

## LOG-LİNEER MODELLER

Log-Linear Modeller kategorik verilerin analizinde sıklıkla kullanılan bir yöntemdir. Daha çok iki veya daha fazla kategorili ikiden fazla değişkenin (çok yönlü kontenjans tabloları) analizinde kullanılır.

Log-Linear Modeller kontenjans tablolarının hücrelerinde yer alan frekansların değişkenlerin kategorilerine ne kadar bağlı olduğunu ortaya çıkaran, kategorik değişkenler arasındaki ilişkileri ve etkileşimleri inceleyen modellerdir.

Log-Linear analiz ki-kare sonuçlarını geliştiren ve ileri düzeyde bilgi veren bir yöntemdir. Tablo boyutu arttıkça ve iç içe faktör yapılarının karmaşıklığı Log-linear analizi ile giderilebilir.

## LOG-LİNEER MODELLER

Log-Lineer Modeller Poisson ve Multinomial dağılımlı verilere uygundur. **Kontenjans tablosundaki sayım verileri (frekanslar) birbirinden bağımsızsa verilerin dağılımı Poisson Dağılımına uygun olduğu varsayılır.**

**Kontenjans tablosundaki sayım verileri (frekanslar) birbirinden bağımsız değilse veriler Multinomial dağılımına uygun olduğu varsayılır.**

Log-Lineer Modeller Ençok Olabilirlik (ML) tahminlerine ve Olabilirlik Oran testlerine dayanmaktadır (Christensen, 1997).

Kategorik verilerin analizi, iki-değişkenli çapraz-sınıflandırmayla başlar. Araştırmacı, iki değişken çifti arasındaki ilişkiyi belirlemeye çalışır. Değişkenlerin istatistiksel olarak bağımsız olup olmadıkları sorusuna, bildiğimiz klasik  $\chi^2$  analiziyle cevap verilebilir. Değişkenler bağımlıysa, yani aralarında istatistiksel olarak önemli bir ilişki varsa, ilişkinin büyüklüğü ya da gücü nedir? Bu sorunun cevabında, sıralanabilen (ordinal) ve sıralanamayan (nominal) kategorik değişkenler için geliştirilen, sadece iki yönlü tablolar için, ilişki ölçütleriyle verebiliriz. Son olarak, iki değişken çifti arasında gözlemlenen ilişki, üç ya da daha fazla değişkenin içerildiği daha kompleks bir ilişkinin parçası olabilir mi? Bu sorunun cevabını ancak kategorik değişkenlerin çok değişkenli analiziyle verebiliriz Reynolds(84).

# Multinomial Dağılım

Multinomial dağılım ikiden fazla ( $k$  tane) mümkün sonuçlu deneyler için uygulanan kesikli bir dağılımdır. Binom dağılımının genelleştirilmiş şeklidir.  $n_i$ ,  $i$ . mümkün sonucun kaç kez meydana geldiğini gösterir ( $i = 1, 2, \dots, k$ ).  $p_i$  her denemede  $i$ . olayın meydana gelme olasılığıdır ve

$\sum_{i=1}^k p_i = 1$ 'dir. Bu durumda  $(n_1, n_2, \dots, n_k)$ ,  $(n, p_1, \dots, p_k)$  parametrelili bir Multinomial dağılım gösterir.

$$(n_1, n_2, \dots, n_k) \sim \text{Mult}(n, p_1, \dots, p_k)$$

# Multinomial Dağılım

Dağılımın olasılık fonksiyonu  $x_i \geq 0$ ,  $x_1 + \dots + x_k = n$  koşulları altında

$$p(n_1 = x_1, \dots, n_k = x_k) = \frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k} \quad (1.5)$$

fonksiyonu ile gösterilir.

Genel olarak her bir mümkün sonuç Binom dağılımı gösterir.

$$n_i \sim \text{Bin}(n, p_i)$$

Dolayısıyla  $n_i$ 'lerin beklenen değerleri ve varyansları Binom dağılımındaki gibi hesaplanır.

$$\begin{aligned} E(n_i) &= np_i \\ \text{Var}(n_i) &= np_i(1 - p_i) \end{aligned} \quad (1.6)$$

**Örnek.** Bir zar 12 kez atılsın. İki kez 1, üç kez 2, bir kez 3, iki kez 4, üç kez 5, bir kez 6 gelmesi olasılığı nedir.

$$A_1 = 2, A_2 = 3, A_3 = 1, A_4 = 2, A_5 = 3, A_6 = 1, n = 12 \text{ ve} \\ p_i = \frac{1}{6}, i = 1, \dots, 6$$

$$\begin{aligned} P &= f(2, 3, 1, 2, 3, 1) \\ &= \frac{12!}{2!3!1!2!3!1!} \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^3 \left(\frac{1}{6}\right)^1 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^3 \left(\frac{1}{6}\right)^1 \\ &= \frac{11!}{12} \left(\frac{1}{6}\right)^{12} \end{aligned}$$

## İki Yönlü Kontenjans Tablolarında Log-Linear Modeller

### -Bağımsız Model (Independence Model)

$I \times J$  boyutlu kontenjans tablolarında değişkenlerin bağımsız olması durumunda ortak olasılıklar  $(p_{ij})$  i.satırın ve j. Sütunun marjinal olasılıklarının çarpımına eşittir.

$$p_{ij} = p_i \cdot p_j \quad i = 1, 2, \dots, I \quad j = 1, 2, \dots, J$$

$$\hat{f}_{ij} = n \times p_i \times p_j \quad \ln \hat{f}_{ij} = \ln n + \ln p_i + \ln p_j$$

Kontenjans tablosundaki satır değişkeni (X) ve sütun değişkeni (Y) olarak tanımlanırsa model aşağıdaki gibi olur.

$$\ln \hat{f}_{ij} = \lambda + \lambda_i^X + \lambda_j^Y$$

$\lambda_i^X$  : i.satırdaki sınıflamanın X değişkenindeki etkisi

$\lambda_j^Y$  : j.sütun sınıflamanın Y değişkenindeki etkisi

Yukarıdaki modelin  $1+(I-1)+(J-1)=I+J-1$  tane bağımsız parametresi bulunmaktadır.

H0: X ve Y değişkenleri arasında bağımlılık yoktur (bağımsızdırlar)

H1: X ve Y değişkenleri bağımlıdır.

Tahmin edilen Log-Linear Modelin uygunluğu Pearson Ki-Kare ve Olabilirlik Oran (G2) testleri kullanılarak test edilir.

## İki Yönlü Kontenjans Tablolarında Log-Linear Modeller

### -Doymuş Model (Saturated Model)

Bağımsız modelin değişkenler arası bağımlılığı da (etkileşim) içeren modele Doymuş model denir.

$$\ln \hat{f}_{ij} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_{ij}^{XY}$$

$$\sum_i \lambda_i^X = \sum_j \lambda_j^Y = 0 \quad \sum_i \lambda_{ij}^{XY} = \sum_j \lambda_{ij}^{XY} = 0$$

$\lambda_{ij}^{XY}$  : iki yönlü ilişki terimi olup, X ile Y değişkenlerinin bağımsızlıktan ne kadar saptıklarını gösterir.

Doymuş modelin  $1+(I-1)+(J-1)+(I-1)(J-1)$  tane parametresi vardır.

# Üç Yönlü Kontenjans Tablolarında Log-Lineer Modeller

## - Koşullu Bağımsızlık Modeli

Üç yönlü kontenjans tablolarında kısmi ilişkilerin yer aldığı Log-Lineer Modeller kullanılır.

$$\ln \hat{f}_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}$$

$\lambda_{ik}^{XZ}$  : Y sabit tutulduğunda X ile Z arasındaki ilişki

$\lambda_{jk}^{YZ}$  : X sabit tutulduğunda Y ile Z arasındaki ilişki

$\lambda_{ij}^{XY} = 0$  kabul edildiğinden modelde yer almamıştır.



# Üç Yönlü Kontenjans Tablolarında Log-Lineer Modeller

- **Karşılıklı Bağımsızlık Modeli (Mutual Independence Model)**  
Model sadece tek yönlü ilişki terimlerini içerir.

$$\ln \hat{f}_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z$$

# Üç Yönlü Kontenjans Tablolarında Log-Lineer Modeller

## - Homojen İlişki Modeli

Her üç değişkenin birbiriyle koşullu bağımlılık gösterdiği modeldir.

$$\ln \hat{f}_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ}$$

# Üç Yönlü Kontenjans Tablolarında Log-Lineer Modeller

## - Doymuş Model (Saturated Model)

Her üç değişkenin birbiriyle koşullu bağımlılık gösterdiği modeldir.

$$\ln \hat{f}_{ijk} = \lambda + \lambda_i^X + \lambda_j^Y + \lambda_k^Z + \lambda_{ij}^{XY} + \lambda_{ik}^{XZ} + \lambda_{jk}^{YZ} + \lambda_{ijk}^{XYZ}$$

# Log-Linear Modelin Uygunluk Testi

Seçilen Log-Linear modelin ne kadar uygun olduğu Ki-Kare uygunluk testi ile yapılır. Seçilen modelle tahmin edilen (beklenen) frekanslar gözlenen frekanslarla karşılaştırılır.

H0: Seçilen model uygundur. (Seçilen modelin doymuş modelden farkı yoktur)

H1: Model uygun değildir.

Uygunluğun araştırılmasında Olabilirlik Oran ( $G^2$ ) ve Pearson Ki – Kare istatistikleri kullanılır.

$$G^2 = 2 \sum n_{ijk} \log \left( \frac{f_{ijk}}{\hat{f}_{ijk}} \right)$$

$$\chi^2 = \sum \frac{(f_{ijk} - \hat{f}_{ijk})^2}{\hat{f}_{ijk}}$$

Serbestlik derecesi (sd) tablodaki hücre sayısı ile modeldeki indirgenmiş parametre sayısı arasındaki farka eşittir. Modelin parametre sayısı arttıkça sd. azalır. Doymuş bir modelin sd sıfırdır. Hesap değeri tablo değerinden küçük çıkarsa H0 red edilemez yani seçilen model uygundur denilir.

