

REGRESYON VE KORELASYON DEĞİŞKENLERARASI BAĞINTI VE İLİŞKİ

- **SİSTEM, ALT SİSTEM ve SİSTEM DİNAMİKLERİ**
- **TERİM ve TANIMLAR**
- **REGRESYON YÖNTEMLERİ**

BASİT DOĞRUSAL REGRESYON

SPSS'de REGRESYON ANALİZİ

- **KORELASYON ve ÇEŞİTLERİ**

**BASİT KORELASYON ANALİZİ (PEARSON
KORELASYON ANALİZİ)**

SPSS'de KORELASYON ANALİZİ

- **ÇOKLU DOĞRUSAL REGRESYON ANALİZİ**
- **ROBUST (SAĞLAM) REGRESYON**
- **LOJİSTİK REGRESYON**

İkili Lojistik Regresyon (BLOGREG) Analizi

Sıralı Lojistik Regresyon (OLOGREG) Analizi

İsimsel Lojistik Regresyon (NLOGREG) Analizi

SİSTEM, ALT SİSTEM ve SİSTEM DİNAMİKLERİ

Doğa bir ana sistemdir. Bu sistemin altsistemleri vardır. Biyolojik, Sosyo-ekonomik vb. sistemler. İNSAN yaşamı da bir sistemdir, Anatomik ve Fizyolojik (Kas-İskelet, Sinir, Dolaşım, Solunum, Boşaltım, Üregenital) Sistemler olarak alt sistemlere ayrılır.

Sistemler denge içinde çalışır. Dinamik sistemde bir takım girdiler (INPUT) ve Çıktılar (OUTPUT) vardır. Girdilerin ve Çıktıların iç ve dış dinamikleri sistemin olumlu (sağlıklı) ya da olumsuz (hastalıklı) davranmasını sağlar. Doğada bir çok değişken birbirlerini etkileyerek değer alırlar. Doğada denge bir etkileşimler zinciri içinde gerçekleşir. Sistemler Dengeli ise Sağlıklı, değilse Hastadır.

SİSTEM MODELLEMESİ

Sistemler matematiksel, istatistiksel ya da benzeşimsel olarak modellenenir.

Bu modelde sonuç/sonuçlar (cevap, response) çıktı, faktörler (faktor, predictor) girdi olarak yer alır. Girdilerin modeldeki etkileri, modelin dinamiğini belirler.

Neden-Sonuç (Girdi-Çıktı) ilişkilerini matematiksel olarak tanımlamak sistemin izlenmesi bakımından önemlidir. Doğada her oluşum bir sistem olarak ele alınarak modellenenir, incelenenir.

Modelleme, Regresyon ve Korelasyon

- ✓ Değişkenler bazı faktör/faktörlerden pozitif ya da negatif yönde etkilenirler. Faktörlerin bazılarının etkisi çok yüksek iken (majör, birincil faktörler), bazılarının etkileri çok düşük düzeydedir (minör, ikincil faktörler).
- ✓ Cevap değişkenleri etkileyen faktörlerin ortaya konması ve faktörlerin etki düzeylerinin belirlenmesi Regresyon ve Korelasyon yöntemleri aracılığı ile yapılır.
- ✓ Regresyon, iki ve daha fazla değişken arasındaki matematiksel bağıntıyı denklemlerle ifade etmek ve değişkenlerin birbirlerinden etkilenme biçimini ve büyüklüğünü ortaya koymak için yararlanılan bir istatistiksel yöntemdir.
- ✓ Korelasyon, değişkenler arasındaki ilişkinin yönünü, derecesini ve önemini ortaya koyan istatistiksel yöntemdir.

TERİM ve TANIMLAR

- ✓ **Faktör:** Bir hastalığın ortaya çıkmasında az ya da çok kesin etkisinin (neden, etken, sebep) bilindiği değişkenlerdir
- ✓ **Risk faktörü:** Bir hastalığın ortaya çıkmasında katkısının olduğu bilinen, fakat bu faktörün mutlaka hastalığa yol açmasının söz konusu olmadığı faktörlerdir. *Sigara, Yaş, Cins, Irk, Kimyasal ajanlar kanser'in birer risk faktörleridir.*
- ✓ **Bağımlı değişken:** Değeri başka değişkenlerin etkileri ile oluşan değişkenlerdir (dependent variable, response variable).
- ✓ **Bağımsız değişken:** Değeri rasgele koşullar altında oluşan değişkenlerdir (Independent variable, Explanatory variable, Factor variable, Predictor variable). Bağımsız değişken, bağımlı değişkenin değişimi üzerinde az ya da çok etkili olan değişkendir.

Model ve Değişkenlerarası Bağıntının Formulasyonu

- ✓ **Model** : Bir problemin çözümünde ya da bir olayın açıklanmasında yararlanılan matematiksel ya da benzetimsel sembolik yaklaşımlara model adı verilir.
- ✓ Model, genelde bir matematiksel eşitlik ya da eşitsizlik biçiminde belirlenir.

$$Y = f(X)$$

$$Y = \beta_0 + \beta_1 X$$

$$Y = \beta_0 + \beta_1 X + \beta_2 X_2 + \dots + \beta_p X_p$$

REGRESYON YÖNTEMLERİ

Regresyon yöntemleri, modeldeki değişken sayısına, değişkenin ölçüm tekniğine göre farklı şekillerde sınıflandırılmaktadır.

