ORİJİNAL ARAŞTIRMA ORIGINAL RESEARCH

Karesel Olumsallık Tablolarında Asimetri ve Çarpık Simetri Modelleri

Asymmetry and Skew-Symmetry Models for Square Contingency Tables

Gökçen ALTUN,[a]
Serpil AKTAŞ[a]

[a]İstatistik Bölümü,
Hacettepe Üniversitesi Fen Fakültesi,
Ankara

Geliş Tarihi/*Received:* 29.01.2016
Kabul Tarihi/*Accepted:* 14.03.2016

Yazışma Adresi/*Correspondence:*
Serpil AKTAŞ
Hacettepe Üniversitesi Fen Fakültesi,
İstatistik Bölümü, Ankara,
TÜRKİYE/TURKEY
spxl@hacettepe.edu.tr

ÖZET Amaç: Karesel olumsallık tabloları bağımlı örneklemlerde ortaya çıkan, satır ve sütun değişkenleri aynı düzeylere sahip olan çapraz tablolarıdır. Bu tablolarının çözümlenmesinde bazı özel modeller kullanılır. Bu modeller çoğunlukla simetri modelleridir. Bu çalışmada iki boyutlu karesel olumsallık tablolarında simetri yapısı bozulduğu durumlar için geliştirilen modeller tanıtılmıştır. Bu modellerde olumsallık tablosundaki değişkenlerin sıralabilir ölçekte olduğu varsayılır. **Gereç ve Yöntemler**: Karesel olumsallık tabloların çözümlenmesinde simetri yapısının bozulduğu durumlar için kullanılan asimetri ve çarpık simetri modelleri 280 kanser hastasına ait veriler üzerinde gösterilmiş ve sonuçlar tartışılmıştır. Bu modellerin çözümlenmesinde SPSS ve SAS programları kullanılmıştır. En uygun model için parametre tahmini ile yorumlama yapılmıştır. **Bulgular:** Veriler simetri yapısını ifaden modellere uyum sağlamamıştır. Asimetri ve çarpık simetri modellerinin bütün alt modelleri veriye uyum sağlarken tekdüze çarpık simetri modeli en uyumlu model olarak bulunmuştur. Tekdüze çarpık simetri modeli varsayımı altında odds oranı 1,34 olarak hesaplanmıştır. **Sonuç:** Sitolojik ve patalojik sınıflandırma arasındaki uyum %74 bulunmuştur, bu değer yüksek uyum olarak değerlendirilebilir. Sitolojik tanı ile şüpheli teşhisi konan vakalar patolojik tanı ile desteklenmeden kesinleştirilemez. Kanser tanısının doğruluğu açısından sitolojik ve patalojik sınıflama arasında mükemmel uyum olması beklendiğinden köşegen dışı elemanların da incelenmesi gerekmektedir. Bu da karesel olumsallık tabloları çözümlemeleri ile mümkündür.

Anahtar Kelimeler: Karesel olumsallık tablosu; simetri modeli; asimetri modeli; çarpık simetri modeli

ABSTRACT Objective: Square contingency tables that arise in dependent samples where the row and column variables have same level. Some specific models used in the analysis of these kinds of tables. These models are mostly in the symmetrical pattern. In this study, the models are introduced for two dimensional square contingency tables where the departure from the symmetry structure is observed. Models described in this paper assume ordinal categories for the contingency table. **Material and Methods:** Asymmetry and skew symmetry models used when the symmetry structure is distorted in analysis of square contingency tables are shown on a 280 cancer patient's dataset and the results are discussed. SPSS and SAS programs are used in the analysis of these models. The interpretations are made with the parameter estimation for the best fitting model. **Results:** Symmetry models do not hold for the data. While the sub models of asymmetry and skew symmetry models hold for the data, the best fitting model is determined as the uniform skew symmetry model for the cancer data. The odds ratio under uniform skew smmetry model is estimated as 1.34. **Conclusion:** The agreement between the cytological and pathalogical diagnosis of uterine cancer is 74% suggesting the high agreement. A cytological diagnosis recorded as suspicious is not considered as diagnostic of cancer unless supported by pathological diagnosis. As it is expected perfect agreement between the cytological classification and pathological classification in terms of accurate diagnosis, the off-diagonal cells should be studied as well. This could be done with the anaysis of square contingency tables.

Key Words: Square contingency tables; symmetry model; asymmetry model; skew symmetry model

doi: 10.5336/biostatic.2016-50376

Turkiye Klinikleri J Biostat 2016;8(2):152-61

Birçok uygulamalı bilim dalında araştırmacılar, değişkenler arasındaki ilişki veya değişken düzeyleri arasındaki farklılığı incelemeyi amaçlar. Düzeyleri sözel olarak ifade edilebilen değişkenler kategorik veri olmakla beraber, düzeyleri sayısal olarak ifade edilebilen değişkenlerin alabileceği değerler sınıflandırılarak kategorik veri haline getirilebilir. Kategorik değişken, her bir gözlemin belirli bir kategoriye, yani sınıfa ait olduğu sınıflanabilen ve sıralanabilen özelliğe sahip değişkenlerdir.

Kategorik değişkenlerin birleşik dağılımı olumsallık tabloları ile özetlenir. Olumsallık tabloları iki ya da daha fazla kategorik değişkenin ortak sıklık dağılımıdır. Olumsallık tablolarında her değişken belirli sayıda düzeye sahiptir. Satır değişkeni düzey sayısı I , sütun değişkeni düzey sayısı I olmak üzere satır ve sütun değişkenlerinin aynı kriterlere göre sınıflandırıldığı bağımlı örneklem çalışmaları $I \times I$ boyutlu karesel tablo çözümlemeleri ile yorumlanır. Karesel olumsallık tabloları hastaların ilk ve son muayene durumları, bir hastalığın iki uzman tarafından teşhis edildiği durumlar; sağ ve sol gözün görme derecesi gibi tekrarlı ölçümlü değişkenler için düzenlenirler. Karesel olumsallık tablolarında çapraz tabloların çözümlenmesinde sıklıkla kullanılan "bağımsızlık çözümlemesi" yapılması tablonun bağımlılık yapısından dolayı uygun değildir.[1] Bu nedenle bazı özel çözümlemeler önerilmiştir.