# Log-Linear Modelde Artıkların İncelenmesi

Tablodaki frekanslar ile uygun model ile tahmin edilen frekanslar arasındaki farklar (artıklar-residuals) modelin ne kadar iyi bir model olduğunu gösterir. Az hücreli kontenjans tablolarında standardize artıkların mutlak değerce 2'den büyük olması; çok hücreli tablolarda ise 3'den büyük olması modelin uygunluğunun zayıf olduğunu gösterir (Agresti, 1996).

Artıkların normal dağılıp dağılmadıklarının incelenmesinde modelin uygunluğu hakkında fikir verir.

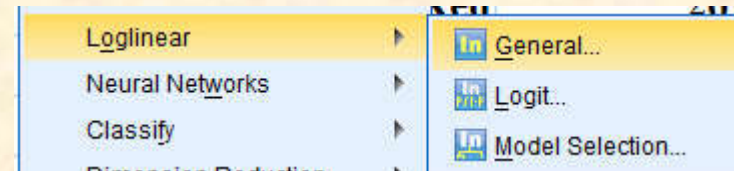
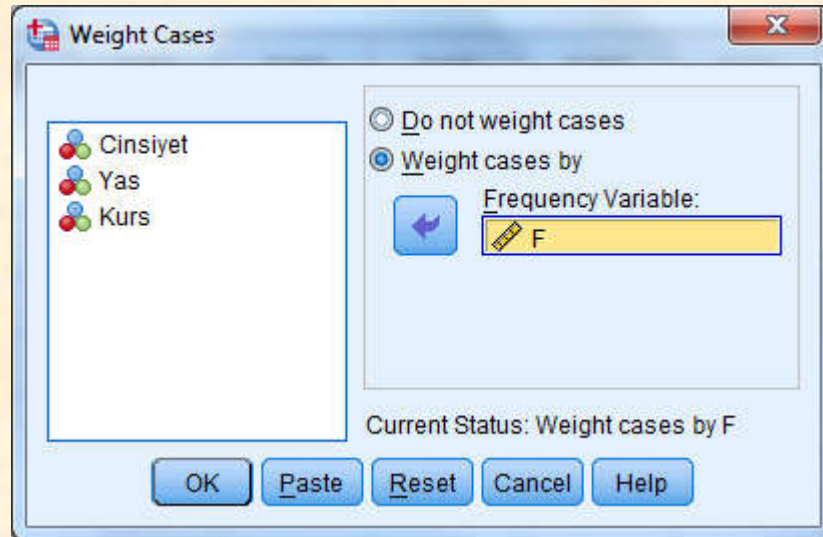
**Örnek.** Bir üniversitede bir kursa kabul edilebilmenin öğrencilerin yaş ve cinsiyete ne kadar bağlı olduğu araştırılmış ve aşağıdaki sonuçlar elde edilmiştir. Veriler birbirinden bağımsız olduğu kabul ediliyor.

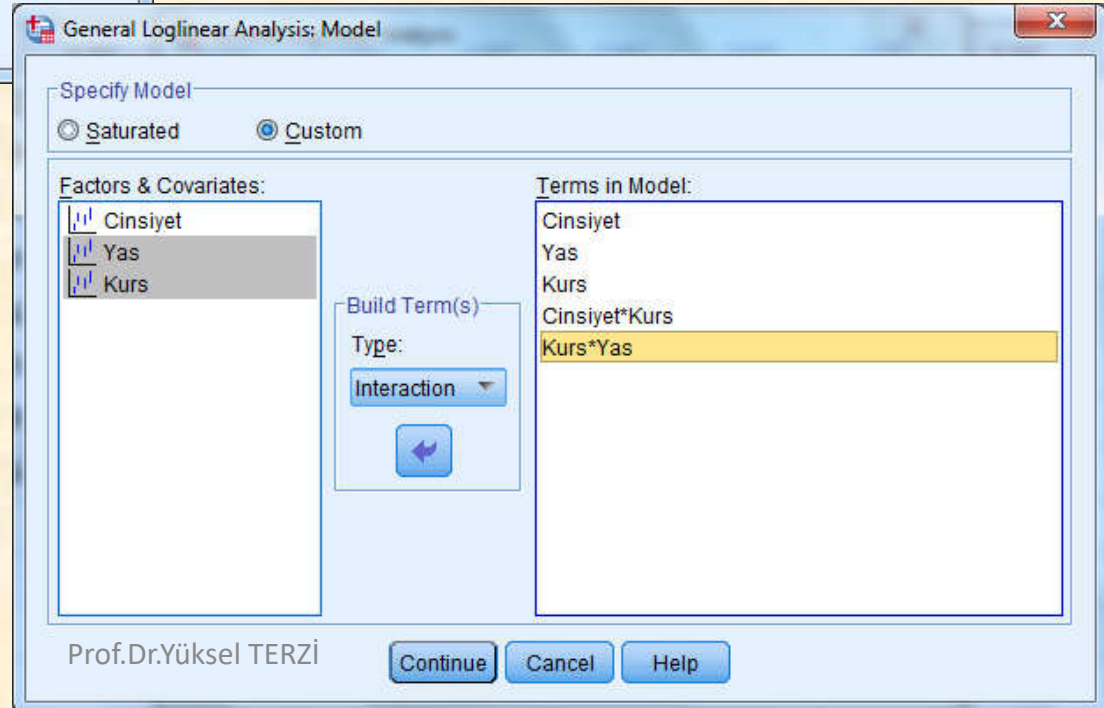
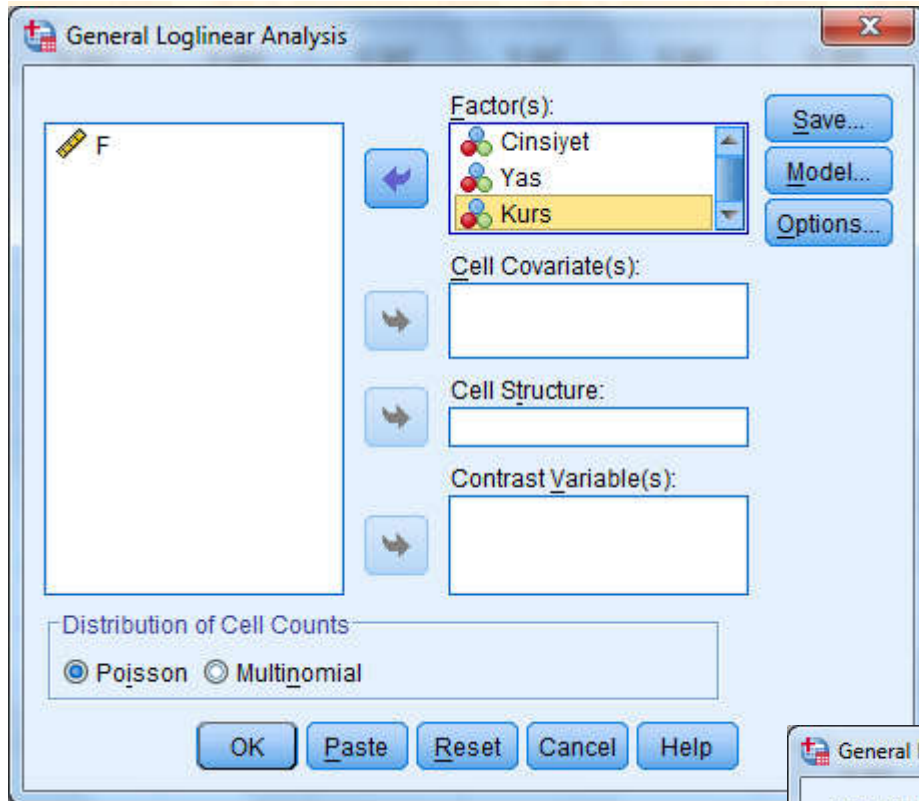
Cinsiyet	Yaş	Kurs	
		Kabul	Red
Erkek	<25	282	189
	≥25	124	269
Kız	<25	24	73
	≥25	19	76

Denemeler sonucunda Koşullu Bağımsızlık modeli en uygun model olarak görülmüştür. Cinsiyet\*Kurs, Cinsiyet\*Yaş, Yaş\*Kurs için Pearson Ki-kare testine bakılır ve önemli olanlar seçilir. Buna göre en uygun model aşağıdaki gibi seçilir.

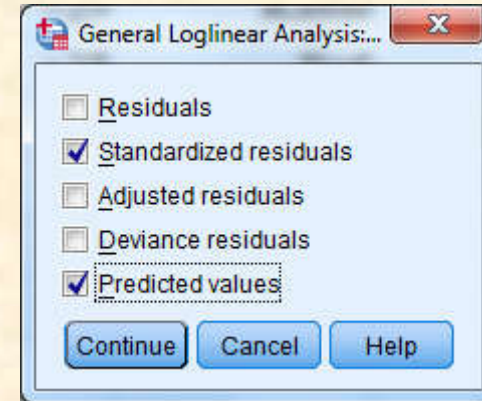
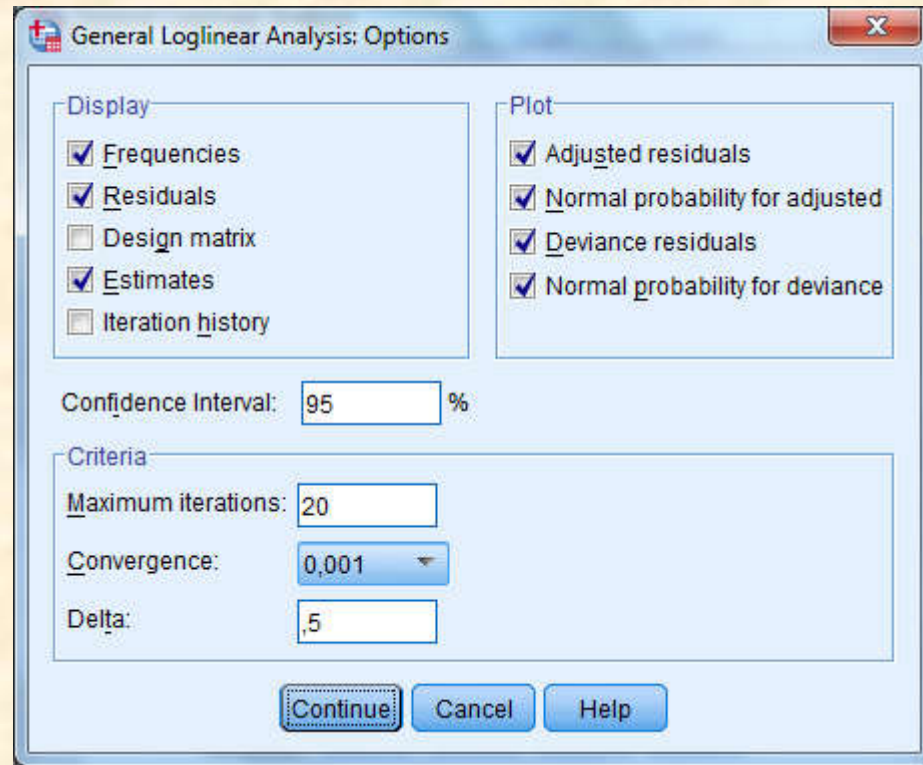
$$\text{Model} = \text{sabit} + \text{Cinsiyet} + \text{Yaş} + \text{Kurs} + (\text{Cinsiyet} * \text{Kurs}) + (\text{Yaş} * \text{Kurs})$$