1. Modeldeki değişkenler sürekli ve değişken sayısı $k=2$ ise (bir bağımlı ($q=1$), bir bağımsız değişken($p=1$)) kurulacak regresyon modelleri

- ✓ Basit Doğrusal Regresyon
- ✓ **Polinomiyal regresyon**
- ✓ **Geometrik Regresyon**
- ✓ **Üssel Regresyon**
- ✓ Basit Eğrisel Regresyon (Nonlinear regresyon)

2. Modeldeki değişkenler sürekli (interval/orantılı) ve sayısı $k > 2$ ise (bir bağımlı($q=1$), iki ve daha fazla bağımsız değişken($p > 2$))

✓ Çoklu (Multiple) Doğrusal Regresyon

✓ Çoklu Doğrusal Olmayan (Multiple Nonlinear) Regresyon

3. Modeldeki bağımlı değişken nominal/ordinal/nominalize interval ölçekli, bağımsız değişkenler orantılı/interval/ordinal/nominal ölçekli ve enaz iki kategorili iseler $q=1$ $p \geq 1$ ise ;

✓ Lojistik Regresyon

✓ Ordinal Regresyon

✓ Robust Regresyon

✓ Poisson Regresyon

BASİT DOĞRUSAL REGRESYON

Y bağımlı (response, dependent) değişken ve X bağımsız (belirleyici, predictor) değişken olmak üzere iki değişken arasındaki sebep-sonuç ilişkisini doğrusal bir model ile ortaya koyan yöntem basit doğrusal regresyon denir.

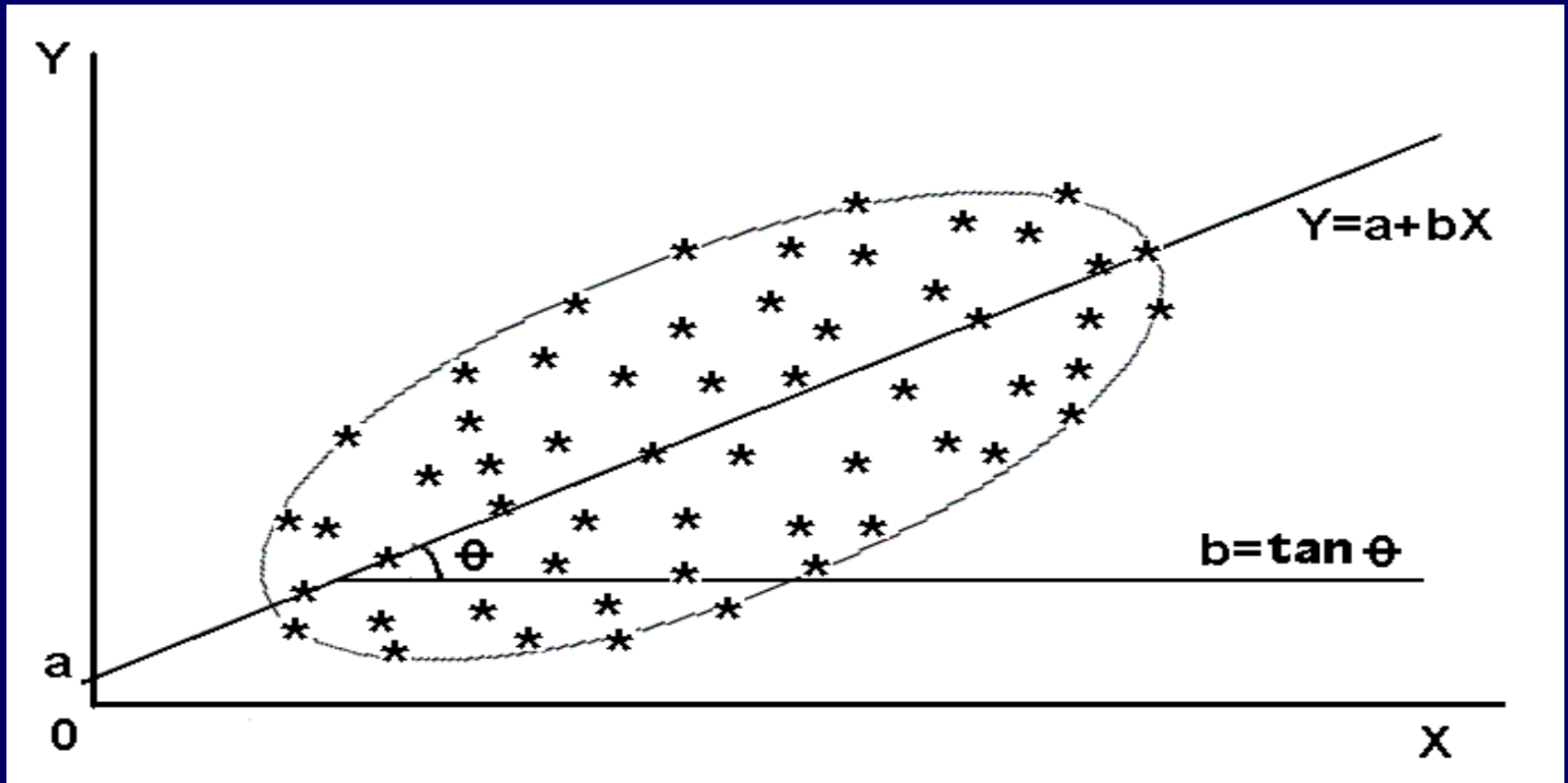
Basit doğrusal regresyon, iki değişken (Y, X) arasındaki neden-sonuç ilişkisini $Y=a+bX$ biçiminde bir denklem (model) ile ortaya koyar.

