullanılabilir.

Uygulamadaki bazı güçlükler rağmen log-linear modellerin, çok boyutlu tabloların analizinde kullanılması yeni bazı olanaklar getirdiğinden yararlı olmaktadır.

Çizelge 15'deki  $3 \times 2 \times 2$  boyutlu olabilirlik tablosunun boyutlarını oluşturan değişkenlerin bağımlılık yapısı araştırılarak bu yapıyı en iyi izah eden model olarak

$$\log \hat{m}_{ijk} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_j^B + \hat{\lambda}_k^C + \hat{\lambda}_{ik}^{AC}$$

bulunmuştur. Bu modele ait  $x^2$  ve  $L^2$  değerleri sırasıyla 1.128 ve 1.126'dır. Modele ait serbestlik derecesi 5'dir. Bu model, A (kuzunun cinsiyeti) ile C (Ana yaşı) değişkeni birbirine bağımlı iken B (doğum şekli) nin bunlardan bağımsız olduğunu ifade eder. Kısmi bağımsızlık söz konusudur.

Çizelge 17'deki  $4 \times 4 \times 2$  boyutlu tablonun bağımlılık yapısını en iyi izah eden model olarak da

$$\log \hat{m}_{ijk} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_j^B + \hat{\lambda}_k^C + \hat{\lambda}_{ij}^{AB}$$

bulunmuştur.  $\chi^2$  ve  $L^2$  istatistikleri sırasıyla 10.285 ve 10.297 dir. Bu modele ait serbestlik derecesi 15 olarak bulunmuştur. Model, A (doğum mevsimi) ile B (ana ırkı) değişkeni arasında bağımlılık söz konusu iken C (buzanın cinsiyeti) nin bunlardan bağımsız olduğunu ifade eder. Yine kısmi bağımsızlık söz konusudur.

THE ANALYSIS OF CATEGORICAL DATA IN THE ANIMAL SCIENCES  
BY LOG-LINEAR MODELS.

SUMMARY

A basic problem in the analysis of qualitative data is an examination of two polytomous variables to see if they are dependent.

The model of independence of two polytomous variables is equivalent to a log-linear model. Maximum-likelihood estimates for this log-linear model have explicit expressions, so that the chi-square tests and adjusted residuals are easily applied to the independence model. Use of adjusted residuals provides a more subtle analysis of sources of departure from independence than that is available from Pearson's familiar chi-square test of independence. Quantitative assessment of the size of departures from independence is at least as important as determination that departures from independence do exist. Such assessment is normally made through examination of the linear combinations of log-cell means known as log cross-product ratios.

Since the log cross-product ratio is a linear function of the logarithms of the cell means of a two-by-two contingency table and since independence of the two dichotomous variables is equivalent to a log cross-product ratio of zero, the hypothesis of independence can be shown to be equivalent to an additive log-linear model. For such tables, maximum-likelihood estimates and chi-square statistics are easily computed. Since both statistics here have only a single degree of freedom, testing independence is simplified; because only one log cross-product ratio need be considered. Since  $x^2 = r_{11}^2 = r_{12}^2 = r_{21}^2 = r_{22}^2$ , residual

analysis is trivial. This situation only arises in two-by-two tables. For such tables, continuity corrections are available to improve assessment of significance levels.

When independence is tested in a table with one dichotomous and one polytomous variable, chi-square tests have multiple degrees of freedom. Here, analysis of residuals is nontrivial, and continuity corrections disappear.

In the situation that two polytomous variables are cross-classified, many log cross-product ratios and many adjusted residuals must be considered. Consequently, problems of simultaneous inference arise. Although the homogeneity hypothesis arises from a different probability model than does the independence hypothesis, they both will be shown to be equivalent to the same additive model. Consequently, maximum-likelihood estimates, chi-square tests, and adjusted residuals for the homogeneity model are the same as those for the independence model. Descriptive models with as many independent parameters as the table has cells are called "saturated". Comparison of the interactions and corresponding estimated standard deviations provides a valuable tool in the determination of which hierarchical models can conceivably fit the data. For this, one obtains a test of whether interactions are zero.

The analysis of multidimensional tables, especially the analysis of the tables which have three or more dimensions, becomes increasingly complex. The use of the log-linear model with categorical data provides use of opportunities of regression analysis techniques. In general, log-linear models can be fit  $k$ -dimensional tables. The basic principals of the methods and interpretations of results are similar for small and large dimensional tables.

The number and degrees of interactions are help-

ful for the investigation of the dependence structure. The assessment of multidimensional tables by using this method protects from making mistakes rising due to employing small dimensional subtables.

The most important problems arising in using log-linear models are related with the selection of most subtle model. Hierarchical models are an attractive family of models for three-dimensional tables. This system can also easily be generalized for more dimensional tables. The number of models required to be examined increases as the dimensions of the tables increase. The issue of the selection of most adequate model requires the investigation of all possible model groups in a systematic way. This sort of systematic approach can not be very useful for the examination of contingency tables in which three or more variables are cross-classified. First of all it is need to eliminate some of the models by using some specific information to decrease the number of examined models. F,  $\chi^2$  and /or  $L^2$  statistics can be employed to test the goodness of fit of the models.

To use log-linear models will be very convenient although some difficulties may arise in applications of the models for three or more dimensional tables.

The most subtle model was found by investigating the dependence structure of the variables which constitute the dimensions of the 3x2x2 contingency tables given in Table 15. The model is

$$\log \hat{m}_{ijk} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_j^B + \hat{\lambda}_k^C + \hat{\lambda}_{ik}^{AC}$$

The calculated  $\chi^2$  and  $L^2$  values for this model are respectively 1.128 and 1.126 with 5 degrees of freedom. This model implies the partial independence of B (parity) from



A (sex of lamb) and C (age of dam). That is to say the partial association between A and C is significant.