	Cinsiyet	Yas	Kurs	F
1	Erkek	<25	Kabul	282
2	Erkek	<25	Red	189
3	Erkek	>25	Kabul	124
4	Erkek	>25	Red	269
5	Kız	<25	Kabul	24
6	Kız	<25	Red	73
7	Kız	>25	Kabul	19
8	Kız	>25	Red	76









Goodness-of-Fit Tests<sup>a,b</sup>

	Value	df	Sig.
Likelihood Ratio	5,899	2	,052
Pearson Chi-Square	6,071	2	,048

a. Model: Poisson

b. Design: Constant + Cinsiyet + Yas + Kurs + Cinsiyet \* Kurs + Yas \* Kurs

**H0: Seçilen model uygundur.**

**P=0,052>0,05 olup model uygun bir modeldir.**

**ZRE\_1**

,319

-,618

-,467

,538

-,980

1,083

1,434

-,944

**Modelin uygunluğu standardize artıklarla da kontrol edilebilir. Standardize artıkların hiç biri 2'den büyük olmadığından model uygundur denilebilir.**

Cell Counts and Residuals<sup>a,b</sup>

Cinsiyet	Yas	Kurs	Observed		Expected		Residual	Standardized Residual	Adjusted Residual	Deviance
			Count	%	Count	%				
Erkek	<25	Red	189	17,9%	197,687	18,7%	-8,687	-,618	-1,654	-,622
		Kabul	282	26,7%	276,695	26,2%	5,305	,319	1,826	,318
	>25	Red	269	25,5%	260,313	24,7%	8,687	,538	1,654	,535
		Kabul	124	11,7%	129,305	12,2%	-5,305	-,467	-1,826	-,470
Kız	<25	Red	73	6,9%	64,313	6,1%	8,687	1,083	1,654	1,060
		Kabul	24	2,3%	29,305	2,8%	-5,305	-,980	-1,826	-1,012
	>25	Red	76	7,2%	84,687	8,0%	-8,687	-,944	-1,654	-,961
		Kabul	19	1,8%	13,695	1,3%	5,305	1,434	1,826	1,353

a. Model: Poisson

b. Design: Constant + Cinsiyet + Yas + Kurs + Cinsiyet \* Kurs + Yas \* Kurs

Parameter Estimates<sup>b,c</sup>

Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Constant	2,617	,167	15,634	,000	2,289	2,945
[Cinsiyet = 1]	2,245	,160	14,000	,000	1,931	2,559
[Cinsiyet = 2]	0 <sup>a</sup>	.	.	.	.	.
[Yas = 1]	,761	,101	7,510	,000	,562	,959
[Yas = 2]	0 <sup>a</sup>	.	.	.	.	.
[Kurs = 0]	1,822	,190	9,605	,000	1,450	2,194
[Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.
[Cinsiyet = 1] * [Kurs = 0]	-1,122	,186	-6,032	,000	-1,487	-,758
[Cinsiyet = 1] * [Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.
[Cinsiyet = 2] * [Kurs = 0]	0 <sup>a</sup>	.	.	.	.	.
[Cinsiyet = 2] * [Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.
[Yas = 1] * [Kurs = 0]	-1,036	,130	-7,951	,000	-1,291	-,781
[Yas = 1] * [Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.
[Yas = 2] * [Kurs = 0]	0 <sup>a</sup>	.	.	.	.	.
[Yas = 2] * [Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

b. Model: Poisson

c. Design: Constant + Cinsiyet + Yas + Kurs + Cinsiyet \* Kurs + Yas \* Kurs

Tüm parametreler kısmi ilişki testi sonucunda anlamlı ( $p < 0,05$ ) yani tablodaki veri kümesinin dağılımını açıkladıkları görülmektedir. **Tablodaki tahminler değişkenlerin ilgili kategorileri arasındaki bağımlılığı, bir başka ifade ile bağımsızlıktan ne kadar saptıklarını gösterir.** Standardize edilmiş Z parametre tahminleri hangi kategoriler arasındaki ilişkilerin daha güçlü olduğu hakkında bilgi verir. Tek yönlü ilişki terimleri incelendiğinde Cinsiyet en yüksek Z değerine (14.0) sahip olduğu yani **Kontenjans tablosundaki frekansları belirleyen en önemli faktör Cinsiyettir.**

Parameter Estimates <sup>b,c</sup>						
Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Constant	2,617	,167	15,634	,000	2,289	2,945
[Cinsiyet = 1]	2,245	,160	14,000	,000	1,931	2,559
[Cinsiyet = 2]	0 <sup>a</sup>	.	.	.	.	.
[Yas = 1]	,761	,101	7,510	,000	,562	,959
[Yas = 2]	0 <sup>a</sup>	.	.	.	.	.
[Kurs = 0]	1,822	,190	9,605	,000	1,450	2,194
[Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.
[Cinsiyet = 1] * [Kurs = 0]	-1,122	,186	-6,032	,000	-1,487	-,758
[Cinsiyet = 1] * [Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.
[Cinsiyet = 2] * [Kurs = 0]	0 <sup>a</sup>	.	.	.	.	.
[Cinsiyet = 2] * [Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.
[Yas = 1] * [Kurs = 0]	-1,036	,130	-7,951	,000	-1,291	-,781
[Yas = 1] * [Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.
[Yas = 2] * [Kurs = 0]	0 <sup>a</sup>	.	.	.	.	.
[Yas = 2] * [Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.

Kontenjans tablosundaki deęişkenlerin kategorileri arasındaki ilişkiler modelin iki-yönlü ilişki terimiyle araştırılmaktadır. **Yaş\*Kurs ve Cinsiyet\*Kurs arasında bağımlılık olduğu görülmektedir (p<0,05)**. Bağımlılık Yaş=1, Kurs=0 yani 25 yaş altı ve Kurşa kabul edilmeyenler (Z=-7,951) ve Cinsiyet=1 (Erkek) ve Kurs=0 (red) (Z=-6,032) kategorilerinde olmuştur. Kurşa kabul yaşa bağlıdır ve bu bağımlılık da çok 25 yaş altındakilerle ilgilidir. Kurşa kabul edilme cinsiyete bağlı olduğu ve bunun kurşa red edilen kategorisi ile erkek öğrenciler arasında olduğu görülmektedir.

Parameter Estimates <sup>b,c</sup>						
Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Constant	2,617	,167	15,634	,000	2,289	2,945
[Cinsiyet = 1]	2,245	,160	14,000	,000	1,931	2,559
[Cinsiyet = 2]	0 <sup>a</sup>	.	.	.	.	.
[Yas = 1]	,761	,101	7,510	,000	,562	,959
[Yas = 2]	0 <sup>a</sup>	.	.	.	.	.
[Kurs = 0]	1,822	,190	9,605	,000	1,450	2,194
[Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.
[Cinsiyet = 1] * [Kurs = 0]	-1,122	,186	-6,032	,000	-1,487	-,758
[Cinsiyet = 1] * [Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.
[Cinsiyet = 2] * [Kurs = 0]	0 <sup>a</sup>	.	.	.	.	.
[Cinsiyet = 2] * [Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.
[Yas = 1] * [Kurs = 0]	-1,036	,130	-7,951	,000	-1,291	-,781
[Yas = 1] * [Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.
[Yas = 2] * [Kurs = 0]	0 <sup>a</sup>	.	.	.	.	.
[Yas = 2] * [Kurs = 1]	0 <sup>a</sup>	.	.	.	.	.

**Tabloda indirgenmiş kategorilerin parametre tahminleri sıfır olarak yer almaktadır. Bu kategoriler referans kategori olarak da adlandırılır. Tüm etkilerin toplamının sıfır olmasından dolayı bu referans kategorilerin parametre tahminleri hesaplanabilir.**

İki yönlü ilişki terimlerinin parametre tahminleri InOdds oranına eşittir. Odds oranı parametre tahminlerinin e tabanlı antilogaritması alınarak elde edilir.

Cinsiyet=1 (Erkek) ve Kurs=0 (red) iki yönlü ilişki terimi için tahmin edilen parametrenin (-1,122) antiln değerinin Cinsiyet ve Kurs değişkenlerinden oluşan kontenjans tablosundan elde edilen odds oranına eşit eşit olduğunu gösterelim.

Count		Kurs		Total
		Red	Kabul	
Cinsiyet	Erkek	458	406	864
	Kız	149	43	192
Total		607	449	1056

$$OR=(A*D)/(B*C)=(458*43)/(149*406)=0,325$$

Cins(1)\*Kurs(0) parametre değeri -1,122 dir.

$\text{Exp}(-1,122)=0,325$  olur. Bu değer yorumu ise Erkek öğrencilerin kursa kabul edilmeme olasılığı kız öğrencilerin kabul edilmeme olasılığının 0,325 katıdır. Yani erkeklerin kursa kabul edilme olasılığı kızlara göre 3 kat daha fazladır.

Yas=1 (<25) ve Kurs =0 (red) için tahmin edilen parametre -1,036 dır.