BASİT DOĞRUSAL REGRESYON

Basit doğrusal regresyon uygulamak için;

1. n birimden Y ve X değişkenleri için veriler toplanır. (X_i, Y_i) . Verilerin aralıklı ya da orantılı ölçekli olması gerekir.
2. Verilerin XY ilişki grafiği çizilir. Grafikteki xy noktaları bir çember ya da elips içine alınır.
3. Eğer noktaları sınırlayan çerçeve bir elips ve elipsin asal (ana) eksenini ikincil (yan) ekseninden daha büyük ise veriler arasında basit doğrusal bir bağıntı olabileceği varsayılır.
4. Verileri temsil eden $Y=a+bX$ doğrusunun denklemi hesaplanır.

BASİT DOĞRUSAL REGRESYON



a ve b Katsayılarının Hesaplanması

$$b = \frac{\sum_{i=1}^n X_i Y_i - \frac{\left(\sum_{i=1}^n X_i\right) \times \left(\sum_{i=1}^n Y_i\right)}{n}}{\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}}$$

$$a = \frac{\sum X_i^2 \sum Y_i - \sum X_i \sum X_i Y_i}{n \sum X_i^2 - (\sum X_i)^2}$$

$$a = \bar{Y} - b\bar{X}$$

a ve b Katsayılarının Hesaplanması

$$\zeta T_{xy} = \sum X_i Y_i - \frac{(\sum X_i)(\sum Y_i)}{n}$$

$$b = \zeta T_{xy} / K T_x$$

$$K T_x = \sum X_i^2 - \frac{(\sum X_i)^2}{n}$$

$$\bar{X} = \sum X / n$$

$$K T_y = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}$$

$$\bar{Y} = \sum Y / n$$

$$a = \bar{Y} - b\bar{X}$$

Modelin Önemliliği

- ✓ $Y=a+bX$ modelinin geçerliliğini belirlemek için Regresyon Analizi yönteminden yararlanılır.
- ✓ Modelin önemliliği, belirlenen model ile Y 'nin değişiminin X tarafından ne kadar açıklanabildiğinin kontrolü yapılır.
- ✓ Modelin önemliliği aynı zamanda eğimin regresyon katsayısının önemliliğini ve iki değişken arasındaki korelasyonun da önemliliğini verir.

Tahminin Varyansı, b'nin Varyansı

$$S^2 = S_{Y.X}^2 = \frac{\sum (Y_i - (a + bX_i))^2}{n - 2}$$

$$s^2 = \frac{1}{n - 2} \left[KT_y - \frac{(CT_{XY})^2}{KT_x} \right]$$

$$s_b^2 = \left\{ \frac{1}{n - 2} \left[KT_y - \frac{(CT_{XY})^2}{KT_x} \right] \right\} / KT_y$$

b'nin Önemliliği

$$T = \frac{b - \beta}{\sqrt{S_b^2}}$$

$$H_0: \beta=0$$

$$sd=n-2$$

$$t < t_{0.05, sd} \quad P > 0.05 \text{ n.s.}$$

$$t > t_{0.05, sd} \quad P < 0.05 *$$

$$t > t_{0.01, sd} \quad P < 0.01 **$$

$$t > t_{0.001, sd} \quad P < 0.001 ***$$

$$s_b^2 = \left\{ \frac{1}{n-2} \left[KT_y - \frac{(CT_{XY})^2}{KT_x} \right] \right\} / KT_Y$$