The most subtle model for data given in Table 17 which is  $4 \times 4 \times 2$  dimensional table involving three variables is

$$\log \hat{m}_{ijk} = \hat{\lambda} + \hat{\lambda}_i^A + \hat{\lambda}_j^B + \hat{\lambda}_k^C + \hat{\lambda}_{ij}^{AB}$$

The related  $\chi^2$  and  $L^2$  statistics are respectively 10.285 and 10.297 with 15 degrees of freedom. The partial independence model also holds with this data set. The variable C (sex of calf) is independent of A (birth of season) and B (breed of dam) while A and B are mutually dependent.

## KAYNAKLAR

- BHAPKAR, V.P., and KOCH, G.G., 1968a. On the hypotheses of 'no interaction' in contingency tables. *Biometrics* 24, 567-594.
- BHAPKAR, V.P., and KOCH, G.G., 1968b. Hypotheses of no interaction in multidimensional contingency tables. *Technometrics* 10, 107-122.
- BIRCH, M.W., 1963. Maximum likelihood in three-way contingency tables. *J.Roy.Statist. Soc.Ser. B.*25, 220-233.
- BISHOP, Y.M.M., 1969. Full contingency tables, logits, and split contingency tables. *Biometrics* 25, 383-399.
- BISHOP, Y.M.M., FIENBERG, S.E., and HOLLAND, P.W., 1975. *Discrete multivariate analysis: theory and practice.* MIT Press, Cambridge, Massachusetts. (530)S.
- BISHOP, Y.M.M., and MOSTELLER, F., 1969. J.P. Bunher, W.H. Forest, Jr., F. Mosteller, and L.D. Vandam (Eds.), *The national halothane study, Chapter IV-3.* U.S. Govt. Printing Office, Washington, D.C. (735)S.
- COCHRAN, W.G., 1954. Some methods for strengthening the common  $\chi^2$  tests. *Biometrics* 10, 417-451.
- GIANOLA, D., 1979. Analysis of discrete variables in animal breeding context. *J. Dairy Sci.* 62 (9), 1471-1478.
- DRAPER, N.R., and SMITH, H., 1966. *Applied regression analysis.* Wiley, New York (417)S.
- FALGON, C., 1979. *Univariate and multivariate models for the analysis of qualitative dependent variables.* Ottawa, Canada; Policy, Planning and Economics Branch, Agriculture Canada.(58)S.
- FIENBERG, S.E., MEYER, M.M., 1983. Log-linear models and catagorical data analysis with psychometric and econometric applications. *J.Econometrics* 22 (1-2), 191-214.
- GART, J.J., and ZWEIFEL, J.R., 1967. On the bias of various estimators of the logit an its variance with application to quantal bioassay *Biometrika* 54, 181-187.

- GOODMAN, L.A., 1970. The multivariate analysis of qualitative data: interactions among multiple classifications. *J. Amer. Statist. Assoc.* 65, 226-256.
- GOODMAN, L.A., 1971a. The analysis of multidimensional contingency tables: stepwise procedures and direct estimation methods for building models for multiple classifications. *Technometric* 13, 33-61.
- GOODMAN, L.A., 1971b. Partitioning of chi-square, analysis of marginal contingency tables, and estimation of expected frequencies in multidimensional contingency tables. *J. Amer. Statist. Assoc.* 66, 339-344.
- GOODMAN, L.A. 1973c. Guided and unguided methods for the selection of models for a set of T multidimensional contingency tables. *J. Amer. Statist. Assoc.* 68, 165-175.
- GRIZZLE, J.E., STARMER, C.F., and KOCH, G.G., 1969. Analysis of categorical data by linear models. *Biometrics* 25, 489-504.
- GRIZZLE, J.E. and WILLIAMS, O.D., 1972. Log-linear models and tests of independence for contingency tables. *Biometrics* 28, 137-156.
- HABERMAN, S.J., 1972. Log-linear fit for contingency tables. *Appl. Statist.* 21, 218-225.
- HABERMAN, S.J., 1973a. Log-linear models for frequency data: sufficient statistics and likelihood equations. *Ann. Statist.* 1, 617-632.
- HABERMAN, S.J., 1973b. The analysis of residuals in cross-classified tables. *Biometrics* 29, 205-220.
- HABERMAN, S.J., 1973c. CTAB: analysis of multidimensional contingency tables by log-linear models: user's guide. International Educational Services, Chicago, Illinois. (245)S.

- HABERMAN, S.J., 1974a. The analysis of frequency data. Univ. of Chicago Press, Chicago, Illinois. (610)S.
- HABERMAN, S.J., 1974b. Log-linear models for frequency tables with ordered classifications. *Biometrics* 30, 589-600.
- HABERMAN, S.J., 1977b. Log-linear models and frequency tables with small expected cell counts. *Ann. Statist.* 5, 1148-1169.
- HABERMAN, S.J., 1978a. Analysis of Qualitative Data: Introductory Topics. Academic Press, New York. (359)S.
- HABERMAN, S.J., 1978b. Analysis of Qualitative Data: II. New Developments. Academic Press, New York. (612)S.
- HALDANE, J.B.S., 1955. The estimation and significance of the logarithm of a ratio of frequencies. *Ann. Hum. Genet.* 20, 309-311.
- JAMES, S., 1979. Multivariate log-linear probability models for the analysis of qualitative data. *Bull. Inst. Internat. Statist.* 48 (2), 3-20.
- MARGOLIN, B.H., and LIGHT, R.J., 1974. An analysis of variance for categorical data, II: small sample comparisons with chi-square and other competitors. *J. Amer. Statist. Assoc.* 69, 755-764.
- OKUR, M.C., 1983. Çok boyutlu tabloların logaritmik doğrusal modeller yardımıyla analizi. E.Ü.Mühendislik Fakültesi Dergisi. Seri: E: Uygulamalı istatistik 1 (1-2), 33-44.
- RUTLEDGE, J.J., and GUNSETT, F.C., 1982. Analysis of categorical data in the animal sciences. *J. Animal. Sci.* 54 (5), 1072-1078.