$\text{Exp}(-1,036)=0,354$  bulunur. 25 yaş altı öğrencilerin kursa kabul edilmeme olasılığı 25 yaş üstündekilerin kabul edilmeme olasılığının 0,354 katıdır. Yani 25 yaş altındakilerin kursa kabul edilme olasılığı 25 yaş üstündekilerden yaklaşık 3 kat daha fazladır.

**Correlations of Parameter Estimates<sup>a,b,c</sup>**

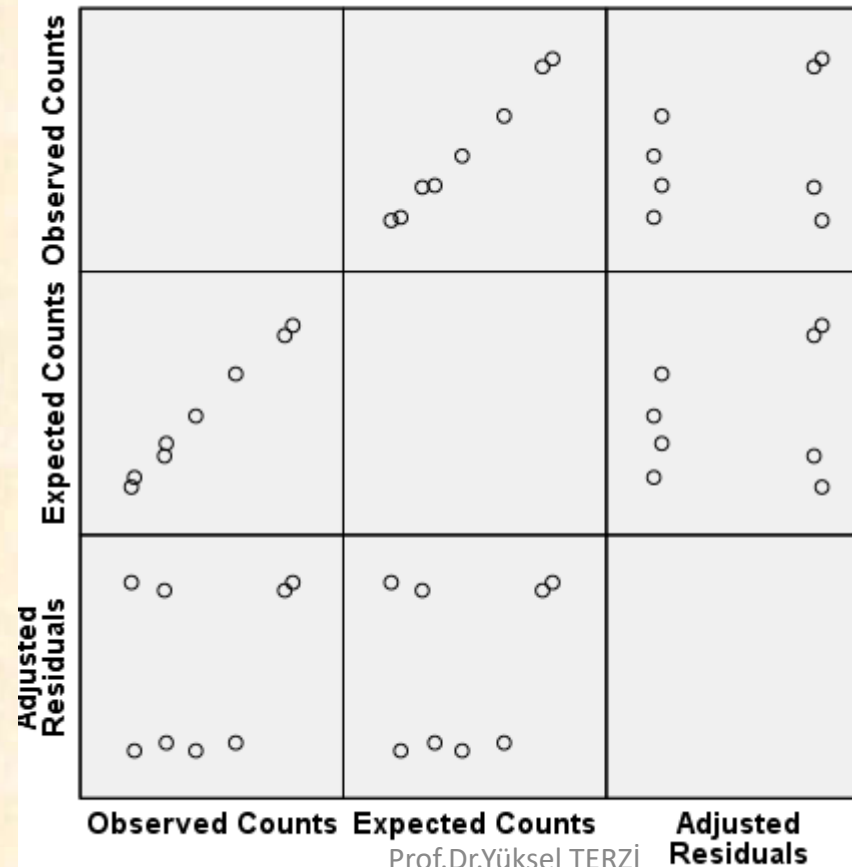
	Constant	[Cinsiyet = 1]	[Yas = 1]	[Kurs = 0]	[Cinsiyet = 1] * [Kurs = 0]	[Yas = 1] * [Kurs = 0]
Constant	1	-,866	-,412	-,882	,747	,321
[Cinsiyet = 1]	-,866	1	,000	,764	-,862	,000
[Yas = 1]	-,412	,000	1	,364	,000	-,777
[Kurs = 0]	-,882	,764	,364	1	-,849	-,400
[Cinsiyet = 1] * [Kurs = 0]	,747	-,862	,000	-,849	1	,000
[Yas = 1] * [Kurs = 0]	,321	,000	-,777	-,400	,000	1