Karesel olumsallık tablolarının çözümlenmesi ilk olarak tam simetri modelinin önerilmesiyle başlamıştır.[1] Daha sonra, yarı simetri (YS), marjinal homojenlik (MH), koşullu simetri (KS), köşegen parametre simetri (KPS) modelleri gibi özel modellerin kullanılmasıyla devam etmiştir.[2] Son yıllarda önerilen bu modellerin yazılım programlarında uygulanabilmesi için tasarım matrisleri elde edilmiştir.[3] Karesel olumsallık tablolarının çözümlenmesinde öncelikle simetri yapısını ifade eden modeller ve eğer simetri yapısı bozuluyorsa bu durumlar için önerilen diğer modeller uygulanmalıdır.

Çalışma kapsamında bu modeller teorik olarak incelenmiş ve gerçek bir veri kümesi üzerinde uygulaması yapılarak sonuçlar tartışılmıştır.

KARESEL OLUMSALLIK TABLOLARI

Olumsallık tabloları iki ya da daha fazla kategorik değişkenin ortak sıklık dağılımıdır. Bir olumsallık tablosu iki ya da daha fazla değişkenin düzeylerinin çapraz sınıflandırılması olarak da düşünülebilir. Satır ve sütun değişkenlerinin düzeylerinin aynı olduğu tablolar $I \times I$ boyutlu karesel olumsallık tablolarıdır (Tablo 1).

KARESEL OLUMSALLIK TABLOLARINDA SİMETRİ YAPISINI İFADE EDEN MODELLER

Karesel olumsallık tablolarında çeşitli durumlar için simetrik yapıyı ifade eden modeller Tam Simetri Modeli, Yarı Simetri Modeli ve Marjinal Homojenlik Modelidir.[4]

Tam Simetri Modeli (TS)

$I \times I$ boyutlu karesel olumsallık tablosunun ana köşegenine göre simetrik gözelere ait olasılıkların ya da beklenen sıklıkların eşitliğini test eden Tam Simetri Modeli (TS)' ne ait hipotez. p_{ij} : Herhangi bir deneğin $i.$ satır $j.$ sütunda yer alma olasılığını göstermek üzere Eşitlik (1)' de ifade edildiği gibidir.

$$H_{TS} : p_{ij} = p_{ji} \qquad i,j=1,2,....,I \tag{1}$$

Tam simetri modeli altında beklenen sıklıklara ilişkin en çok olabilirlik tahmin eşitlikleri m_{ij} : i. satır, j. sütuna ait beklenen sıklık değeri olmak üzere Eşitlik (2)'de verilmiştir.

TABLO 1: Karesel olumsallık tablosunun gösterimi

Satır Değişkeni	Sütün Değişkeni				Toplam
	1	2	$\cdots$	I	
1	x_{11}	x_{12}	$\cdots$	x_{1I}	$x_{1.}$
2	x_{21}	x_{22}	$\cdots$	x_{2I}	$x_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
I	x_{I1}	x_{I2}	$\cdots$	x_{II}	$x_{I.}$
Toplam	$x_{.1}$	$x_{.2}$	$\cdots$	$x_{.I}$	$x_{..}$

$p_{ij} = \frac{x_{ij}}{x_{..}}$, $x_{i.} = \sum_{j=1}^{I} x_{ij}$, $x_{.j} = \sum_{i=1}^{I} x_{ij}$, $p_{i.} = \frac{x_{i.}}{x_{..}}$,

$p_{.j} = \frac{x_{.j}}{x_{..}}$, $p_{ij/i.} = \frac{x_{ij}}{x_{i.}}$, $p_{ij/.j} = \frac{x_{ij}}{x_{.j}}$

$$\hat{m}_{ij} = \begin{cases} \dfrac{x_{ij} + x_{ji}}{2} & \mathrm{i} \neq \mathrm{j} \\ x_{ii} & \mathrm{i} = \mathrm{j} \end{cases} \qquad (2)$$

Tam simetri modelinin serbestlik derecesi; $sd_{TS} = \frac{I(I-1)}{2}$ dir. Tam simetri modeline ilişkin parametrelerin gözelere göre dağılımı matris yapısında verilmiştir. TS gösterimiyle ifade edilen bu matrisi oluşturan TS_{ij} elemanları tüm (i,j) ler için $k = |i - j|$ olmak üzere aşağıda verilen eşitlik ile ifade edilir ;

$$TS_{ij} = \begin{cases} (k+1) - (i+1)(\frac{1}{2}i+1) + (\mathrm{I}+3)(i+1) - 3 - 2I & i \le j \\ (k+1) - (j+1)(\frac{1}{2}j+1) + (\mathrm{I}+3)(j+1) - 3 - 2I & i > j \end{cases} \qquad (3)$$

Bilgisayar ortamındaki çözümlemelerde faktör değişkeni olarak belirtilen TS matris yapısı simetri modelinin uygulanması için gereklidir ve 4×4 karesel olumsallık tablosu için TS matris gösterimi aşağıdaki gibidir;

$$TS = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 6 & 7 \\ 3 & 6 & 8 & 9 \\ 4 & 7 & 9 & 10 \end{bmatrix}$$

Yarı Simetri Modeli (YS)

Caussinus (1966) tarafında tanımlanan, ana köşegenin bir tarafındaki gözelerden elde edilen odds oranlarının, köşegenin diğer tarafındaki gözelerden elde edilen odds oranlarına eşitliğini test eden Yarı Simetri Modeli (YS)' ne ait hipotez Eşitlik (4)'de verilmiştir.[5]

$$H_{YS} : p_{ij} = \alpha_j p_{ji} \qquad \mathrm{i} \neq \mathrm{j} \qquad (4)$$