Modelin Önemliliğinin Belirlenmesi

Genel $KT = \text{Regresyon } KT + \text{Artık } KT$

$GKT = RKT + AKT$

$$RKT = (CT_{XY})^2 / KT_X$$

$rsd=1$

$$AKT = KT_Y - RKT$$

$asd=n-2$

$$GKT = KT_Y$$

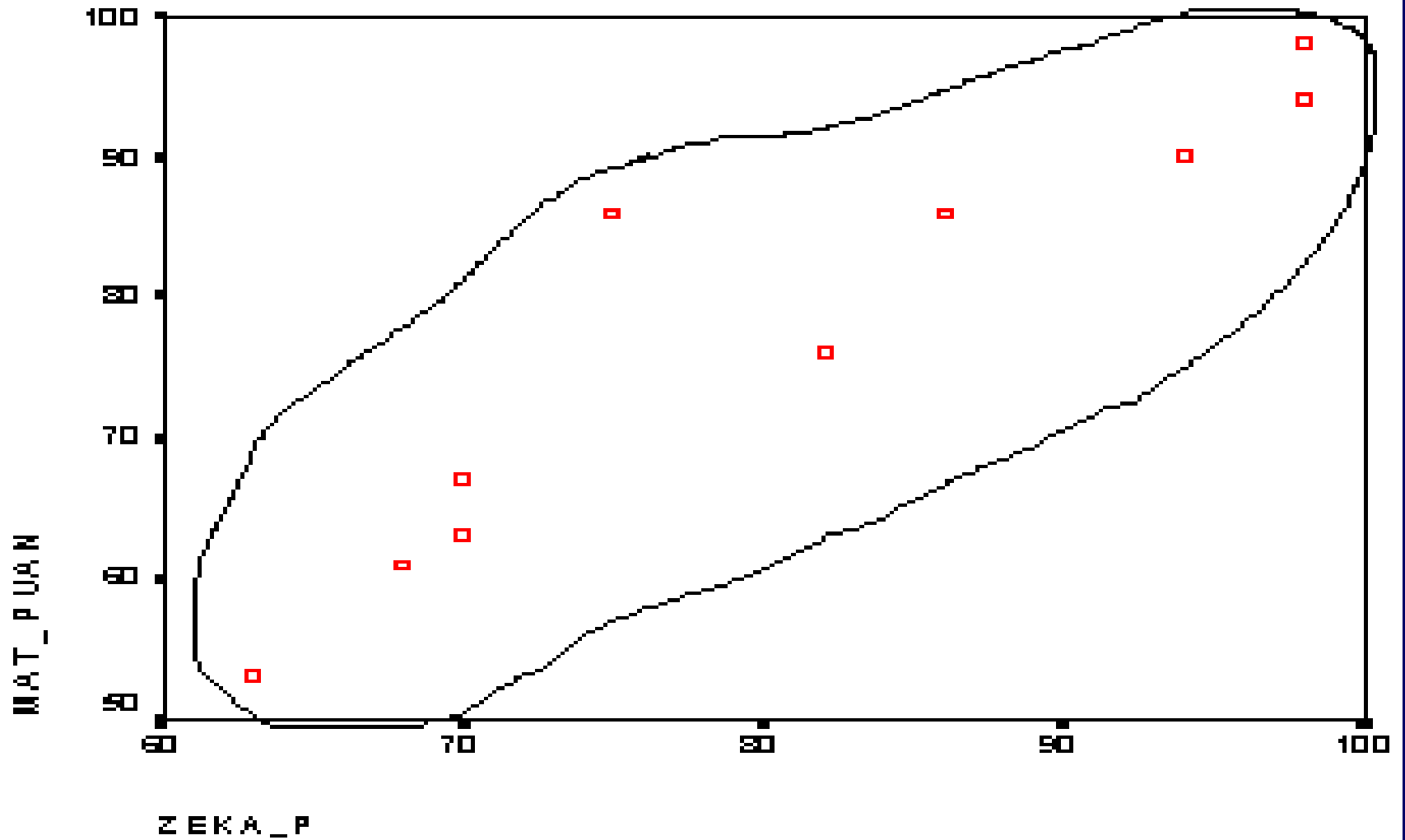
Regresyon Analizi Tablosu

DK	sd	KT	KO	F	p
Regresyon	1	RKT	RKO	RKO/AKO	
Artık	n-2	AKT	AKO	-	-
Genel	n-1	GK _Y	-	-	-

$F(rsd, asd) < F(0.05, rsd, asd)$ $P > 0.05$ ns. Model önemsiz
 $F(rsd, asd) > F(0.05, rsd, asd)$ $P < 0.05$ * Model önemli.
 $F(rsd, asd) < F(0.01, rsd, asd)$ $P < 0.01$ ** Model önemli.
 $F(rsd, asd) < F(0.001, rsd, asd)$ $P < 0.001$ *** Model önemli.

10 Lise Öğrencisinin Matematik ve Zeka Puanları

Öğr. No	Mat_P (Y)	Zeka_P (X)
1	86	75
2	67	70
3	90	94
4	94	98
5	53	63
6	61	68
7	86	86
8	76	82
9	98	98
10	63	70
T	774	804



Şekil: 10 Orta Okul Öğrencisinin Matematik ve Zeka Puanları İlişki Grafiği

Tablo: 10 Lise Öğrencisinin Matematik, Zeka Puanları ve Gerekli Hesaplamalar

Öğr. No	Mat_P (Y)	Zeka_P (X)	Y2	X2	XY
1	86	75	7396	5625	6450
2	67	70	4489	4900	4690
3	90	94	8100	8836	8460
4	94	98	8836	9604	9212
5	53	63	2809	3969	3339
6	61	68	3721	4624	4148
7	86	86	7396	7396	7396
8	76	82	5776	6724	6232
9	98	98	9604	9604	9604
10	63	70	3969	4900	4410
T	774	804	62096	66182	63941