a. Model: Poisson

b. Design: Constant + Cinsiyet + Yas + Kurs + Cinsiyet \* Kurs + Yas \* Kurs

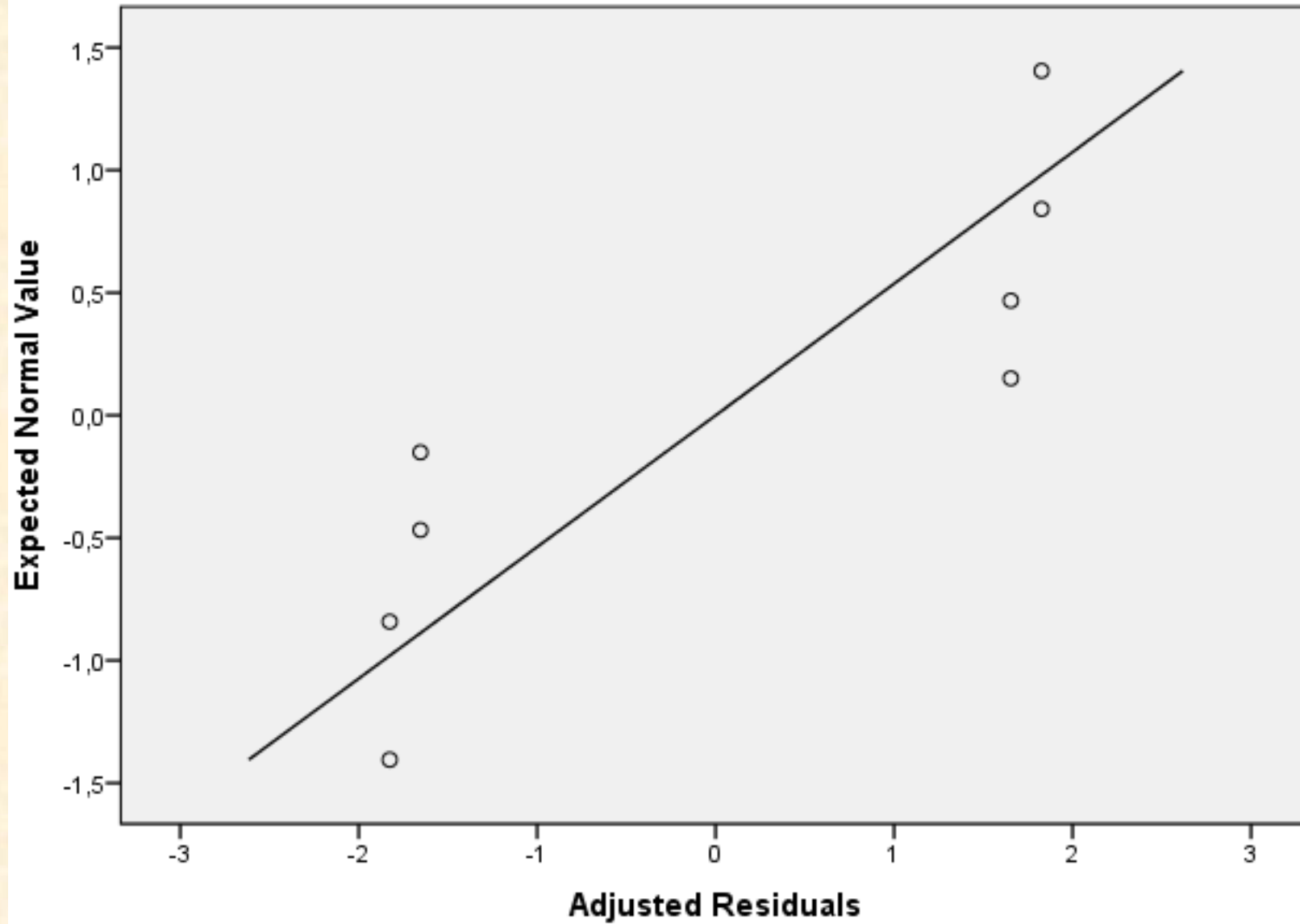
c. Redundant p

**Poisson Model**



Analysis weighted by F

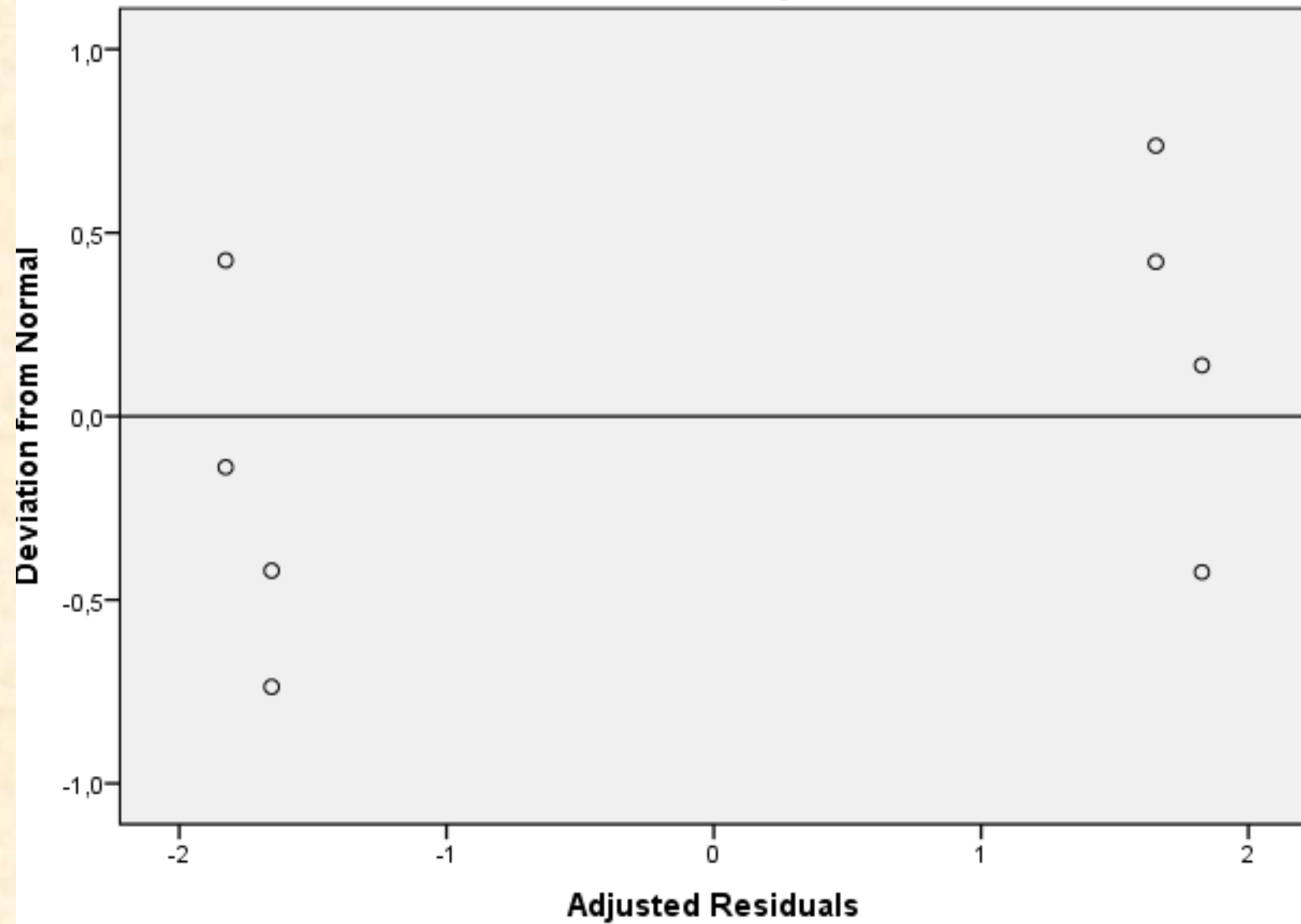
Normal Q-Q Plot of Adjusted Residuals



Analysis weighted by F  
Prof.Dr.Yüksel TERZİ

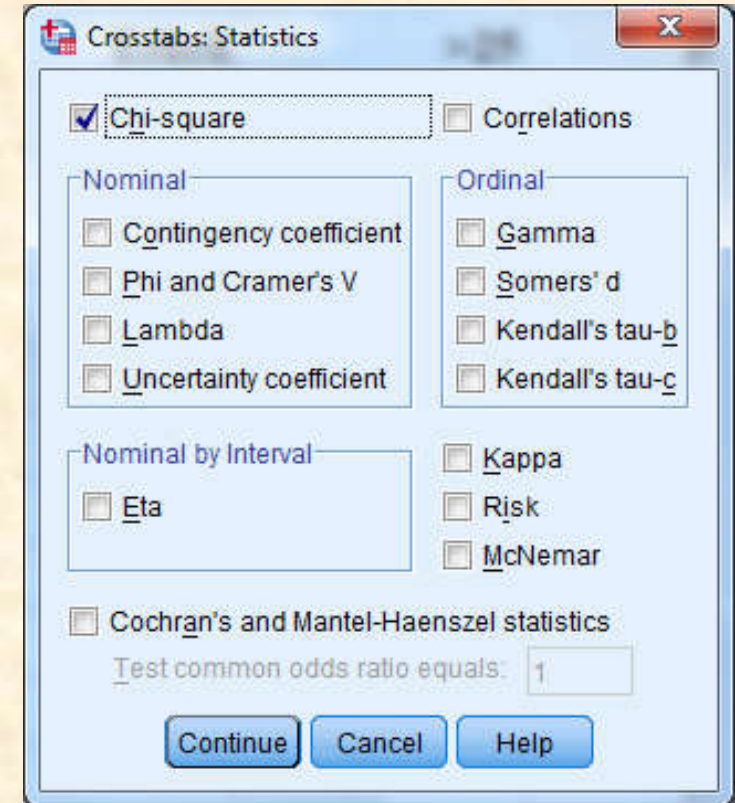
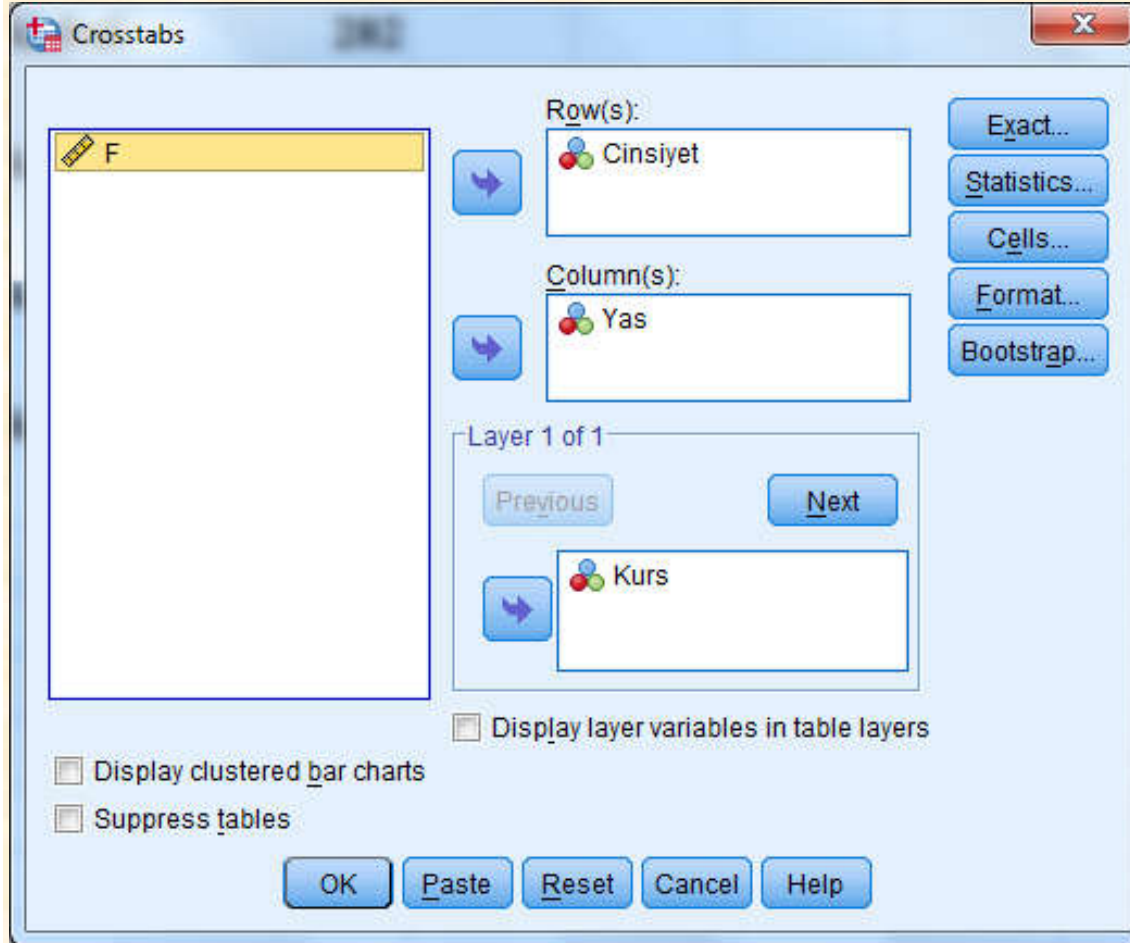


**Detrended Normal Q-Q Plot of Adjusted Residuals**



Analysis weighted by F

Model koşullu bağımsızlık modeli olduğundan Kurs değişkeni kontrol altında tutulduğunda Cinsiyet ve Yaş değişkenleri bağımsızdır.



**Cinsiyet \* Yas \* Kurs Crosstabulation**

Count

			Yas		Total
			<25	>25	
Red	Cinsiyet	Erkek	189	269	458
		Kiz	73	76	149
	Total		262	345	607
Kabul	Cinsiyet	Erkek	282	124	406
		Kiz	24	19	43
	Total		306	143	449
Total	Cinsiyet	Erkek	471	393	864
		Kiz	97	95	192
	Total		568	488	1056

**Chi-Square Tests**

Kurs		Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Red	Pearson Chi-Square	2,736 <sup>c</sup>	1	,098	,106	,060
	Continuity Correction <sup>b</sup>	2,430	1	,119		
	Likelihood Ratio	2,721	1	,099		
	Fisher's Exact Test					
	Linear-by-Linear Association	2,732	1	,098		
	N of Valid Cases	607				
Kabul	Pearson Chi-Square	3,335 <sup>d</sup>	1	,068	,084	,051
	Continuity Correction <sup>b</sup>	2,736	1	,098		
	Likelihood Ratio	3,178	1	,075		
	Fisher's Exact Test					
	Linear-by-Linear Association	3,327	1	,068		
	N of Valid Cases	449				
Total	Pearson Chi-Square	1,008 <sup>a</sup>	1	,315	,337	,178
	Continuity Correction <sup>b</sup>	,853	1	,356		
	Likelihood Ratio	1,006	1	,316		
	Fisher's Exact Test					
	Linear-by-Linear Association	1,007	1	,316		
	N of Valid Cases	1056				

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 88,73.