## TEŞEKKÜR

Bu çalışmayı bana yüksek lisans tezi olarak verip yürütmemi sağlayan ve tüm çalışma boyunca yol gösteren Sayın Hocam Doç.Dr.Yüksel BEK'e, çalışmam esnasında bölüm imkanlarını kullanmam için fırsat veren Sayın Hocam Prof.Dr. Lütfi ÖZCAN'a ve yardımlarını esirgemeyen tüm hocalarıma, ayrıca bu tez çalışmasının başından sonuna kadar düzenleme ve gerekli kontrollerde yardımları geçen Ar.Gör.Ercan EFE ve tezin daktilo edilmesinde büyük titizlik gösteren Sekreter Behiye LÜBİÇ'e teşekkür etmeyi bir borç bilirim.



## ÖZGEÇMİŞ

1962 yılında Adana ilinin Karataş ilçesinde doğdum. İlk, orta ve lise öğrenimimi Karataş ilçesinde yaptım. Üniversitelerarası merkezi yerleştirme sınavı sonrasında 1979-1980 ders yılında Ç.Ü.Ziraat Fakültesinin Zootekni Bölümüne kaydolarak 1982-1983 döneminde lisans eğitimini Fakülte ikincisi ve bölüm ikincisi olarak bitirdim. 1982-1983 ders yılında Ç.Ü.Fen Bilimleri Enstitüsü Zootekni Anabilim Dalı Yüksek Lisans eğitim programına başladım.



BU PROGRAM LOG-LINEAR MODEL İÇİN MAKSİMUM LIKELIHOOD  
TAHMİNİNİ YAPMAKTADIR.

$\text{LOG}(M(I,J)/Z(I,J))=A(J)+U(I,J)$ , BURADA  
 $U(I,J)=B(1)*X(I,J,1)+\dots+B(NV)*X(I,J,NV)$ ; DIR.  
BURADA NR SIRALI VE NC SÜTUNLU BİR TABLE N(I,J)  
VERİLMEKTEDİR. N(I,J) NİN BEKLENEN DEĞERİ M(I,J) DIR.  
NORMAL OLARAK Z(I,J) POZİTİFDİR. EĞER Z(I,J) POZİTİF DEĞİLSE  
N(I,J) HÜCRESI ANALİZ ESNASINDA İHMAL EDİLİR. BU, EKSİK TABLO  
ANALİZİNİ GÖSTERİR.

KONTROL KARTLARININ TARIFI İÇİN PROGRAM TANIMINA BAKINIZ.

BU PROGRAMDA, EĞER Z(I,J) 1 İSE İL 0 DIR. Z(I,J) NİN  
BAŞKA BİR SEÇİMİNDE İSE POZİTİFTİR.  
EĞER, B(K,0), 1<=K<=NV BAŞLANGIÇ TAHMİNLERİ PROGRAM  
TARAFINDAN BULUNUYorsa PARAMETRE 0' A ESİTTİR.  
EĞER KULLANICI BAŞLANGIÇ TAHMİNLERİNİ VERECEKSE  
0 ZAMAN PARAMETRE POZİTİF OLACAKTIR.  
EĞER, IR POZİTİF İSE DÜZELTİLMİŞ REZİDÜYALLER VERİLİR.  
EĞER, IC POZİTİF İSE GENELLEŞTİRİLMİŞ REZİDÜYALLER VERİLİR.  
GEREKTEĞİNDE MAX İTERASYON YAPILIR. EĞER KULLANICI MX'İ  
POZİTİF OLARAK TANIMLARSA, MAX DEĞER MX OLUR.  
EĞER MX POZİTİF DEĞİLSE, MAX DEĞER 10 OLUR.  
EĞER  $U(I,J,T)-U(I,J,T-1)$  NİN BÜYÜKLÜĞÜ DEĞİŞMESSE İTERASYON  
DURUR. YANI, BÜYÜKLÜK  $10**IT$  ALINIR. NORMAL OLARAK  $IT=-3$  DÜR.  
FAKAT KULLANICI BU SEÇİMİ DEĞİŞTİREBİLİR..

REAL WORK(10000), FMT(20), TITLE(20)  
LOGICAL ONE, INIT  
DATA MXS, TOLS, NW/10, 0.001, 10000/

PARAMETRELERİN OKUNMASI.

READ(5,1) NR, NC, NV, IL, IS, IR, IC, MX, IT  
1 FORMAT(9I4)  
IF(NR.LE.0) NR=1  
IF(NC.LE.0) NC=1  
ONE = .TRUE.  
INIT = .TRUE.  
IF(IL.GT.0) ONE = .FALSE.  
IF(IS.GT.0) INIT = .FALSE.  
IF(MX.LE.0) MX=MXS  
IF(IT.NE.0) TOL = 10.0\*\*IT  
IF(IT.EQ.0) TOL = TOLS

VERİ İÇİN BİR FORMAT DEĞİŞKENİNİN OKUTULMASI.

READ(5,2) FMT  
YAPILAN İŞ İÇİN BİR İSİM OKUTULMASI. 80 KARAKTERE KADAR  
OLABİLİR.