$$\zeta T_{xy} = \sum X_i Y_i - \frac{(\sum X_i)(\sum Y_i)}{n}$$

$$CT_{xy} = 63941 - \frac{(804)(774)}{10} = 1711.4$$

$$KT_x = \sum X_i^2 - \frac{(\sum X_i)^2}{n}$$

$$KT_x = 66182 - \frac{(804)^2}{10} = 1540.4$$

$$b = \zeta T_{xy} / KT_x$$

$$b = 1711.4 / 1540.4$$

$$b = 1.111$$

$$\bar{X} = 804 / 10 = 80.4$$

$$\bar{Y} = 774 / 10 = 77.4$$

$$a = \bar{Y} - b\bar{X}$$

$$a = 77.4 - 1.11 * 80.4$$

$$a = -11.92$$

Regresyon Denklemi

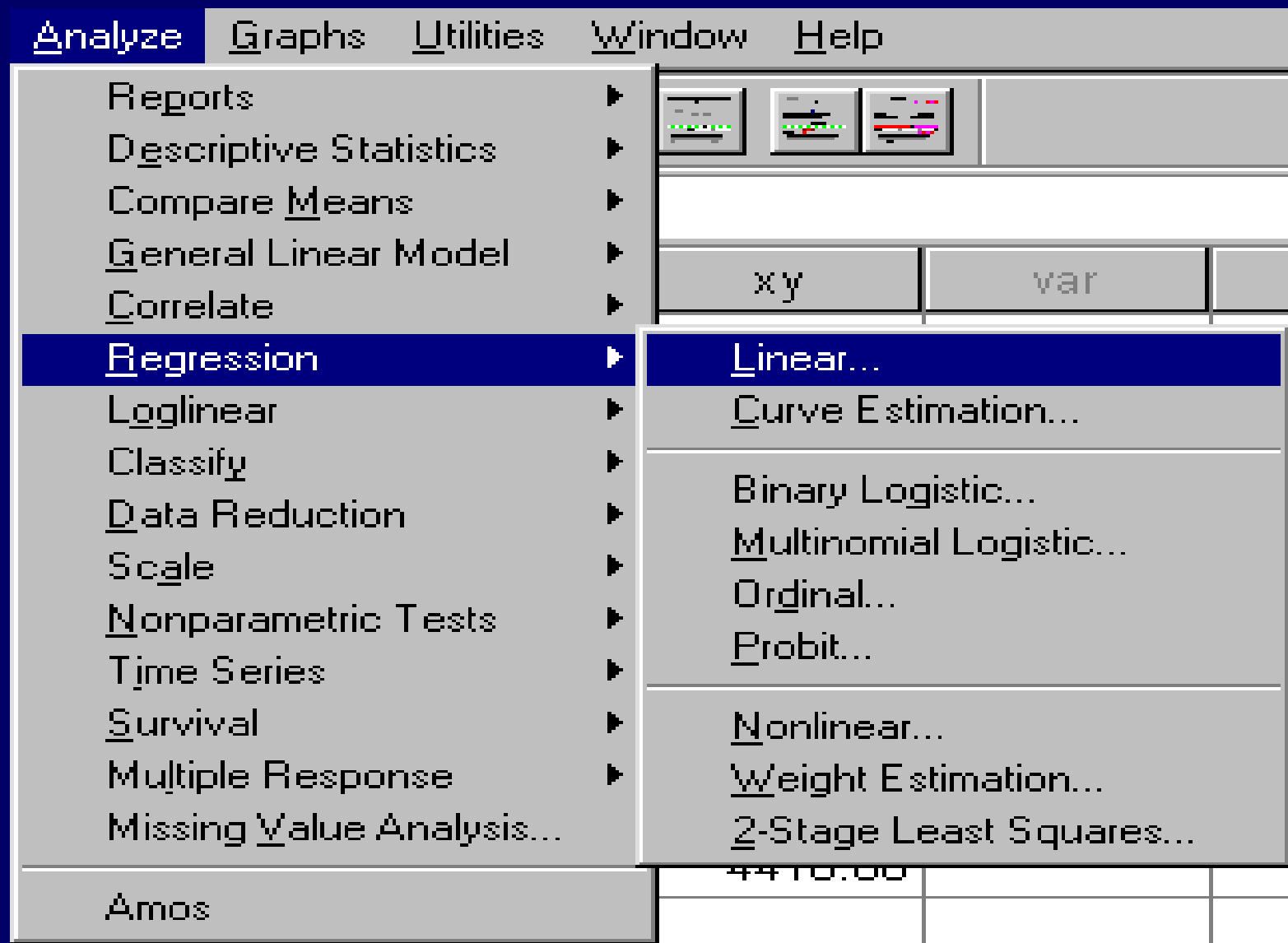
$$Y = -11.92 + 1.111 * X$$

SPSS'de REGRESYON ANALİZİ

- ✓ SPSS veri sayfasında X ve Y verilerini farklı sütunlara giriniz
- ✓ **Analyze > Regression > Linear** seçeneklerini tıklayınız.
- ✓ İşlem penceresinde X ve Y değişkenlerini doğru tanımlayarak alanlara taşıyınız.
- ✓ **OK** tıklayınız.

	mat_puan	zeka_p	y2	x2	xy	
1	86.00	75.00	7396.00	5625.00	6450.00	
2	67.00	70.00	4489.00	4900.00	4690.00	
3	90.00	94.00	8100.00	8836.00	8460.00	
4	94.00	98.00	8836.00	9604.00	9212.00	
5	53.00	63.00	2809.00	3969.00	3339.00	
6	61.00	68.00	3721.00	4624.00	4148.00	
7	86.00	86.00	7396.00	7396.00	7396.00	
8	76.00	82.00	5776.00	6724.00	6232.00	
9	98.00	98.00	9604.00	9604.00	9604.00	
10	63.00	70.00	3969.00	4900.00	4410.00	
11						

SPSS'de REGRESYON



zeka_p



Dependent:

mat_puan

OK

Paste

Reset

Cancel

Help

Previous

Block 1 of 1

Next

Independent(s):



zeka_p

Method: Enter

Selection Variable:



Rule...

Case Labels:



WLS >>

Statistics...

Plots...

Save...

Options...

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1901.383	1	1901.383	52.997	.000 ^a
	Residual	287.017	8	35.877		
	Total	2188.400	9			

a. Predictors: (Constant), ZEKA_P

b. Dependent Variable: MAT_PUAN

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-11.925	12.415		-.961	.365
	ZEKA_P	1.111	.153	.932	7.280	.000

a. Dependent Variable: MAT_PUAN

KORELASYON ve ÇEŞİTLERİ

Korelasyon (Correlation), değişkenler arasındaki ilişkinin yönünü, derecesini ve önemini ortaya koyan istatistiksel yöntemdir. Değişkenlerin sayısına ve hesaplama biçimine göre;

- ✓ İkili (Bivariate) Korelasyon
- ✓ Kısmi (Partial) Korelasyon
- ✓ Çoklu (Multiple) Korelasyon
- ✓ Setlerarası (Canonical) Korelasyon

isimleri ile anılır.