b. Computed only for a 2x2 table

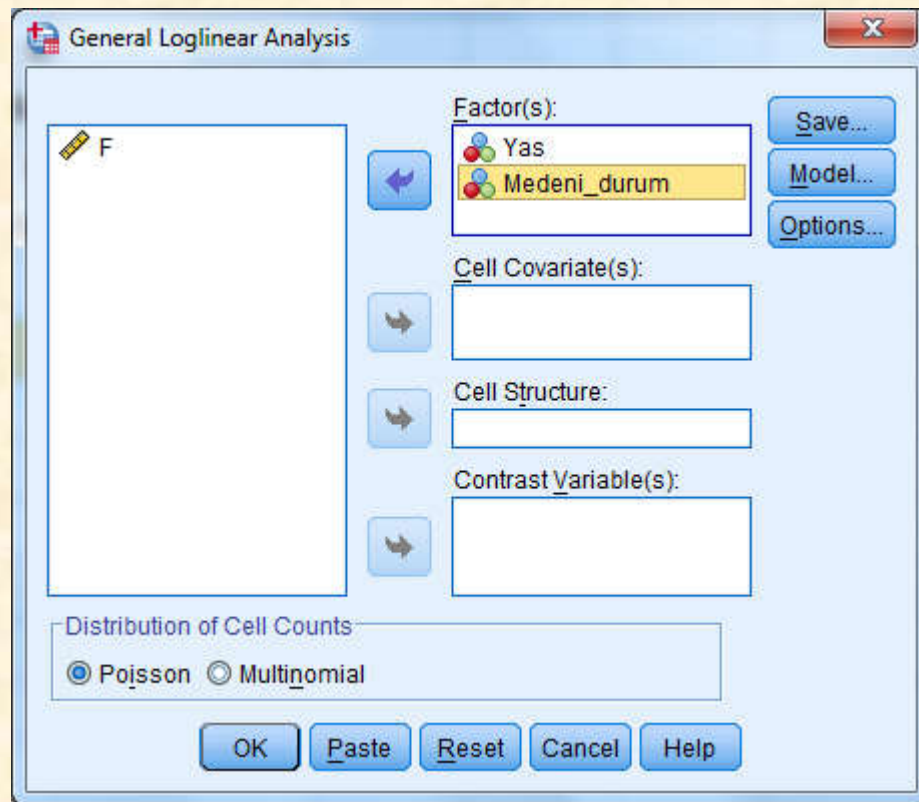
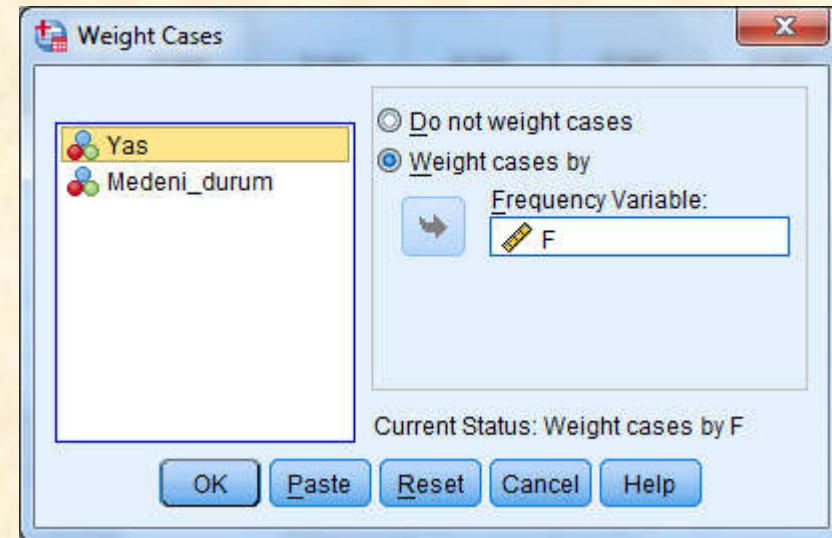
c. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 64,31.

d. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 13,69.

**Örnek.** 208 kişi üzerinde yapılan bir çalışmada yaş kategorileri ile medeni durum arasındaki ilişki araştırılmış ve aşağıdaki sonuçlar elde edilmiştir. Yaş ile medeni durum arasındaki etkileşim önemli midir?

Yaş	Medeni Durum	
	Bekar	Evli
<30	6	86
>=30	8	108

	Yas	Medeni_duru m	F
1	<30	Bekar	6
2	<30	Evli	86
3	>=30	Bekar	8
4	>=30	Evli	108



# UYGUNLUK ANALİZİ (CORRESPONDENCE ANALYSIS)

Uygunluk analizi (UA) veri yapısının ortaya çıkarılmasında kullanılan ve herhangi bir dağılıma bağlı olmaksızın uygulanan çok değişkenli bir istatistiktir. Uygunluk analizinin uygulanabilmesi için verilerin negatif olmaması gerekir (Greenacre, 1984).

UA bir kontenjans tablosundaki satır ve sütunlar arasındaki ilişkileri az boyutlu bir grafik üzerinde gösterimini sağlar. Veri kümesinde sapan değerlerin (outlier) yer alması ve boyut sayısına karar verilmesi analizin güçlükleri olarak kabul edilir. Grafikte benzer dağılımlara sahip kategoriler birbirine yakın noktalar olarak grafik üzerinde yer alırken, dağılımları benzemeyenler grafik üzerinde birbirinden uzak konumda bulunurlar. Grafiğin oluşturulmasında önce satır ve sütun profilleri ayrı ayrı oluşturulur. Daha sonra bu noktalar tek bir grafikte bir araya getirilir (Clausen, 1998).

**Grafik gösterimi değişkenler ve kategoriler arasındaki karmaşık ilişkilerin açıklanmasında kolaylık sağlar.**

# UYGUNLUK ANALİZİ (CORRESPONDENCE ANALYSIS)

UA Log-Linear Modeller ve Gizli Sınıf Analizi (Latent Class Analysis) sonuçlarını tamamlayan bir analizdir. UA değişkenler arasındaki etkileşimlerin anlamlılık testlerini içermez ancak Log-Linear Modeller ile ortaya çıkarılan karmaşık etkileşimlerin yorumlanmasında yardımcı olmaktadır.

**Uygunluk Analizine başlamadan önce Pearson Ki-Kare testinin yapılması önerilmektedir.**

Uygunluk Analizi ikiden fazla değişken içeren çok yönlü kontenjans tablolarına da uygulanabilmekte ve bu durumda ismi «Çok Değişkenli Uygunluk Analizi» olarak adlandırılır. Çok yönlü kontenjans tabloları iki yönlü kontenjans tablolarına dönüştürülür. Örneğin A,B,C değişkenlerin yer aldığı üç yönlü bir kontenjans tablosu ile (A,B), (A,C) ve (B,C) gibi yeni değişkenler elde edilir ve bu değişkenler **interaktif** değişken olarak adlandırılır.

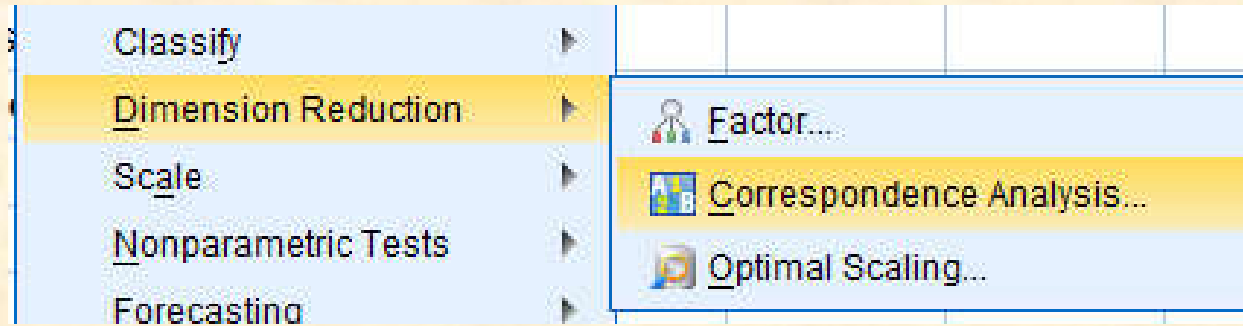
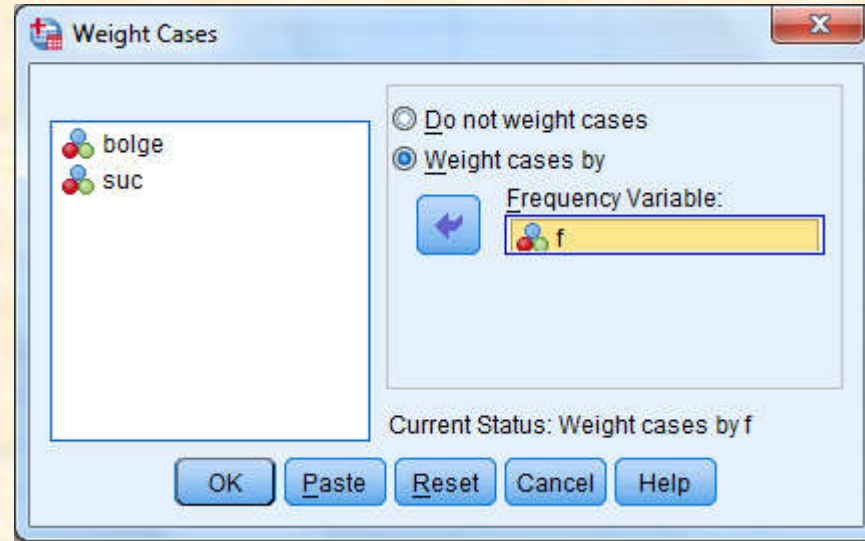
Kontenjans tablosu 3 farklı iki yönlü kontenjans tablosuna dönüştürülebilir (A ve BC, B ve AC, C ve AB). Hangi kontenjans tablosunun temel alınacağına karar vermede **interaktif değişkenin aralarında en az ilişkili değişkenden oluşmasına özen gösterilir. Değişkenlerden biri bağımlı kabul edildiğinde birleşecek değişkenler kendi aralarında en az ilişkili bağımsız değişkenlerden seçilir.**