READ(5,2) TITLE  
2 FORMAT(20A4)

WRITE(6,11) TITLE  
11 FORMAT(1H1,20A4)

-WORK- ALANININ BELIRLENMESI: İLK OLARAK TABLONUN  
BÜYÜKLÜĞÜ BULUNUR.  
NSIZE =NR\*NC

NF:BEKLENEN DEĞERLER (M) TABLOSUNDAKI -WORK- ALANININ  
YERİNİ GÖSTERİR.  
EĞER I1>0 İSE,NZ:Z TABLOSUNUN YERİNİ GÖSTERİR.

NF = NSIZE+1  
IF(ONE) NZ = NF  
IF(.NOT.ONE) NZ = NF+NSIZE

NX: X TABLOSUNUN YERİNİ GÖSTERİR.  
NT: ORTALAMA TABLOSU THETA NIN YERİNİ GÖSTERİR.

NX = NZ+NSIZE  
NX1 = NX-1  
NT = NX+NSIZE\*NV

NS DEKI KOVARYANS MATRISI İÇİN YER AYRILMASI.

NS = NT+NC\*NV

NL DE DEĞİŞKEN İSİMLERİ İÇİN YER AYRILMASI.

NL = NS+NV\*NV

NR DE REZİDÜYALLERİN YERLESTİRİLMESİ.

NRS = NL+NV+NV  
NR2 = NRS-1  
NEN =NR2+NSIZE

NB DE PARAMETRE TAHMINLERİNİN YERLESTİRİLMESİ.

NB =NRS+NSIZE  
ND = NB+NV  
ND1 = ND-1  
NE = ND+NV-1

FAZLA YER İÇİN İSTEK OLUP OLMADIĞININ KONTROLÜ.

IF(NE.GT.NW) WRITE(6,3)  
3 FORMAT(38H ANALİZ İÇİN ÇOK FAZLA YER GEREKİYOR..)  
IF(NE.GT.NW) RETURN

DEĞİŞKEN İSİMLERİNİN OKUTULMASI.

IF(NV.LE.0) GO TO 10  
KK=NL  
4 LL=KK+19  
IF(LL.GT.NR2) LL = NR2



READ(5,2) (WORK(I),I=KK,LL)  
 KK=LL+1  
 IF(LL.LT.NR2) GO TO 4

TABLONUN OKUTULMASI.

10 READ(5,FMT) (WORK(I),I=1,NSIZE)

Z NİN OKUTULUP OKUTULMADIĞINA BAK.X IN OKUTULMASI.

IF(.NOT.ONE) READ(5,FMT) (WORK(I),I=NZ,NX1)  
 IF(NV.LE.0) GO TO 8

KK = NX  
 LL = KK+NSIZE-1  
 DO 5 J=1,NV

READ(5,FMT) (WORK(I),I=KK,LL)  
 KK = KK+NSIZE  
 LL = LL+NSIZE

5 CONTINUE

ARZU EDİLİYORSA B NİN BAŞLANGIÇ DEĞERLERİNİN OKUTULMASI.

IF(INIT) GO TO 8

KK = NB

6 LL = KK+7

IF(LL.GE.ND) LL= ND1

READ(5,7) (WORK(I),I=KK,LL)

7 FORMAT(8F10.5)

KK = LL+1

IF(LL.LT.ND1) GO TO 6

TAHMİNLERİN HESAPLANMASI.

8 CALL CONTIN(NR,NC,NV,MX,TOL,WORK(1),WORK(NF),WORK(NZ),  
 \* WORK(NX),WORK(NT),WORK(NB),WORK(ND),WORK(NRS),WORK(NS),  
 \* NDF,INIT,ONE,IFault)

NETİCELERİN YAZDIRILMASI.

IF(IFault.EQ.1) WRITE(6,9)

9 FORMAT(17H GIRİŞ HATALIDIR.)

IF(IFault.GT.0) RETURN

IF(NV.LE.0) GO TO 27

DO 12 J=1,NV

KK = NS+(J-1)\*(NV+1)

LL = ND+J-1

WORK(LL) = SQRT(WORK(KK))

12 CONTINUE

WRITE(6,13)

13 FORMAT(33H KATSAYILAR VE STANDART HATALAR. )

KK = 1

14 LL = KK+6

IF(LL.GT.NV) LL = NV

KK1 = NL+KK+KK-2

LL1 = NL+LL+LL-1

```

KK2 = NB+KK-1
LL2 = NB+LL-1
KK3 = ND+KK-1
LL3 = ND+LL-1
WRITE(6,15) (WORK(I),I=KK1,LL1)
15 FORMAT(/1X,7(2X,2A4))
WRITE(6,16) (WORK(I),I=KK2,LL2)
16 FORMAT(1X,7F10.5)
WRITE(6,16) (WORK(I),I=KK3,LL3)
KK = KK+7
IF(LL.LT.NV) GO TO 14

```

C  
C  
C

UYUM MÜKEMMELİĞİ İÇİN BİLGİNİN YAZDIRILMASI.

```

27 CALL CHISQRATIO,CHI,NSIZE,WORK(1),WORK(NF)
WRITE(6,17) RATIO,CHI,NDF
17 FORMAT(/32H LIKELIHOOD ORANI KHI-KARESİ = ,F10.5/
*          32H PEARSON KHI-KARESİ = ,F10.5/
*          32H SERBESTLIK DERECESESİ = ,I4)