Yarı simetri modeli altında beklenen sıklıklara (m_{ij}) ilişkin en çok olabilirlik tahminleri aşağıda verilen kısıtları sağlamalıdır:

$$\begin{aligned} \hat{m}_{i.} &= x_{i.}, \quad \mathrm{i}=1,2,...,\mathrm{I} \\ \hat{m}_{.i} &= x_{.i}, \quad \mathrm{i}=1,2,...,\mathrm{I} \\ \hat{m}_{ij} + \hat{m}_{ji} &= x_{ij} + x_{ji}, \quad \mathrm{i} \neq \mathrm{j} \end{aligned} \qquad (5)$$

Yarı simetri modelinin serbestlik derecesi; $sd_{YS} = \frac{(I-1)(I-2)}{2}$ dir.

Marjinal Homojenlik Modeli (MH)

Satır ve sütun değişkenlerinin aynı düzeylerine ait marjinal olasılıklarının birbirine eşitliğini test eden Marjinal Homojenlik Modeli (MH) Stuart (1955) tarafından önerilmiştir. MH' ne ait hipotez Eşitlik (6)' da ifade edildiği gibidir.[6]

X: satır değişkeni, *Y*: sütun değişkeni olmak üzere

$$H_{MH} : p_{i.} = p_{.i} \quad \mathrm{i}=1,2...,\mathrm{I} \qquad (6)$$

ile ifade edilir. Burada,

$$p_{i.} = \sum_{t=1}^{I} p_{it} \qquad p_{.i} = \sum_{s=1}^{I} p_{si}$$

F_i^X ve F_i^Y terimleri birikimli marjinal olasılıkları göstermek üzere,

$$F_i^X = \Pr(X \le i) = \sum_{k=1}^{i} p_{k.} \qquad \mathrm{i}=1,2,....,\mathrm{I}-1$$

$$F_i^Y = \Pr(Y \le i) = \sum_{k=1}^{i} p_{.k} \qquad \mathrm{i}=1,2,....,\mathrm{I}-1$$

MH modeline ait farklı gösterimler Eşitlik (7) ve (8)' de ifade edildiği gibidir;

$$H_{MH} : \Pr(X = i) = \Pr(Y = i) \quad \mathrm{i}=1,2,....,\mathrm{I} \qquad (7)$$

$$H_{MH} : F_i^X = F_i^Y, \quad \mathrm{i}=1,2,....,\mathrm{I}-1 \qquad (8)$$

Marjinal homojenlik modelinin serbestlik derecesi $(I-1)$ dir.

TS, YS ve MH modelleri karesel olumsallık tablolarındaki simetri yapısını çeşitli durumlarda tanımlamaktadır. Simetri yapısına uygunluk TS modeline uyumla sağlanır. Caussinus (1965)'in tanımına göre veri analizinde TS modeline uyum yoksa, bunun neden kaynaklandığını görmek ya da farklı durumlar için simetri yapısını incelemek için simetrik yapıyı ifade eden diğer iki model MH ve YS modellerine bakılır.[5]

Simetri yapısını tanımlayan modeller uyumlu olmadığında veya sapmalar meydana geldiğinde ise bu sapmaları tanımlayan ve sapmaların derecelerini belirleyen simetrik olmayan yapıları ifade eden modeller önerilmiştir.

Karesel Olumsallık Tablolarında Simetri Yapısından Sapan Modeller

Karesel olumsallık tablolarında 2. Bölümde verilen modellere uyum sağlanmadığında simetri modellerinden sapma olduğu söylenir. Bu durumda Asimetrik modeller daha uygun sonuçlar verir. Bu bölümde asimetri modelleri tanıtılmıştır.

Asimetri Modelleri

Tam simetri modelinden sapmaları ifade eden model Goodman (1985) tarafından asimetrik model olarak tanımlanmıştır.[7] Asimetrik modeller tam simetri modelini temel aldıkları için tam simetri modeli "Boş Asimetri" modeli olarak da adlandırılır.

Üçgen Asimetri Modeli (ÜA)

Goodman (1985), ρ_{ij} simetri parametresi τ_1 ve τ_2 koşul parametresi olmak üzere üçgen asimetri modelini ifade etmiştir. ÜA modeli alt ve üst üçgenlere ait beklenen olasılıkların eşitliğini test eder.[7] Modele ait hipotez $\rho_{ij} = \rho_{ji}$ koşulu altında Eşitlik (9)' da ifade edildiği gibidir;

$$H_{ÜA} : p_{ij} = \begin{cases} \rho_{ij}\tau_1 & i > j \\ \rho_{ij}\tau_2 & i < j \end{cases} \qquad \text{i,j=1,2,....,I} \qquad (9)$$

Üçgen asimetri modeli McCullagh (1978) tarafından tanıtılan koşullu simetri modeline (KS) eşittir.[8] Koşullu simetri modeline ait hipotez Eşitlik (10)' da verildiği gibidir;

$$H_{KS} : p_{ij} = \gamma p_{ji} \quad \text{i<j} \qquad (10)$$

Üçgen asimetri modeli altında beklenen sıklıklara ait en çok olabilirlik tahmin eşitlikleri Eşitlik (11)' de verilmiştir ;

$$\hat{m}_{ij} = \begin{cases} \dfrac{\hat{\gamma}}{1+\hat{\gamma}}(x_{ij} + x_{ji}) & \text{i<j} \\ x_{ii} & \text{i=j} \\ \dfrac{1}{1+\hat{\gamma}}(x_{ij} + x_{ji}) & \text{i>j} \end{cases} \qquad \hat{\gamma} = \frac{\sum_{i<j} x_{ij}}{\sum_{i>j} x_{ij}} \qquad (11)$$

Üçgen asimetri modelinin serbestlik derecesi; $sd_{ÜA} = \frac{(I+1)(I-2)}{2}$ dir. Üçgen asimetri modeline ilişkin parametrelerin gözelere göre dağılımı matris yapısında verilmiştir.