BASİT KORELASYON ANALİZİ (PEARSON KORELASYON ANALİZİ)

İki değişken arasındaki ilişkiyi, önemini, yönünü inceleyen korelasyon yöntemidir. Korelasyon, korelasyon katsayısı ile ölçülür. r_{XY} ile gösterilir. Önemlilik t testi ile belirlenir.

$$r_{XY} = \frac{\sum_{i=1}^n X_i Y_i - \frac{\left(\sum_{i=1}^n X_i\right) \times \left(\sum_{i=1}^n Y_i\right)}{n}}{\sqrt{\left(\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}\right) \left(\sum_{i=1}^n Y_i^2 - \frac{\left(\sum_{i=1}^n Y_i\right)^2}{n}\right)}}$$

$$t = \frac{r\sqrt{(n-2)}}{\sqrt{1-r^2}} \quad \text{sd} = n - 2$$

Korelasyon Katsayısının Hesaplanması

$$CT_{xy} = 1711.4$$

$$KT_x = 1540.4$$

$$KT_Y = 62096 - \frac{(774)^2}{10}$$

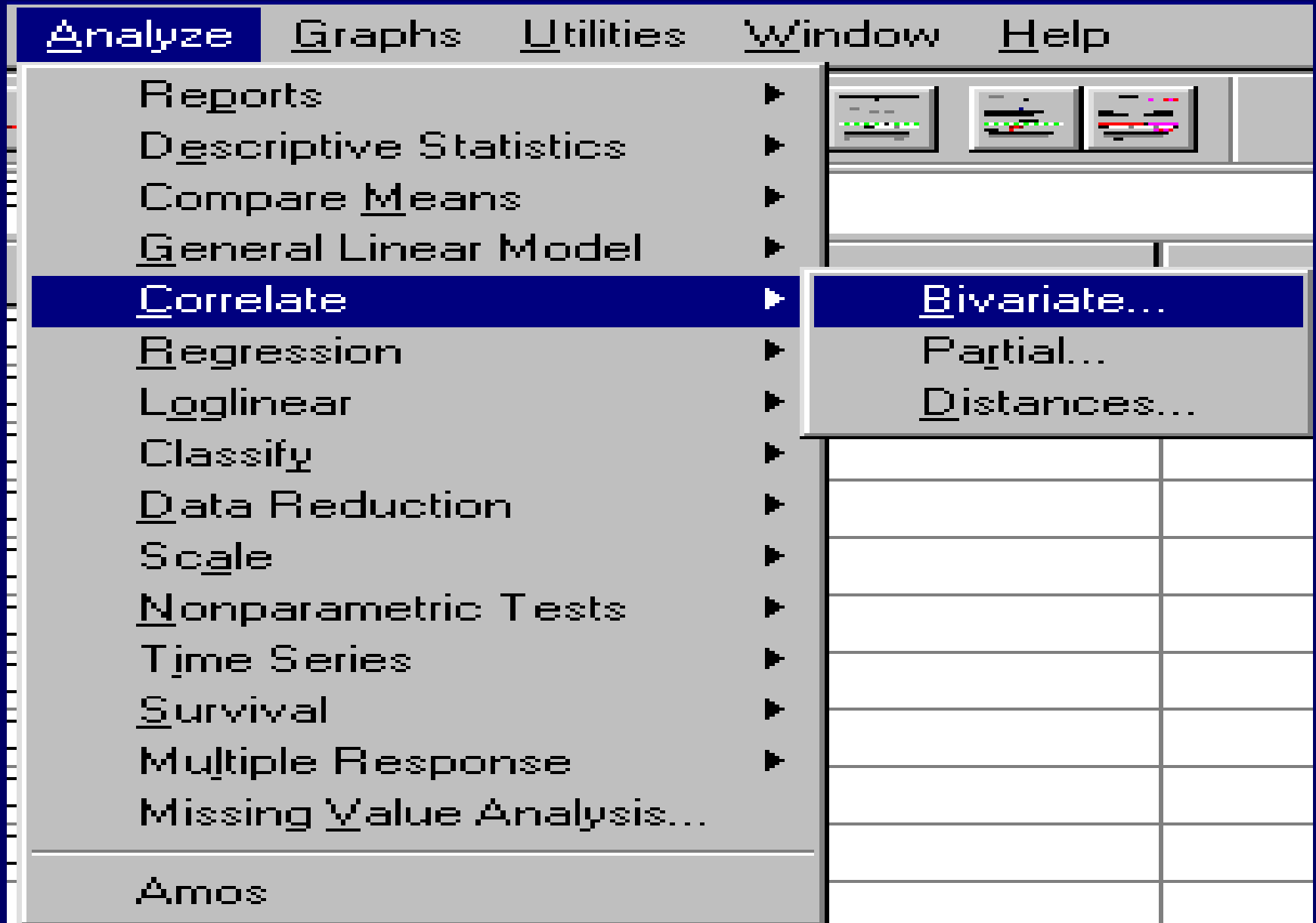
$$KT_Y = 2188.4$$

$$r = \frac{1711.4}{\sqrt{1540.4 * 2188.4}} = \frac{1711.4}{\sqrt{1836.03}} = 0.932$$

$$t = \frac{0.932 * \sqrt{(10 - 2)}}{\sqrt{1 - (0.932)^2}} = 7.27$$

$$t = 7.27, \quad sd = 8, \quad P < 0.001 ***$$

SPSS'de KORELASYON ANALİZİ



Bivariate Correlations



Empty box for output or additional variables.