```

C  
C  
C

ARZU EDİLİRSE DÜZELTİLMİŞ REZİDÜYALLERİN ELDE EDİLMESİ.

```

IF(IR.LE.0) GO TO 22
CALL RESID(NR,NC,NV,WORK(1),WORK(NF),WORK(NS),WORK(NX),
* WORK(NT),WORK(NRS))
WRITE(6,18)
18 FORMAT(/11H I J,16H GÖZLENEN DEĞER,
* 16H BEKLENEN DEĞER,19H DÜZELTİ. REZİDÜY.)
KK1 = 1
KK2 = NF
KK3 = NRS
DO 20 J=1,NC
DO 19 I=1,NR
WRITE(6,21) I,J,WORK(KK1),WORK(KK2),WORK(KK3)
21 FORMAT(1X,2I5,2F16.3,F19.3)
KK1 = KK1+1
KK2 = KK2+1
KK3 = KK3+1
19 CONTINUE
20 CONTINUE

```

C  
C  
C

ARZU EDİLİRSE, GENELLEŞTİRİLMİŞ REZİDÜYALLERİN ELDE EDİLMESİ.

```

22 IF(IC.LE.0) RETURN
WRITE(6,23)
23 FORMAT(/11H DEĞİŞKEN,16H GÖZLENEN DEĞER,
* 16H BEKLENEN DEĞER,19H DÜZELTİL. REZİDÜY.)
DO 30 K=1,IC
READ(5,2) NAME1,NAME2
READ(5,FMT) (WORK(I),I=NRS,NEV)
CALL GENRES(NR,NC,NV,WORK(1),WORK(NF),WORK(NX),
* WORK(NT),WORK(NS),WORK(NRS),WORK(ND),OBS,EXP,RES)
WRITE(6,24) NAME1,NAME2,OBS,EXP,RES
24 FORMAT(3X,2A4,2F16.3,F19.3)
30 CONTINUE

```

RETURN  
END

C  
C  
SUBROUTINE CONTIN(NR,NC,NV,MX,TOL,TABLE,FIT,Z,X,THETA,B,D,U,S,  
\* NDF,INIT,ONE,IFAU)LT)

C  
C BU PROGRAM LOG-LINEAR MODELLERDE PARAMETRELERIN  
C MAXIMUM-LIKELIHOOD TAHMİNLERİNİ HESAPLAR.  
C NEWTON-RAPHSON ALGORİTMA S^ MULTİNOMİAL DAĞ^L^ İ İN  
C KULLAN^LMAKTAD^R.ÜZERİNDE ÇALIŞILAN FREKANS TABLOSU NR  
C SİRALI NC SÜTUNLU BİR DİZEDİR. TABLO (I,J) ^NİN HER HÜCRESİNE  
C X(I,J,K)^NİN NV GÖZLEMİ VE Z(I,J)^NİN BİR GÖZLEMİ TEKABÜL EDER.  
C TAHMİN EDİLEN BEKLENEN DEĞER-FIT(I,J)-AŞAĞIDAKİ MODELİ SAĞLAR.  
C  $FIT(I,J) = Z(I,J) * EXP(A(J) + U(I,J))$   
C VE  $U(I,J) = B(1) * X(I,J,1) + \dots + B(NV) * X(I,J,NV)$ .  
C B^NİN ASİMTOTİK KOVARYANS MATRİSİ S DİR.  
C U^NUN KOORDİNATLARINDAKİ HİÇBİR DEĞİŞİKLİK TOL DAN DAHA BÜYÜK  
C OLMAYINCAYA VEYA MAX İTERASYON YAPILINCAYA KADAR İTERASYONA  
C DEVAM EDİLİR.

C EN SON İTERASYONDA B ^DEKİ DEĞİŞİME D ^DİR.  
C EĞER INIT .TRUE.İSE,B^NİN BAŞLANGIÇ DEĞERLERİ İ^NİN 1 DEN NR YE,  
C J NİN 1 DEN NC YE KADAR DEĞERLERİ İÇİN  $0.5 + TABLE(^,J)$  ^NİN  
C LOGARİTMALARINA DAYANAN TARTILI REGRESYON İLE BULUNUR.  
C EĞER INIT .FALSE.İSE,BAŞLANGIÇ DEĞERLERİNİ KULLANICI VERİR.  
C NORMAL İŞLEMLER ALTINDA IFAULT 0 ^A ÇEVİRİLİR.  
C EĞER TABLONUN VEYA Z NİN ELAMANLARINDAN NEGATİF OLANLAR VARSA  
C IFAULT 1 OLUR.E ER ONE .TRUE.İSE, T M Z(I,J) LER 1 VARSAY^L^R.  
C THETA(J,K) TARTILI ORTALAMADIR.  
C  $(X(I,J,K) * FIT(1,J) + \dots + X(NR,J,K) * FIT(NR,J)) / (FIT(1,J) + \dots$   
C  $+ FIT(NR,J))$ ,  
C NDF Kİ-KARE TESTLERİ İÇİN SERBESTLİK DERESESİDİR.

C  
C REAL TABLE(NR,NC),FIT(NR,NC),Z(NR,NC),U(NR,NC),X(NR,NC,NV),  
C \* THETA(NC,NV),B(NV),D(NV),S(NV,NV)  
C LOGICAL INIT,INIT1,ONE  
C DOUBLE PRECISION SUM,SUM1

C  
C ALGORİTMA BAŞLANGICI.