ÜA gösterimiyle ifade edilen bu matrisi oluşturan $ÜA_{ij}$ elemanları tüm (i,j) ler aşağıda verilen eşitlik ile ifade edilir;

$$ÜA_{ij} = \begin{cases} 3 & \text{i<j} \\ 2 & \text{i>j} \\ 1 & \text{i=j} \end{cases} \qquad (12)$$

4×4 karesel olumsallık tablosu için ÜA matris gösterimi aşağıdaki gibidir;

$$ÜA = \begin{bmatrix} 1 & 3 & 3 & 3 \\ 2 & 1 & 3 & 3 \\ 2 & 2 & 1 & 3 \\ 2 & 2 & 2 & 1 \end{bmatrix}$$

Köşegen Asimetri Modeli (KA)

Goodman (1985) tarafından KS modelinin uzantısı olarak tanımlanan köşegen asimetri modeline ait hipotez Eşitlik (13) 'de ifade edildiği gibidir;[7]

$$H_{KA} : p_{ij} = p_{ji}\delta_k \qquad k = j-1 = \pm 1, \pm 2, ..., (I-1) \qquad (13)$$

Eşitlik (13)' de yer alan δ_k parametreleri bir gözlem değerinin k=j-i için, (j,i) gözesi yerine (i,j) gözesinde olmasının odds değerini ifade etmektedir.

Köşegen asimetri modeli altında beklenen sıklıklara ait en çok olabilirlik tahmin eşitlikleri Eşitlik (14)'de verilmiştir ;

$$\hat{m}_{ij} = \begin{cases} \dfrac{\hat{\delta}_{j-i}}{1+\hat{\delta}_{j-i}}(x_{ij} + x_{ji}) & \text{i<j} \\ x_{ii} & \text{i=j} \\ \dfrac{1}{1+\hat{\delta}_{i-j}}(x_{ij} + x_{ji}) & \text{i>j} \end{cases} \qquad (14)$$

$$\hat{\delta}_k = \frac{x^+_{(k)}}{x^-_{(k)}} \qquad x^+_{(k)} = \sum_{t=1}^{I-k} x_{t,t+k} \qquad x^-_{(k)} = \sum_{t=1}^{I-k} x_{t+k,t}$$

Köşegen asimetri modelinin serbestlik derecesi; $sd_{KA} = \frac{(I-1)\,(I-2)}{2}$ dir. Köşegen asimetri modeline ilişkin parametrelerin gözelere göre dağılımı matris yapısında verilmiştir.

KA gösterimiyle ifade edilen bu matrisi oluşturan KA_{ij} elemanları tüm (i,j) ler aşağıda verilen eşitlik ile ifade edilir;

$$KA_{ij} = \begin{cases} |i-j| & \text{i<j} \\ |i-j| + \text{I -1} & \text{i>j} \\ \text{2I-1} & \text{i=j} \end{cases} \quad (15)$$

4×4 karesel olumsallık tablosu için KA matris gösterimi aşağıdaki gibidir;

$$KA = \begin{bmatrix} 7 & 1 & 2 & 3 \\ 4 & 7 & 1 & 2 \\ 5 & 4 & 7 & 1 \\ 6 & 5 & 4 & 7 \end{bmatrix}$$

Sabit Uzaklık Asimetri Modeli (SUA)

Sabit uzaklık asimetri modelinde köşegen asimetri modelindeki parametrelerin logaritmaları $\left[\ln\left(\delta_k\right)\right]$ doğrusal yapıya sahiptir. Modele ait hipotez Eşitlik (16)'da ifade edilmiştir.

$$H_{SUA}: p_{ij} = p_{ji}\delta^{|\text{j-i}|} \qquad \text{j} \geq \text{i} \quad (16)$$

Sabit uzaklık asimetri modeli altında beklenen sıklıkların en çok olabilirlik tahmin eşitlikleri Eşitlik (17)'de verilmiştir;

$$\hat{m}_{ij} = \begin{cases} \dfrac{\hat{\delta}^{j-i}}{1+\hat{\delta}^{j-i}}(x_{ij}+x_{ji}) & \text{i<j} \\ x_{ii} & \text{i=j} \\ \dfrac{1}{1+\hat{\delta}^{j-i}}(x_{ij}+x_{ji}) & \text{i>j} \end{cases} \quad (17)$$

Sabit uzaklık asimetri modelinin serbestlik derecesi; $sd_{SUA} = \frac{(I+1)(I-2)}{2}$ dir. Sabit uzaklık asimetri modeline ilişkin parametrelerin gözelere göre dağılımı matris yapısında verilmiştir. SUA gösterimiyle ifade edilen bu matrisi oluşturan SUA_{ij} elemanları tüm (i,j) ler aşağıda verilen eşitlik ile ifade edilir;

$$SUA_{ij} = \begin{cases} |i-j|+1 & \text{i<j} \\ 1 & \text{i>j} \end{cases} \quad (18)$$

4×4 karesel olumsallık tablosu için SUA matris gösterimi aşağıdaki gibidir;

$$SUA = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 2 & 3 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Odds Asimetri 1 Modeli (OA1)

Üçgen asimetri modelinin uzantısı olarak Tomizawa (1985) tarafından tanımlanan odds asimetri 1 modeline ait hipotez Eşitlik (19)'da ifade edildiği gibidir;[4]

$$H_{OA1}: \frac{p_{ij}}{p_{(i,j+1)}} = \frac{p_{ji}}{p_{(j+1,i)}} \qquad \text{i<j} \quad (19)$$

Odds asimetri 1 modeline göre, olumsallık tablosunun sağ üst üçgeninde i. satır düzeyinde, sütun düzeyinin (j+1) yerine j'de olmasının odds oranı, sol alt üçgenin i. sütun düzeyinde satır düzeyinin (j+1) yerine j'de olmasının odds oranına eşittir.