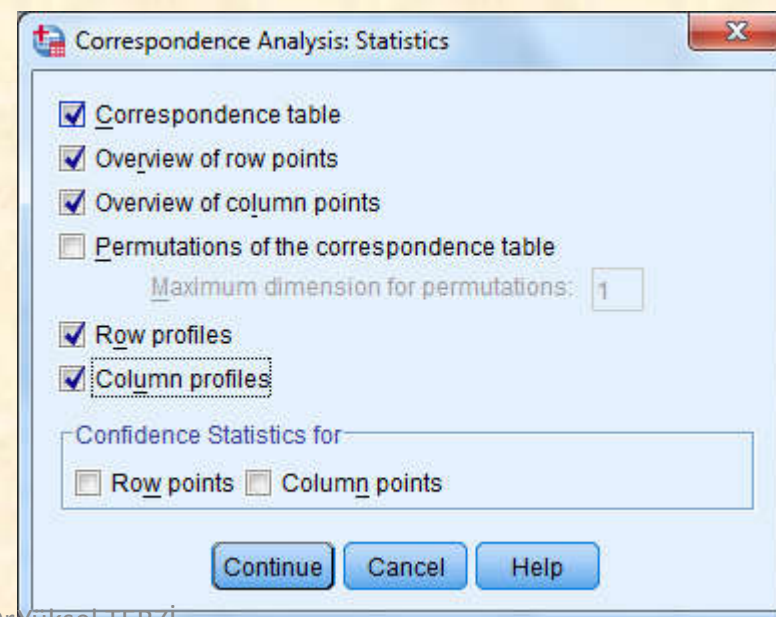
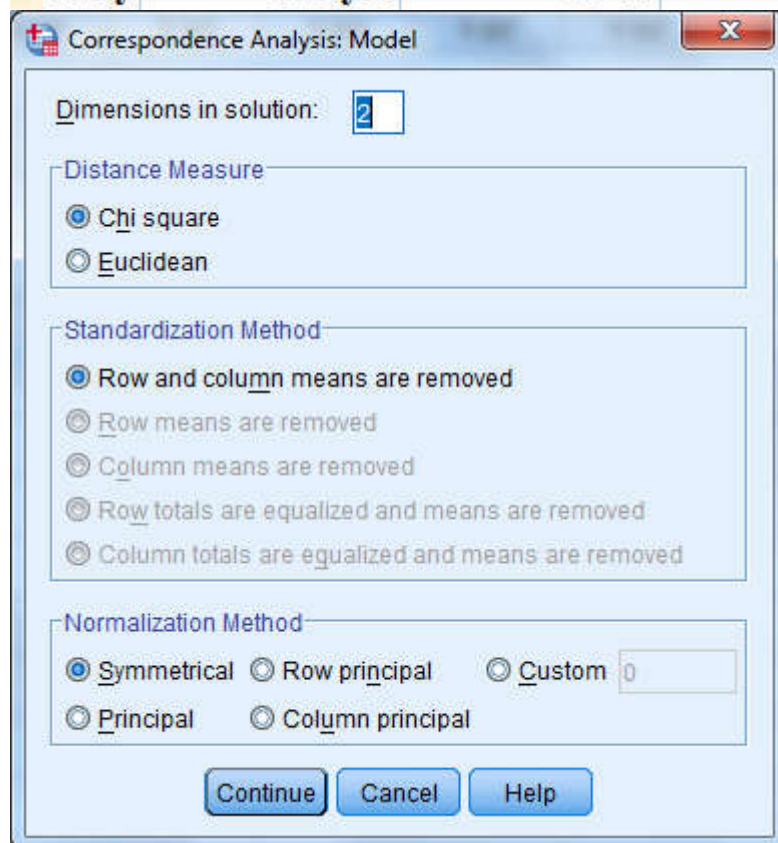
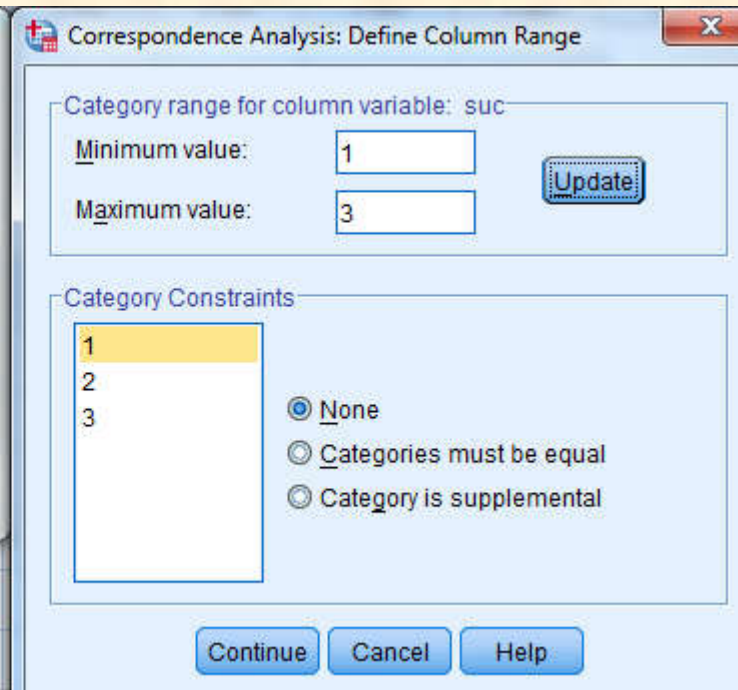
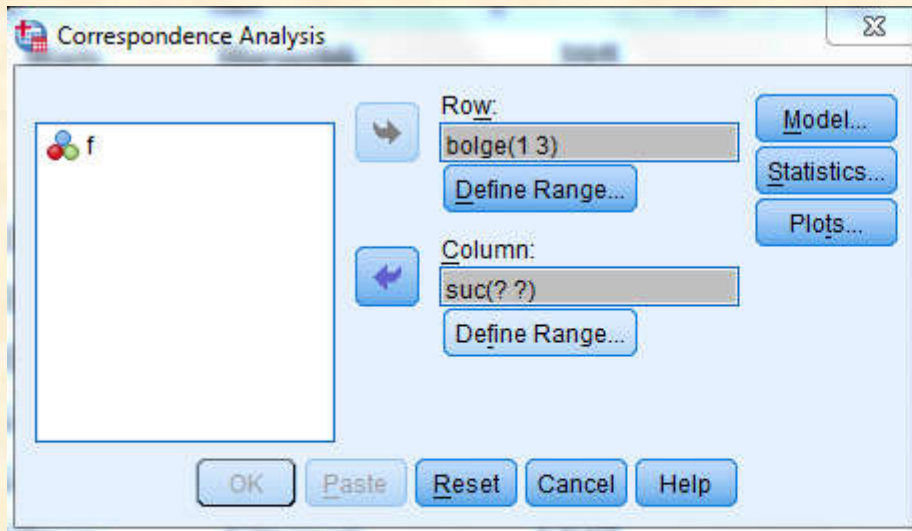
**Örnek.** Üç farklı suç tipini işleyen 8193 suçlunun bir ülkenin farklı üç bölgesine ait verileri aşağıdaki gibi bulunmuştur.

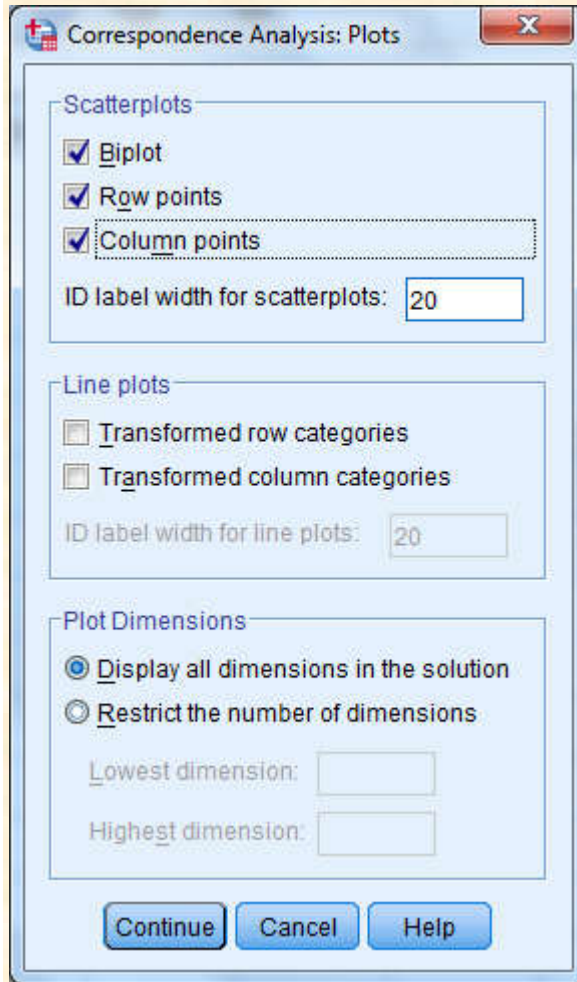
<b>Bölgeler</b>	<b>Suç Tipi</b>			<b>Toplam</b>
	Hırsızlık	Gasp	Cinayet	
Batı	395	2456	1758	4609
Doğu	147	153	916	1216
Güney	694	327	1347	2368
Toplam	1236	2936	4021	8193



	bolge	suc	f
1	Bati	Hirsizlik	395
2	Bati	Gasp	2456
3	Bati	Cinayet	1758
4	Doğu	Hirsizlik	147
5	Doğu	Gasp	153
6	Doğu	Cinayet	916
7	Güney	Hirsizlik	694
8	Güney	Gasp	327
9	Güney	Cinayet	1347







**Correspondence Table**

bolge	suc			
	Hırsızlık	Gasp	Cinayet	Active Margin
Bati	395	2456	1758	4609
Doğu	147	153	916	1216
Güney	694	327	1347	2368
Active Margin	1236	2936	4021	8193

**Row Profiles**

bolge	suc			
	Hırsızlık	Gasp	Cinayet	Active Margin
Bati	,086	,533	,381	1,000
Doğu	,121	,126	,753	1,000
Güney	,293	,138	,569	1,000
Mass	,151	,358	,491	

**Column Profiles**

bolge	suc			
	Hırsızlık	Gasp	Cinayet	Mass
Bati	,320	,837	,437	,563
Doğu	,119	,052	,228	,148
Güney	,561	,111	,335	,289
Active Margin	1,000	1,000	1,000	

Satır ve sütun profilleri  
ve oranları

### Summary

Dimension	Singular Value	Inertia	Chi Square	Sig.	Proportion of Inertia		Confidence Singular Value	
					Accounted for	Cumulative	Standard Deviation	Correlation
1	,421	,177			,874	,874	,009	,112
2	,160	,025			,126	1,000	,012	
Total		,203	1662,625	,000 <sup>a</sup>	1,000	1,000		

a. 4 degrees of freedom

İki boyutlu gösterimde Ki-kare istatistiği  $p=0,000<0,05$  önemli bulunmuştur. İki özdeğer 0,874 ve 0,126 bulunmuştur. İki boyutlu çözümde ilk boyut toplam değişkenliğin %87,4'ünü (0,177/0,203), ikinci boyut ise %12,6'sını (0,025/0,203) açıklamaktadır.