C  
C F = 1.0E30  
C IFAULT = 0  
C NR1 = NR  
C NC1 = NC  
C NV1 = NV  
C NDF = (NR1-1)\*NC1  
C INIT1 = INIT  
C IF(NV1.LE.0) INIT1 = .FALSE.  
C IF(.NOT.INIT1.AND.ONE) GO TO 5  
C DO 2 J=1,NC1  
C SUM = 0.000  
C SUM1 = 0.000  
C K = 0  
C IF(ONE) K = 1  
C DO 1 I=1,NR1

```

IF(TABLE(I,J).LT.0.0) IFAULT = 1
IF(IFAULT.GT.0) RETURN
IF(INIT1) FIT(I,J) = TABLE(I,J)+0.5
IF(INIT1) U(I,J) = ALOG(FIT(I,J))
IF(ONE) GO TO 1
IF(Z(I,J).LT.0.0) IFAULT = 1
IF(IFAULT.GT.0) RETURN
IF(Z(I,J).EQ.0.0) FIT(I,J) = 0.0
IF(Z(I,J).EQ.0.0) NDF = NDF-1
IF(INIT1.AND.Z(I,J).GT.0.0) U(I,J) = U(I,J)-ALOG(Z(I,J))
IF(Z(I,J).GT.0.0) K = 1
1 CONTINUE
NDF = NDF+1-K
DO 101 I=1,NR1
SUM = SUM+FIT(I,J)
SUM1 = SUM1+FIT(I,J)*U(I,J)
101 CONTINUE
IF(SUM.LE.0.000) GO TO 2
W = SUM1/SUM
DO 102 I=1,NR1
102 U(I,J) = U(I,J)-W
2 CONTINUE
IF(.NOT.INIT1) GO TO 5
DO 4 K=1,NV1
B(K) = 0.0
SUM = 0.000
DO 3 J=1,NC1
DO 3 I=1,NR1
3 SUM = SUM+FIT(I,J)*U(I,J)*X(I,J,K)
D(K) = SUM
4 CONTINUE

C
C
C   BİR İTERASYONUN YAPILMASI.

5 DO 36 IT=1,MX
IF(IT.EQ.1.AND..NOT.INIT1) GO TO 20
DO 10 J=1,NC1
SUM = 0.000
DO 7 I=1,NR1
7 SUM = SUM+FIT(I,J)
W = SUM
DO 9 K=1,NV1
SUM = 0.000
DO 8 I=1,NR1
8 SUM = SUM+FIT(I,J)*X(I,J,K)
IF(W.GT.0.0) THETA(J,K) = SUM/W
IF(W.LE.0.0) THETA(J,K) = 0.0
9 CONTINUE
10 CONTINUE

C
C
C   ÇAPRAZ ÇARPIMLARIN TARTILI TOPLAMLARININ BULUNMASI.

DO 15 K=1,NV1
DO 14 L=1,K
SUM = 0.000

```

```

DO 13 J=1,NC1
DO 12 I=1,NR1
12 SUM = SUM+(X(I,J,K)-THETA(J,K))*(X(I,J,L)-THETA(J,L))*FIT(I,J)
13 CONTINUE
S(K,L) = SUM
14 CONTINUE
15 CONTINUE

```

C  
C  
C S \*NIN CHOLESKY DEKOMPOZISYONU.

CALL CHOL (NV1,S,S,IRANK)

C  
C  
C DAHA ILERI ADIMLARIN GEREKLI OLUP OLMADIĞININ KONTROLÜ.

IF(F.LT.TOL) GO TO 37

C  
C  
C B \*NIN YENİ DEĞERİNİN ELDE EDİLMESİ.

CALL SOLVE(NV1,S,D,D)

DO 16 K=1,NV1

16 B(K) = B(K)+D(K)

C  
C  
C U\* NUN AYARLANMASI,UYDURULMASI VE KONVERJENS İÇİN KONTROL EDİLMESİ.

IF(IT.GT.1) F=0.0

20 DO 33 J=1,NC1

DO 30 I=1,NR1

SUM = 0.000

IF(NV1.LE.0) GO TO 22

DO 21 K=1,NV1

21 SUM = SUM+B(K)\*X(I,J,K)

IF(IT.EQ.1) GO TO 22

E = SUM

E = ABS(E-U(I,J))

IF(E.GT.F) F=E

22 U(I,J) = SUM

FIT(I,J) = EXP(U(I,J))

IF(ONE) GO TO 30

IF(Z(I,J).LT.0.0) IFAULT = 1

IF(IFFAULT.GT.0) RETURN

FIT(I,J) = Z(I,J)\*FIT(I,J)

30 CONTINUE

SUM = 0.000

W = 0.0

DO 31 I=1,NR1

IF(.NOT.ONE.AND.Z(I,J).LE.0.0) GO TO 31

W = W+TABLE(I,J)

SUM = SUM+FIT(I,J)

31 CONTINUE

IF(SUM.GT.0.000) W=W/SUM

IF(SUM.LE.0.000) W=0.0

DO 32 I=1,NR1

32 FIT(I,J) = W\*FIT(I,J)

33 CONTINUE

IF(NV1.LE.0) RETURN

```
C
C   UYDURULAN VE GÖZLENEN LİNEAR KOMBİNASYONLAR ARASINDAKİ
C   FARKLARIN HAZIRLANMASI.
C
DO 35 K=1,NV1
SUM = 0.000
DO 34 J=1,NC1
DO 34 I=1,NR1
IF(FIT(I,J).GT.0.0) SUM = SUM+(TABLE(I,J)-FIT(I,J))*X(I,J,K)
34 CONTINUE
D(K) = SUM
35 CONTINUE
36 CONTINUE
37 CALL INVERT(NV1,S,S)
NDF = NDF-IRANK
RETURN
END

C
C   SUBROUTINE SOLVE(N,LU,B,X)
C
C   I=1,..,N KADAR U(I,I) N`N S`F`R OLMAYAN DEĞERLERİ İÇİN
C   LUX = B NİN ÇÖZÜMÜ.
REAL X(N),LU(N,N),B(N)
DOUBLE PRECISION SUM
NN = N

C
C   DAHA UZUN ÜÇGEN İÇİN ÇÖZÜM.
C
X(1) = B(1)
IF(LU(1,1).LE.0.0) X(1) = 0.0
IF(NN.EQ.1) GO TO 3
DO 2 I=2,NN
SUM = 0.000
I1 = I-1
DO 1 J=1,I1
1 SUM = SUM+LU(I,J)*X(J)
X(I) =B(I)-SUM
IF(LU(I,I).LE.0.0) X(I) = 0.0
2 CONTINUE

C
C   ÜST ÜÇGEN İÇİN ÇÖZÜM .
C
3 I = NN
IF(LU(I,I).GT.0.0) X(I) = X(I)/LU(I,I)
IF(LU(I,I).LE.0.0) X(I) = 0.0
IF(NN.EQ.1) RETURN
DO 5 K=2,NN
SUM = 0.000
I1 = I
I =I-1
IF(LU(I,I).LE.0.0) X(I) =0.0
IF(LU(I,I).LE.0.0) GO TO 5
DO 4 J=I1,NN
4 SUM = SUM+LU(I,J)*X(J)
```

```

X(I) = (X(I)-SUM)/LU(I,I)
5 CONTINUE
RETURN
END