Odds asimetri 1 modeli altında beklenen sıklıklara ait en çok olabilirlik tahmin eşitlikleri Eşitlik (20)'de verilmiştir.[4]

$$\hat{m}_{ij} = \begin{cases} \dfrac{\left(x_{ij}+x_{ji}\right)b_i}{\left(b_i+c_i\right)} & \text{i<j} \\ x_{ii} & \text{i=j} \\ \dfrac{\left(x_{ij}+x_{ji}\right)c_i}{\left(b_i+c_i\right)} & \text{i>j} \end{cases} \quad (20)$$

Odds Asimetri 1 modelinin serbestlik derecesi $sd_{OA1} = \frac{(I-1)(I-2)}{2}$ dir. Odds asimetri 1 modeline ilişkin parametrelerin gözelere göre dağılımı matris yapısında verilmiştir. OA1 gösterimiyle ifade edilen bu matrisi oluşturan $OA1_{ij}$ elemanları tüm (i,j) ler aşağıda verilen eşitlik ile ifade edilir;

$$(OA1)_{ij} = \begin{cases} i & \text{i<j} \\ (I-1)+j & \text{i>j} \\ 2I-1 & \text{i=j} \end{cases} \quad (21)$$

4×4 karesel olumsallık tablosu için OA1 matris gösterimi aşağıdaki gibidir;

$$OA1 = \begin{bmatrix} 7 & 1 & 1 & 1 \\ 4 & 7 & 2 & 2 \\ 4 & 5 & 7 & 3 \\ 4 & 5 & 6 & 7 \end{bmatrix}$$

Odds Asimetri 2 Modeli (OA2)

Üçgen asimetri modelinin uzantısı olarak Tomizawa (1985) tarafından tanımlanan odds asimetri 2 modeline ait hipotez Eşitlik (22)'de ifade edildiği gibidir;[4]

$$H_{OA2}: \frac{p_{(i-1,j)}}{p_{ij}} = \frac{p_{(j,i-1)}}{p_{ji}} \qquad \text{i<j} \qquad (22)$$

Odds asimetri 2 modeline göre, olumsallık tablosunun sağ üst üçgeninde j. sütun düzeyinde, satır düzeyinin i yerine (i-1)'de olmasının odds oranı, sol alt üçgenin j. satır düzeyinde sütun düzeyinin i yerine (i-1)'de olmasının odds oranına eşittir.

Odds asimetri 2 modeli altında beklenen sıklıklara ait en çok olabilirlik tahmin eşitlikleri Eşitlik (23)' de verilmiştir

$$\hat{m}_{ij} = \begin{cases} \dfrac{\left(x_{ij}+x_{ji}\right)b_i}{\left(b_i+c_i\right)} & \text{i<j} \\ x_{ii} & \text{i=j} \\ \dfrac{\left(x_{ij}+x_{ji}\right)c_i}{\left(b_i+c_i\right)} & \text{i>j} \end{cases} \qquad (23)$$

Odds asimetri 2 modeline ait serbestlik derecesi $sd_{OA2} = \frac{(I-1)(I-2)}{2}$ dir. Odds asimetri 2 modeline ilişkin parametrelerin gözelere göre dağılımı matris yapısında verilmiştir. OA2 gösterimiyle ifade edilen bu matrisi oluşturan $\mathrm{OA2}_{ij}$ elemanları tüm (i,j) ler aşağıda verilen eşitlik ile ifade edilir;

$$(OA2)_{ij} = \begin{cases} (2I-j) & \text{i<j} \\ (I+1)-i & \text{i>j} \\ 2I-1 & \text{i=j} \end{cases} \quad (24)$$

4×4 karesel olumsallık tablosu için OA2 matris gösterimi aşağıdaki gibidir;

$$\mathrm{OA2} = \begin{bmatrix} 7 & 6 & 5 & 4 \\ 3 & 7 & 5 & 4 \\ 2 & 2 & 7 & 4 \\ 1 & 1 & 1 & 7 \end{bmatrix}$$

İki Oran Parametre Asimetri Modeli (2OPA)

İki oran parametre asimetri modeline ait hipotez Eşitlik (25)'de verildiği gibidir.

$$H_{2OPA}: p_{ij} = p_{ji}\ \gamma\ \delta^{j-i} \qquad \text{i<j} \qquad (25)$$

2 oran parametre asimetri modeli altında beklenen sıklıklara ait en çok olabilirlik tahminleri eşitlikleri Eşitlik (26)'da verilen kısıtları sağlamalıdır;

$$\begin{aligned} &\hat{m}_{ij} + \hat{m}_{ji} = x_{ij} + x_{ji} \qquad 1 \le \text{i,j} \le I \\ &\sum_{i=1}^{I}\sum_{j=1}^{I} j\hat{m}_{ij} = \sum_{i=1}^{I}\sum_{j=1}^{I} jx_{ij} \\ &\sum_{i<j} \hat{m}_{ij}\left\{I-2\left(j-i\right)\right\} = \sum_{i<j} x_{ij}\left\{I-2\left(j-i\right)\right\} \end{aligned} \qquad (26)$$

2 oran parametre asimetri modelinin serbestlik derecesi; $sd_{2OPA} = \frac{I^2-I-4}{2}$ dir Asimetri modelleri tam simetri modelini temel aldıkları için bu modelle birlikte ifade edilmelidir. Asimetri modellerinin bilgisayar ortamındaki çözümü için özetlenen Tablo 2'de matris gösterimleri "TERİMLER" başlığı altında gösterilmiştir

Çarpık Simetri Modelleri

Yarı simetri modelinden sapmaları ifade eden çarpık simetri modelleri Yamagushi (1990) tarafından tanımlanmıştır.[9] Çarpık simetri modelleri yarı simetri modelini temel alırlar. Bu nedenle yarı simetri modeli boş çarpık simetri modeli olarak adlandırılır.