## Satır deęişkeninin (Bölge) X ve Y koordinatları, Satırların boyutlara-boyutların satırlara katkısı

Overview Row Points<sup>a</sup>

bolge	Mass	Score in Dimension		Inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		Total
					1	2	1	2	
Batı	,563	-,571	-,023	,077	,436	,002	,999	,001	1,000
Doęu	,148	,647	,870	,044	,147	,704	,593	,407	1,000
Güney	,289	,779	-,403	,081	,417	,294	,908	,092	1,000
Active Total	1,000			,203	1,000	1,000			

a. Symmetrical normalization

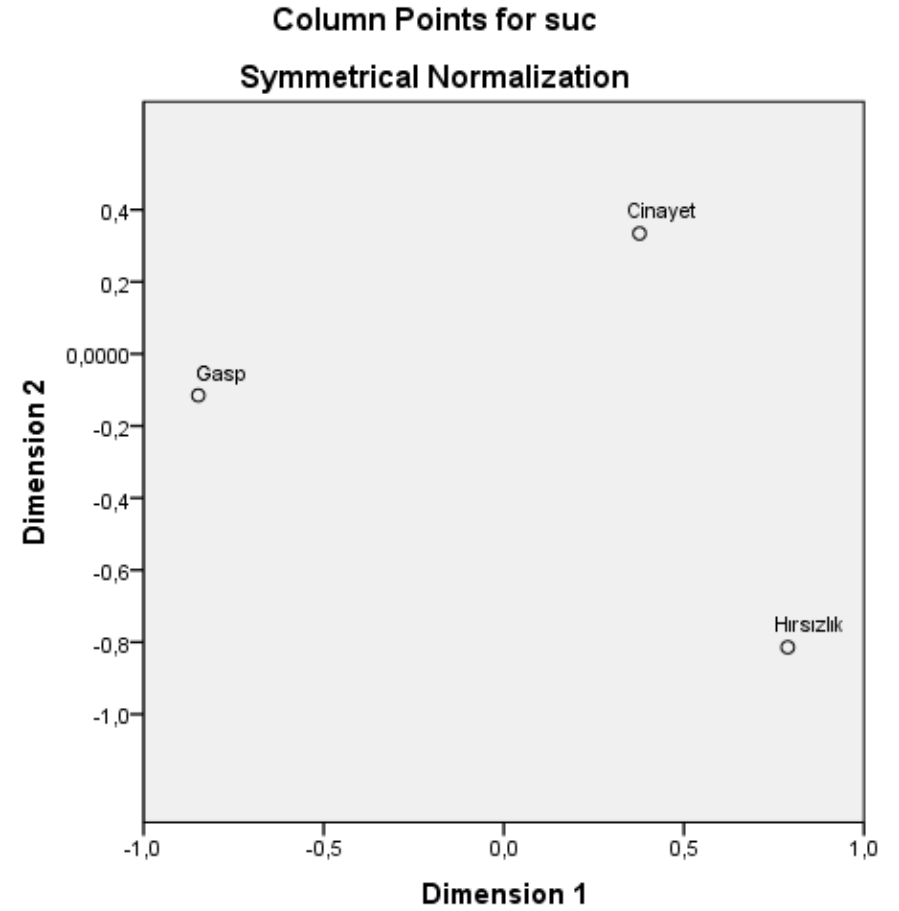
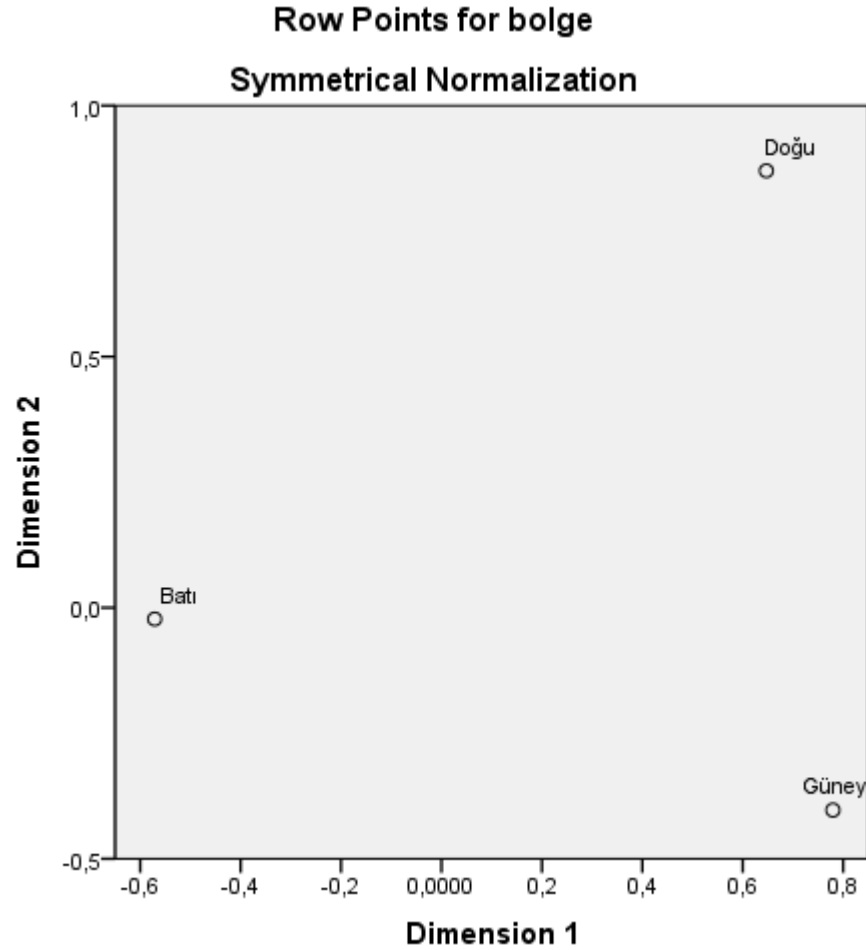
## Sütun deęişkeninin (Suç) X ve Y koordinatları, Sütunların boyutlara-boyutların sütunlara katkısı

Overview Column Points<sup>a</sup>

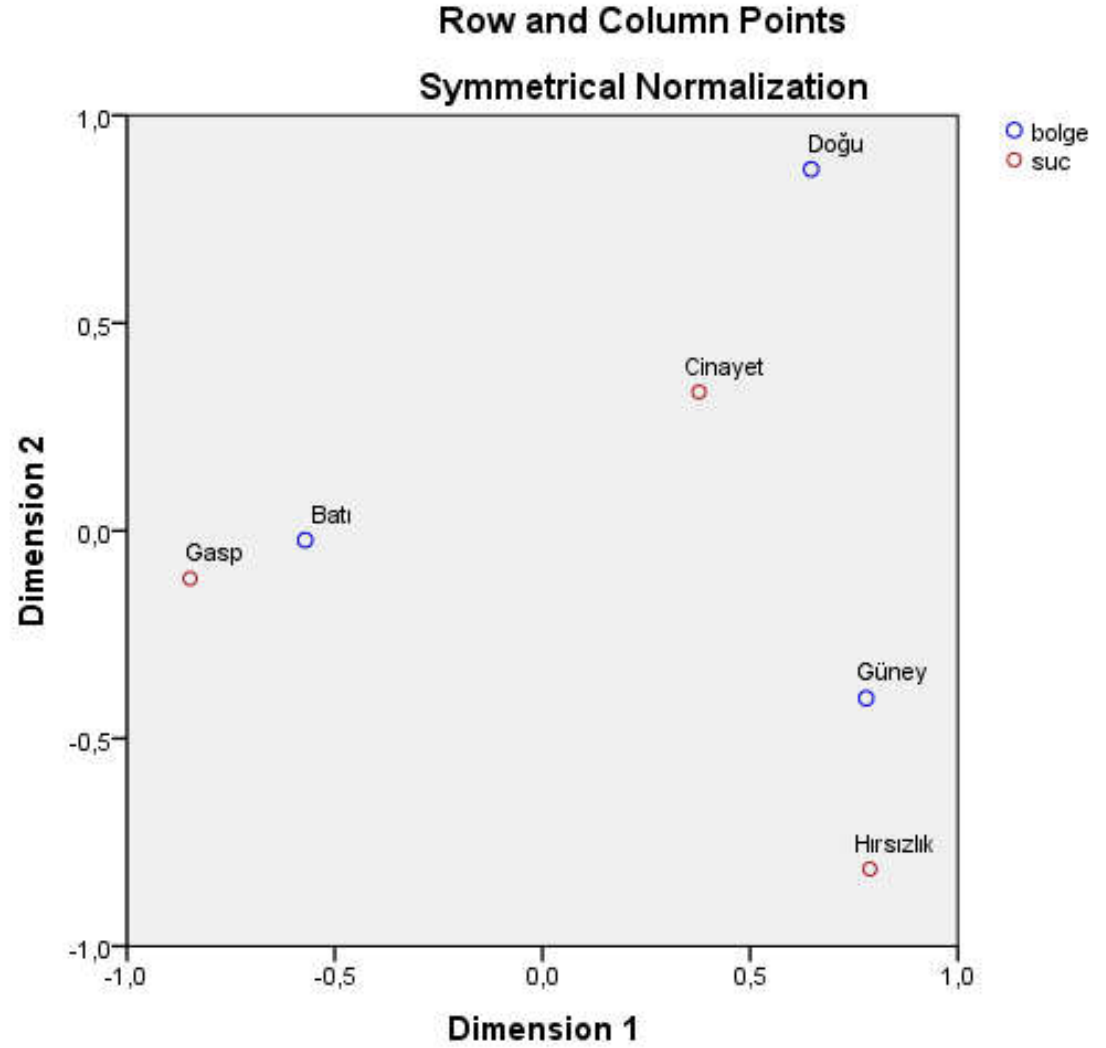
suc	Mass	Score in Dimension		Inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		Total
					1	2	1	2	
Hırsızlık	,151	,788	-,814	,055	,223	,627	,712	,288	1,000
Gasp	,358	-,848	-,115	,109	,612	,030	,993	,007	1,000
Cinayet	,491	,377	,334	,038	,166	,344	,770	,230	1,000
Active Total	1,000			,203	1,000	1,000			

a. Symmetrical normalization

# Kategoriler arasındaki ilişkinin grafiksel gösterimi (satırlar arası-sütunlar arası uzaklıklar)



## Bölge-Suç'un birlikte gösterimi



Batı ile Gasp noktaları birbirine yakındır. Yani Gasp suçunun işleyenlerin daha çok Batı bölgesinde toplandıkları söylenebilir. Yine Güney bölgesinde daha çok Hırsızlık yapıldığı, Doğu bölgesinde ise Cinayet daha çok işlenmiştir söylenebilir.