```

```

SUBROUTINE CHISQ(RATIO,CHI,NSIZE,TABLE,FIT)

```

```

LIKELIHOOD ORANI(RATIO) VE PEARSON(CHI)KHI-KARESİ İSTATİS-
TİKLERİNİN HESAPLANMASI.GÖZLENEN DEĞERLER TABLOSU NSIZE
ELAMANA SAHIPTIR.AYNI ŞEKİLDE,UYDURULAN HÜCRE ORTALAMALARI
TABLOSUDA NSIZE ELEMANA SAHIPTIR.

```

```

INTEGER NSIZE
REAL TABLE(NSIZE),FIT(NSIZE)
DOUBLE PRECISION SUM,SUM1
SUM = 0.000
SUM1 = 0.000
DO 1 I=1,NSIZE

```

```

NOTE THAT FIT(I) =0 IMPLIES TABLE(I) =0.

```

```

IF(FIT(I).LE.0.0) GO TO 1
SUM1 = SUM1+(TABLE(I)-FIT(I))*2/FIT(I)
IF(TABLE(I).LE.0.0) GO TO 1
SUM =SUM+TABLE(I)*ALOG(TABLE(I)/FIT(I))
1 CONTINUE
CHI= SUM1
RATIO = 2.0*SUM
RETURN
END

```

```

SUBROUTINE RESID(NR,NC,NV,TABLE,FIT,S,X,THETA,R)

```

```

RESID ALT PROGRAMI UYDURULAN FIT TABLOSU ,X VE THETA DIZE-
LERİ VE CONTIN ALT PROGRAMININ ASİMTOTİK KOVARYANS MATRİSİ S
İÇİN DÜZELTİLMİŞ REZİDÜYALLERİ R HESAPLAR.
PARAMETRELER CONTİNDEKİ GİBİ TANIMLANIR.
REAL TABLE(NR,NC),FIT(NR,NC),S(NV,NV),X(NR,NC,NV),R(NR,NC),
* THETA(NC,NV)
DO 8 J=1,NC
SUM1 = 0.0
DO 10 I=1,NR
10 SUM1 = SUM1+FIT(I,J)
11 DO 7 I=1,NR
SUM = 0.0
IF(NV.LE.1) GO TO 3
DO 2 K=2,NV
LL= K-1
DO 1 L=1,LL
SUM =SUM+S(K,L)*(X(I,J,K)-THETA(J,K))*(X(I,J,L)-THETA(J,L))
1 CONTINUE
2 CONTINUE
SUM = 2.0*SUM

```

```

3 IF(NV.LE.0) GO TO 5
  DO 4 K=1,NV
    SUM = SUM+S(K,K)*(X(I,J,K)-THETA(J,K))**2
4 CONTINUE
5 SUM =FIT(I,J)*(1.0-FIT(I,J)*SUM)
  IF(SUM1.GT.0.0) SUM = SUM-FIT(I,J)*FIT(I,J)/SUM1
  R(I,J) = 0.0
  IF(SUM.LE.0.0) GO TO 7
  R(I,J) = (TABLE(I,J)-FIT(I,J))/SQRT(SUM)
7 CONTINUE
8 CONTINUE
  RETURN
  END

```

```

C
C
SUBROUTINE INVERT(N,TRI,XINV)

```

```

C
C
MODIFIYE EDİLMİŞ CHOLESKY DEKOMPOZİSYON TRI İLE BİR MATRİSİN
C
C
İNVERSİNİN HESAPLANMASI.ÇIKIŞ - XINV- 'DİR.

```

```

C
C
REAL TRI(N,N),XINV(N,N)
DOUBLE PRECISION SUM
NN = N

```

```

C
C
KÖŞEĞEN ELEMANLARIN İNVERSi

```

```

C
C
DO 1 I=1,NN
  IF(TRI(I,I).GT.0.0) XINV(I,I) = 1.0/TRI(I,I)
  IF(TRI(I,I).LE.0.0) XINV(I,I) = 0.0
1 CONTINUE

```

```

C
C
EĞER N= 1 İSE,TAMAM` BITİRİLMİ OLUR.

```

```

C
C
IF(NN.LE.1) RETURN

```

```

C
C
KÖŞEĞEN HARİCİ ELEMANLARIN İNVERSi

```

```

C
C
DO 5 I=2,NN
  I1 = I-1
  DO 4 K=1,I1
    SUM = -TRI(I,K)
    IF(K.GE.I1) GO TO 3
    K1 = K+1
    DO 2 J=K1,I1
2 SUM = SUM-TRI(I,J)*XINV(J,K)
3 XINV(I,K) = SUM
  XINV(K,I) =XINV(I,K)*XINV(I,I)
4 CONTINUE
5 CONTINUE

```

```

C
C
TÜM İNVERSİN BULUNMASI.