TABLO 2: Asimetri modellerinin bilgisayar ortamındaki çözümüne ait terimler.

Model	Terimler
Üçgen Asimetri	TS + ÜA
Köşegen Asimetri	TS + KA
Sabit Uzaklık Asimetri	TS + SUA
Odds Asimetri 1	TS + OA1
Odds Asimetri 2	TS + OA2
2 Oran Parametre Asimetri	TS + ÜA + SUA

(TS:tam simetri, ÜA:üçgen asimetri, KA:köşegen asimetri, SUA: sabit uzaklık asimetri, OA1: odds asimetri 1, OA2:odds asimetri 2).

$\Omega_{(ij,st)}$, $i.$ ve $j.$ satırlar ile, $s.$ ve $t.$ sütunlardan oluşan 2×2 alt tablolar için odds oranı olarak tanımlansın. Bu durumda,

$$\Theta_{(ij,st)}=\frac{p_{is}p_{jt}}{p_{js}p_{it}} \text{ iken,}$$

$$\Omega_{(ij,st)}=\frac{\Theta_{(ij,st)}}{\Theta_{(sj,ij)}} \text{ dir.} \tag{27}$$

$\Omega_{(ij,st)}$ tahminleri bazı modeller altında m_{ij} beklenen sıklıklarından aşağıdaki gibi elde edilebilir.

$$\hat{\Omega}_{(ij,st)}=\frac{m_{is}\times m_{jk}\times m_{ti}\times m_{sj}}{m_{si}\times m_{tj}\times m_{it}\times m_{js}} \tag{28}$$

$s=j$ ve $t=k$ olduğu durum için Eşitlik (4.20) aşağıdaki gibi elde edilir.

$$\hat{\Omega}_{(ij,jk)}=\frac{m_{ij}\times m_{jk}\times m_{ki}\times m_{jj}}{m_{ji}\times m_{kj}\times m_{ik}\times m_{jj}}=\frac{m_{ij}\times m_{jk}\times m_{ki}}{m_{ji}\times m_{kj}\times m_{ik}},\quad i<j<k \tag{29}$$

$\hat{\Omega}_{(ij,jk)}=e^{\Phi_{(ij,jk)}}$ olmak üzere boş çarpık simetri modeli (yarı simetri) modeli altında $\Omega_{(ij,jk)}=1$ ve $\Phi_{(ij,jk)}=0$ dır. Aşağıda ifade edilen çarpık simetri modellerinde $\Omega_{(ij,jk)}\neq 1$ ve $\Phi_{(ij,jk)}\neq 0$ dır.[10]

Tekdüze Çarpık Simetri Modeli (TÇS)

Tekdüze çarpık simetri modeline ait hipotez Eşitlik (30)' da ifade edildiği gibidir;

$$H_{TÇS}: p_{ij}=\gamma\,\alpha_j\,p_{ji} \qquad i<j \tag{30}$$

Tekdüze çarpık simetri modelinin serbestlik derecesi $sd_{TÇSD}=\frac{I(I-3)}{2}$ dir.

Köşegen Parametre Çarpık Simetri Modeli (KPÇS)

Köşegen asimetri modelinden yola çıkarak elde edilen köşegen parametre çarpık simetri modeline ait hipotez $k=j-i$ olmak üzere Eşitlik (31)' de verildiği gibidir;

$$H_{KPÇS}: p_{ij}=\alpha_j\,p_{ji}\delta_k \qquad i<j \tag{31}$$

Köşegen parametre simetri modeline ait serbestlik derecesi $sd_{KPÇS}=\frac{(I-2)(I-3)}{2}$ dir .

Orta Değer Etkili Çarpık Simetri Modeli (OEÇS)

$\theta_{(i<j;k<l)}$ odds oranı olmak üzere orta değer etkili çarpık simetri modeli olumsallık tablosunun sağ üst üçgeninde yer alan odds oranının $\left(\theta_{(i<j;k<l)}\ 1\le i<j<k<l\le I\right)$, sol alt üçgeninde yer alan odds oranına $\left(\theta_{(k<l;i<j)}\right)$ eşitliğini test eder. Orta değer etkili çarpık simetri modeline ait hipotez Eşitlik (32)' de verildiği gibidir;

$$H_{OEÇS}: p_{ij}=\alpha_i\beta_j\delta_{ij} \qquad i<j \tag{32}$$

Orta değer etkili çarpık simetri modelinin serbestlik derecesi $sd_{OEÇS}=\frac{(I-2)(I-3)}{2}$ dir.

Çarpık simetri modelleri yarı simetri modelini temel aldıkları için bu modelle birlikte ifade edilmelidir. Çarpık simetri modellerinin bilgisayar ortamındaki çözümü için özetlenen Tablo 3'te matris gösterimleri "TERİMLER" başlığı altında gösterilmiştir

SAYISAL ÖRNEK

Modellerin uygulaması, Osius (1997) makalesinden alınan 6x6 karesel tablosu üzerinde yapılmıştır. Bu çalışmada önleyici check-up kapsamında smear testi yaptıran kadınlardan alınan örnekler histolijik olarak analiz edilmiş ve sitolojik sınıflandırmaya göre kategorize edilmiştir.[11] Bu sınıflanırmalar:

Pap I	Normal hücre
Pap II	hafif inflame
PapIII D L	hafif displazi,
PapIII D M	orta displazi,
PapIV A S	şiddetli displazi,

TABLO 3: Çarpık simetri modellerinin bilgisayar ortamındaki çözümüne ait terimler	
Model	**Terimler**
Tekdüze Çarpık Simetri Düzey	Satır Değişkeni + Sütun Değişkeni + TS + ÜA
Köşegen Parametre Çarpık Simetri	Satır Değişkeni + Sütun Değişkeni + TS + KA
Ortadeğer Etkili Çarpık Simetri	Satır Değişkeni + Sütun Değişkeni + TS + OA1

(TS:tam simetri, ÜA:üçgen asimetri, KA:köşegen asimetri, OA1:odds asimetri 1).