Variables:

mat_puan
zeka_p

OK

Paste

Reset

Cancel

Help

Correlation Coefficients

☒ Pearson ☐ Kendall's tau-b ☐ Spearman

Test of Significance

☒ Two-tailed ☐ One-tailed

☒ Flag significant correlations

Options...

Correlations

		MAT PUAN	ZEKA P
MAT_PUAN	Pearson Correlation	1.000	.932**
	Sig. (2-tailed)	.	.000
	N	10	10
ZEKA_P	Pearson Correlation	.932**	1.000
	Sig. (2-tailed)	.000	.
	N	10	10

** . Correlation is significant at the 0.01 level (2-tailed).

Verilere basit doğrusal regresyon uygulanıyor ise korelasyon analizi sonuçları da regresyon çıktısı içinde yer alır.

Mat_P ve Zeka_P verileri Örneğimize regresyon uygulaması tekrarlanırsa sonuçlar aşağıdaki gibi elde edilir.

Correlations

		MAT_PUAN	ZEKA_P
MAT_PUAN	Pearson Correlation	1.000	.932**
	Sig. (2-tailed)	.	.000
	N	10	10
ZEKA_P	Pearson Correlation	.932**	1.000
	Sig. (2-tailed)	.000	.
	N	10	10

** . Correlation is significant at the 0.01 level (2-tailed).

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.932 ^a	.869	.852	5.9898

a. Predictors: (Constant), ZEKA_P

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-11.925	12.415		-.961	.365
	ZEKA_P	1.111	.153	.932	7.280	.000

a. Dependent Variable: MAT_PUAN

ÇOKLU DOĞRUSAL REGRESYON ANALİZİ

Y bağımlı değişken ve $X_1, X_2, \dots, X_p$ bağımsız değişkenler olmak üzere değişkenler arasındaki sebep-sonuç ilişkisini matematiksel bir model olarak ortaya koyan yönteme **çoklu regresyon analizi** adı verilir.

Bir bağımlı değişken ile bu değişkenin değişimi üzerinde etkide bulunan p sayıda bağımsız değişken arasındaki ilişkinin düzeyini belirleyen yönteme ise **çoklu korelasyon analizi** denilmektedir.

Genellikle çoklu regresyon ve korelasyon analizi birlikte ele alınan ve hesaplamaları birlikte yapılan karma yöntemlerdir.

ÇOKLU DOĞRUSAL REGRESYON

Çoklu Doğrusal Regresyon Analizi, Y ile iki ve daha fazla açıklayıcı değişken arasındaki ilişkiyi

$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_pX_p$ biçiminde inceler.

Verilere uyan modelin açıklayıcılık yüzdesi belirtme katsayısı R^2 ile belirlenir.

Regresyon analizi, modelin tutarlılığını; tahmin gücünü ve her bir değişkenin Y üzerindeki açıklayıcılığını test eder. Modelin belirleyicilik gücünü ifade eden R^2 , aşağıdaki gibi hesaplanır.

$$R^2 = \frac{RKT}{KT_Y} = \frac{\text{Regresyon Kareler Toplamı}}{\text{Genel Kareler Toplamı}}$$

ÇOKLU DOĞRUSAL REGRESYON

Çoklu belirtme katsayısı (R^2) modele yeni bir değişken eklendiğinde artış gösterir. Modele yeni bir değişken eklenmesine rağmen paydanın değeri değişmezken payın değeri artar. Bu nedenle R^2 hesaplanırken değişken sayısına göre düzeltme yapılması gerekir.

Düzeltilmiş R^2 değeri ($R^2_{\text{düz}}$),

$$R^2_{\text{düz}} = 1 - \left[\frac{\sum e^2 / (N - k)}{\sum y^2 / (N - 1)} \right] \quad \text{ya da}$$

$$R^2_{\text{düz}} = 1 - \frac{(1 - R^2)(N - 1)}{N - k - 1} \quad \text{biçiminde hesaplanır}$$

ÇOKLU DOĞRUSAL REGRESYON

Örnek: Rasgele seçilen 16 bireyin Günlük İçtiği Sigara Sayısı, YAŞ, BOY, AĞIRLIK ve SKB değerleri aşağıdaki tabloda verilmiştir

GİSS	YAS (yil)	BOY (cm)	AGIRLIK (Kg)	SKB (mm/Hg)
10	51	166	67.0	115
15	64	165	61.0	122
20	46	174	83.0	130
25	39	168	78.9	126
0	58	162	67.0	110
30	54	178	90.0	141
12	31	171	77.7	124
40	67	173	89.3	150
0	48	165	70.0	110
10	78	152	58.0	119
18	39	177	82.5	130
20	51	166	63.0	120
45	73	178	93.1	149
27	53	174	89.0	125
30	56	169	72.0	125
5	47	159	64.0	114

ÇOKLU DOĞRUSAL REGRESYON

SPSS'de çoklu regresyon analizi uygulamak için **SKB, YAS, BOY, KILO ve GISS** değerleri ayrı sütunlara girilir. Çoklu regresyon analizi uygulamak için *Statistics>Regression>Linear* seçenekleri tıklanır. Açılan ekranda *Dependent* alanına SKB değişkeni, *Independent(s)* alanına ise diğer dört değişken taşınır. *OK* tıklanır.