```

```

C
C
DO 9 I=1,NN
  I1 = I+1
DO 8 J=1,I

```



```

SUM = XINV(J,I)
IF(I.EQ.NN) GO TO 7
DO 6 K=I1,NN
6 SUM = SUM+XINV(K,I)*XINV(J,K)
7 XINV(I,J) = SUM
8 CONTINUE
9 CONTINUE
DO 10 I=2,NN
I1 = I-1
DO 10 J=1,I1
10 XINV(J,I) = XINV(I,J)
RETURN
END

```

```

SUBROUTINE CHOL(N,SYM,TRI,IRANK)

```

```

N X N BOYUTLU SYM MATRISININ MODIFIYE EDILMIS CHOLESKY
DEKOMPOZISYONUNUN HESAPLANMASI.SYM MATRISI NEGATIF OLMAYAN
SONLU MATRISTIR.SONUC - TRI - DIR.IRANK= SYM 'NIN RANK'IDIR.

```

```

REAL SYM(N,N),TRI(N,N)
DOUBLE PRECISION SUM
DATA TOL/1.0E-4/
NN = N
IRANK = 0
DO 6 I=1,NN
X = TOL*SYM(I,I)
TRI(1,I) = SYM(I,1)
IF(I.EQ.1) GO TO 4
IF(TRI(1,1).GT.0.0) TRI(I,1) = TRI(1,I)/TRI(1,1)
DO 3 J=2,I
IF(TRI(J,J).LE.0.0) GO TO 3
SUM = 0.000
K = J-1
DO 2 L=1,K
2 SUM = SUM+TRI(I,L)*TRI(L,J)
TRI(J,I) = SYM(I,J)-SUM
IF(J.LT.I) TRI(I,J) = TRI(J,I)/TRI(J,J)
3 CONTINUE
4 IF(TRI(I,I).GT.X) IRANK = IRANK+1
IF(TRI(I,I).GT.X) GO TO 6
DO 5 J=1,NN
TRI(I,J) = 0.0
TRI(J,I) = 0.0
5 CONTINUE
6 CONTINUE
RETURN
END

```

```

SUBROUTINE GENRES(NR,NC,NV,TABLE,FIT,X,THETA,S,C,D,OBS,EXP,RES)

```

```

GENRES ALT PROGRAMI GENELLESTIRILMIS REZIDUYALLERI HESAPLAR.
NR,NC,NV,TABLE,FIT,X,THETA, VE S CONTIN'DEKI GIBI TANIMLANIR.

```

C      C=KULLANILAN TARTILAR DIZESİDİR. D =NV UZUNLUĞUNDA -WORK-DIZESİ  
 C      DIR. ÇIKIŞTA:GÖZLENEN LINEAR KOMBİNASYON OBS,TAHMİN EDİLEN  
 C      BEKLENEN DEĞER EXP,VE DÜZELTİLMİŞ REZİDÜYALLER RES ELDE  
 C      EDİLMEKTEDİR.

C      REAL TABLE(NR,NC),FIT(NR,NC),X(NR,NC,NV),THETA(NC,NV),S(NV,NV),  
 \*      C(NR,NC),D(NV)

C      DOUBLE PRECISION SUM,SUM1,SUM2

C      BAŞLANGIÇ DEĞERİ VERİLMESİ.

C      NV1 = NV

C      NR1 = NR

C      NC1 = NC

C      ASİMTOTİK VARYANS TAHMİNLERİNİN ELDE EDİLMESİ.

C      SUM2 = 0.000

C      DO 2 J=1,NC1

C      SUM = 0.000

C      SUM1 = 0.000

C      DO 1 I=1,NR1

C      SUM1 = SUM1+FIT(I,J)\*C(I,J)

1      SUM = SUM+FIT(I,J)

C      SUM2 = SUM2+SUM1\*SUM1/SUM

2      CONTINUE

C      IF(NV1.LE.0) GO TO 14

C      DO 9 K=1,NV1

C      SUM = 0.000

C      DO 8 J=1,NC1

C      DO 7 I=1,NR1

C      SUM = SUM+(X(I,J,K)-THETA(J,K))\*FIT(I,J)\*C(I,J)

7      CONTINUE

8      CONTINUE

C      D(K) = SUM

9      CONTINUE

C      SUM = 0.000

C      IF(NV1.LE.1) GO TO 12

C      DO 11 K=2,NV1

C      LL = K-1

C      DO 10 L=1,LL

10      SUM = SUM+S(K,L)\*D(K)\*D(L)

11      CONTINUE

C      SUM =SUM+SUM

12      DO 13 K=1,NV1

13      SUM = SUM+S(K,K)\*D(K)\*D(K)

14      SUM1 = 0.000

C      DO 16 J=1,NC1

C      DO 15 I=1,NR1

C      SUM1 = SUM1+FIT(I,J)\*C(I,J)\*\*2

15      CONTINUE

16      CONTINUE

C      ST = SUM1-SUM-SUM2

C      GÖZLENEN VE BEKLENEN DEĞERLERİN ELDE EDİLMESİ.

EK I- DEVAM

FILE: SERAP      FORTRAN H      C.U . VM/SP

C

```
RES =0.0
SUM =0.000
SUM1 =0.000
DO 18 J=1,NC1
DO 17 I=1,NR1
SUM =SUM+TABLE(I,J)*C(I,J)
SUM1 =SUM1+FIT(I,J)*C(I,J)
17 CONTINUE
18 CONTINUE
OBS =SUM
EXP = SUM1
```

C

C

C

DÜZELTİLMİŞ REZİDÜYALLERİN ELDE EDİLMESİ.

```
IF(ST.LE.0.0) RETURN
ST = SQRT(ST)
RES = (OBS-EXP)/ST
RETURN
END
```