PapIV AC karsinom in situ,

PapIV B başlangıç inflitrasyon karsinom in situ,

Pap V epitel doku karsinom veya adenoid karsinom

PAP III, IV ve V kategorileri şüpheli kanser olabileceği için aynı zamanda biyopsi önerilmektedir. Dolayısıyla sitolojik sınıflama ile en son elde edilen patalojik sınıflandırma arasında doğruluk bakımından bir karşılaştırma yapılmak istenmektedir. Sitolojik ve patalojik sınıflandırmalar Tablo 4'te verilmiştir. PAP III ve daha üstü durumlarda biyopsi istendiği için PAP I ve PAP II sınıfları tabloya dahil edilmemiştir.

Sitolojik ve patolojik sınıflamalar arasındaki uyuma bakıldığında Kappa katsayısı %74 hesaplanmıştır (P<0.01). Buradan iki teşhis yöntemi arasında mükemmel bir uyumdan söz edilemeyeceği görülür. Kalite hedeflerinde sitolojik tanı ile patalojik tanı arasındaki uyumun %100'e yakın olması beklendiğinden, bu durumda köşegen dışı değerlerde de anlamlı bir yorum olup olmayacağı araştırılmalıdır. Karesel tablo çözümlemeleri köşegen dışı elemanları dikkate aldığından bu tablolar için geliştirilen modellerden bazılarının çözümlemeleri verilmiştir. Tablo 4'de verilen karesel tabloya 3.Bölümde verilen simetri modelleri uygulanmış ve modellere ilişkin serbestlik dereceleri (sd), olabilirlik oran istatistikleri (G^2) ve P değerleri elde edilmiştir ve Tablo 5'te özetlenmiştir.

Veriler simetri yapısını ifade eden modellerden TS ve YS modeline uyum sağlamaktadır. Fakat veri için simetri yapısı sağlandı denilemez. Bunun nedeni marjinal homojenlik modeline uyum olmamasıdır. YS modeline yüksek ölçüde uyumlu bulunmasının nedeni ise verinin, daha sonra tartışılacak olunan YS modelini temel alan çarpık simetri modellerine uyum gösterecek olmasıdır. Bu durumda verilere Asimetri ve Çarpık Simetri modelleri uygulanmıştır (Tablo 6, 7).

Asimetri ve çarpık simetri alt modellerinin hepsi veriye uyum sağlamıştır. En küçük AIC değerine sahip olan model ise TÇS modelidir. Dolayısıyla veriye en iyi uyumu sağlayan model 9 serbestlik derecesi ve G^2=1,791 ile TÇS modelidir. TÇS modeli beklenen sıklıklar elde edilmiştir (Tablo 8).

TÇS modeli altında $\log(\hat{\gamma}) = 0.292$ yani $\hat{\gamma}$=1,34'tür. Bu durumda,

$$m_{ij} m_{jk} m_{ki} = \hat{\gamma} m_{ji} m_{kj} m_{ik} \quad , \quad (1 \leq i < j < k \leq I)$$

$\frac{m_{ij} m_{jk} m_{ki}}{m_{ji} m_{kj} m_{ik}} = \hat{\gamma}$ eşitliğinden odds oranlarına geçilirse

$$\gamma = \hat{\Omega}_{(12,23)} = \hat{\Omega}_{(13,34)} = \hat{\Omega}_{(14,45)} = \hat{\Omega}_{(15,56)} =$$
$$\hat{\Omega}_{(23,34)} = \hat{\Omega}_{(24,45)} = \hat{\Omega}_{(25,56)} = \hat{\Omega}_{(34,45)} = \hat{\Omega}_{(35,56)} = \hat{\Omega}_{(45,56)} = 1.34$$

olduğu görülür. TÇS modelinde $\hat{\gamma}$ > 1 olduğu için sitolojik tanı yönteminin tanımladığı sınıflandırmaları patolojik sınıflandırmanın derece

TABLO 4: Sitolojik ve patalojik sınıflandırma.

Sitolojik Sınıflandırma	**Patolojik Sınıflandırma**						**Toplam**
	1	**2**	**3**	**4**	**5**	**6**	
1	12	5	0	1	0	1	**19**
2	6	32	13	7	0	2	**60**
3	0	8	43	18	1	1	**71**
4	0	0	9	36	2	1	**48**
5	0	0	1	6	0	1	**8**
6	0	1	1	7	1	64	**74**
Toplam	**18**	**47**	**67**	**75**	**3**	**70**	**280**

TABLO 5: Simetri yapısını ifade eden modellerinin sd, G^2 ve P değerleri.

Model	Sd	G^2	P-değeri
TS	15	18,909	0,218
YS	10	2,207	0,994
MH	5	16,81	0,04*

*P<0,05.

TABLO 6: Asimetri alt modellerinin sd, G^2 ve P değerleri

Asimetri Alt Modelleri	sd	G^2	P-değeri	AIC
ÜA	14	17,550	**0,228**	-10,45
KA	10	17,240	**0,069**	-2,76
SUA	14	17,765	**0,218**	-10,235
OA1	10	4,586	**0,917**	-15,414
OA2	10	7,704	**0,658**	-12,296
2OPA	13	17,532	**0,176**	-8,468

TABLO 7: Çarpık simetri alt modellerinin sd, G^2 ve P değerleri.

Çarpık Simetri Alt Modelleri	sd	G^2	P-değeri	AIC
TÇS	9	1,791	**0,994**	**-16,209**
KPÇS	6	1,706	**0,945**	-10,294
OEÇS	6	1,109	**0,981**	-10,891

lendirmeyi daha üst sınıfa koyduğu söylenebilir. Yani patalojik teşhis daha hassas bir sınıflandırma yapmıştır ve sitolojik tanıdan kaçabilecek gözlemleri yakalamada daha başarılıdır.

SONUÇ VE TARTIŞMA

Aynı hastada aynı dokudan alınmış örneklerde gerçekleştirilmiş sitolojik ve patolojik incelemelere ilişkin tanılar arasındaki uyumu değerlendirmek üzere geliştirilmiş ölçüm araçları kullanılmaktadır. Pratikte aynı hasta ve aynı dokuda sitolojik ve patolojik tanı açısından uyum: Aynı hasta ve aynı dokuda sitolojik ve patolojik tanı açısından uyumlu olan vaka sayısı/Aynı hasta ve aynı dokuda sitolojik ve patolojik inceleme yapılmış vaka sayısı) x 100 formülunden hesaplanmaktadır. Bu çalışmada sitolojik sınıflama ile patalojik sınıflandırma arasındaki uyuma ilişkin oluşturulmuş bir 6x6 karesel tablo üzerinde modeller incelenmiştir. Öncelikle karesel olumsallık tablolarında simetri yapısını ifade eden modeller ile araştırmacının denemesi gereken simetrik olmayan modeller tanıtılmıştır. Simetri yapısının bozulduğu çözümlemelerle belirlendiği için veri kümesine simetri yapısına uyum sağlanmadığı taktirde kullanılan modellerin uygulaması gösterilmiştir.

Aktaş ve Saraçbaşı (2009) aynı veri üzerinde uyumsuzluk modellerinin çözümlemesini yapmışlardır. Bu modeller varsayımı altında hesaplanan odds oranlarına göre sitolojik tanı yönteminin tanımladığı sınıflandırmaları patolojik sınıflandırmanın derecelendirmeyi bir üst dereceye koyduğu gösterilmiştir. Osius (1997) sitolojik ve patalojik sınıflandırma verisini 1972-1983 arası için 6 ardışık tabloda ifade etmiş ve her iki yöntem arasındaki uyumda yıllara göre değişim olup olmadığını incelemiştir. Yıllar ilerledikçe her iki teşhis yöntemi arasındaki uyumun arttığını göstermiştir.

Karesel tablo çözümlemelerinde çalışmanın amacına göre test edilecek modeller için: TS, YS, MH; ÜA, KA, SUA, OA1, OA2, 2OPA ve TÇS,

TABLO 8: TÇS modeli altında beklenen sıklıklar.

Sistolojik Sınıflandırma	Patolojik Sınıflandırma					
	1	2	3	4	5	6
1	12	6,422	0,379	0,253	0,367	0,580
2	4,578	32	7,723	1,216	0,356	1,126
3	0,621	13,277	43	8,321	1,120	1,161
4	1,247	6,284	18,679	36	6,342	6,448
5	0,633	0,644	0,880	1,658	0,5	1,185
6	0,920	1,874	0,839	1,552	0,815	64

KPÇS, OEÇS modelleri sıralaması uygulanmalıdır. En uygun model beklenen sıklıkları üzerinden ya da parametre tahminleri kullanılarak odds oranları hesaplanmalı ve tablo ayrıntılı olarak yorumlanmalıdır.

Bu çalışmada yanlış bir uygulama olarak her türlü çapraz tablo çözümlemelerinde bağımsızlık çözümlemesi yapmak yerine özel olarak karesel tablolarının çözümlemesinde önerilmiş olan bazı özel modellerin uygulaması gösterilmiştir.

KAYNAKLAR

1. Bishop, Y. M., Fienberg, S., Holland, P. W., Discrete Multivariate Analysis, Theory and Practice. MIT Press. London; 1975. p. 557
2. Agresti, A., An Introduction Categorical Data Analysis. 2nd ed. New Jersey: John Wiley & Sons; 2007. p. 141-142.
3. Lawal H. B., Categorical Data Analysis with SAS and SPSS Aplications. Lawrence Erlbaum Associates Publishers, New Jersey, 2003. p. 449-451
4. Tomizawa, S., Miyamoto, N., Funato, R. Conditional Difference Asymmetry Model for Square Contingency Tables with Nominal Categories. Journal of Applied Statistics,2004; 31 (3): 271-277.
5. Caussinus, H. Contributions a l'analyse Statistique des Tableaux de Correlation. Ann. Fac. Sci. Univ. Toulouse,1965; 29: 77-182.
6. Stuart, A. A Test for Homogeneity of the Marginal Distributions in a Two-way Classification. Biometrika,1955; 42 (3) :412-416.
7. Yamaguchi, K. Some models for the analysis of asymmetric association in square contingency tables with ordered categories . Sociological Methodology, 1990;20, 181–212.
8. McCullagh, P. A class of Parametric Models for the Analysis of Square Contingency Tables with Ordered Categories. Biometrika,1978; 65 (2): 413-418.
9. Yamaguchi, K. Some models for the analysis of asymmetric association in square contingency tables with ordered categories . Sociological Methodology, 1990;20, 181–212.
10. Lawal, H. B. Review of Non-Independence, Asymmetry, Skew-Symmetry and Point-Symmetry Models in the Analysis of Social Mobility Data, Quality&Quantity, 2004;38 (3): 259-289
11. Osius G. Log-Linear Models for Association and Agreement in Stratified Square Contingency Tables. Computational Statistics, 1997; 12(2):311-328.
12. Aktaş, S., Saraçbaşı S., Estimation of symmetric disagreement using a uniform a association model for ordinal agreement data. AStA Advances in Statistical Analysis, 2009; 93:335-343.