ÇOKLU DOĞRUSAL REGRESYON

Untitled - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

5:

	yas	boy	giss	var
1	51,00	166,0		
2	64,00	165,0		
3	46,00	174,0		
4	39,00	168,0		
5	58,00	162,0		
6	54,00	178,0		
7	31,00	171,0		
8	67,00	173,0		
9	48,00	165,00	70,00	110,00
10	78,00	152,00	58,00	119,00
11	39,00	177,00	82,50	130,00

Regression

- Linear...
- Curve Estimation...
- Binary Logistic...
- Multinomial Logistic...
- Ordinal...
- Probit...
- Nonlinear...
- Weight Estimation...
- 2-Stage Least Squares...

ÇOKLU DOĞRUSAL REGRESYON

Linear Regression

yas
boy
agirlik
giss

Dependent:
skb

Previous Block 1 of 1 Next

Independent(s):
yas
boy
agirlik

Method: Enter

Selection Variable:

Case Labels:

WLS >> Statistics... Plots... Save... Options...

OK
Paste
Reset
Cancel
Help

Rule...

ÇOKLU DOĞRUSAL REGRESYON

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	GISS, YAS, AGIRLIK, BOY ^a	,	Enter

a. All requested variables entered.

b. Dependent Variable: SKB

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,942 ^a	,887	,846	4,7942

a. Predictors: (Constant), GISS, YAS, AGIRLIK, BOY

ÇOKLU DOĞRUSAL REGRESYON

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1990,918	4	497,729	21,655	,000 ^a
	Residual	252,832	11	22,985		
	Total	2243,750	15			

a. Predictors: (Constant), GISS, YAS, AGIRLIK, BOY

b. Dependent Variable: SKB

ÇOKLU DOĞRUSAL REGRESYON

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	53,292	65,883		,809	,436
YAS	,161	,139	,166	1,162	,270
BOY	,170	,448	,100	,379	,712
AGIRLIK	,333	,260	,315	1,281	,226
GISS	,521	,182	,562	2,872	,015

a. Dependent Variable: SKB

ROBUST (SAĞLAM) REGRESYON

Sıralama puanları kullanılarak kategorik verilere regresyon modeli uygulamak için ***ROBUST regresyon*** yaklaşımı kullanılır.

Robust regresyon analizi, SPSS paket programında bulunmadığından dolayı bu analizi diğer bir istatistik paket programı olan MINITAB ile inceleyeceğiz.

ROBUST (SAĞLAM) REGRESYON

MINITAB'de verilere sağlam (robust) regresyon uygulamak için sıralı ya da aralıklı ölçekli veriler sıralama puanlarına dönüştürülür. MINITAB'de *Robust regresyon* menü seçenekleri ile uygulanamamaktadır. Bunun için MTB> iletişiminde iken **RREGRES** komutundan yararlanılır. Komut yazılımı; **>RREGRES dep_var, pred_numb, predictor(s)** biçimindedir.

REGRESS komutundan sonra bağımlı değişken, açıklayıcı değişken sayısı ve açıklayıcı değişkenlerin sıralama puanlarının yer aldığı sütun numaraları yazılır.

ROBUST (SAĞLAM) REGRESYON

Örnek: 12 öğrencinin yılsonu başarı puanları, sosyal etkinlik skorları, sosyo-ekonomik düzey skorları, babanın eğitim düzeyi ve yıl içi devamsızlık gün sayıları verilmiştir. Veriler aralıklı ölçekli skor değerlerdir.

Öğrencilerin yılsonu başarı puanlarını etkileyen faktörlerin regresyon modelini ROBUST regresyon uygulanması ile bulalım.

ROBUST (SAĞLAM) REGRESYON

Tablo: 12 bireyin başarı, sosyal etk., sosyo-eko., babanın eğitim düzeyi ve yıl içi devamsızlık gün sayısı değerleri

Birey No	Sosyal Etk. Puanı	Sosyo-Eko. Düzeyi	Babanın Eğ. Düzeyi	Başarı Puanı	Yıl içi dev. gün sayısı
1	2	3	2	67	2
2	3	4	4	78	4
3	2	5	4	79	5
4	4	4	5	66	3
5	5	3	3	56	4
6	3	6	5	90	5
7	4	2	2	45	6
8	5	3	3	53	4
9	6	5	5	67	1
10	3	6	5	78	1
11	2	4	4	58	2
12	1	5	5	61	4

ROBUST (SAĞLAM) REGRESYON

MTB > rank c1 c6

MTB > rank c2 c7

MTB > rank c3 c8

MTB > rank c4 c9

MTB > rank c5 c10

MTB > rregres c9 4 c6-c8 c10

ROBUST (SAĞLAM) REGRESYON

The regression equation is

$$C9 = 0.69 + 0.108 C6 + 1.52 C7 - 0.787 C8 + 0.0222 C10$$

Predictor	Coefficient		StDev	
	Rank	Least-sq	Rank	Least-sq
Constant	0.692	1.921	3.959	3.177
C6	0.1079	0.0070	0.3007	0.2414
C7	1.5218	1.3555	0.6150	0.4936
C8	-0.7873	-0.6660	0.6187	0.4966
C10	0.0222	0.0080	0.2838	0.2278
Hodges-Lehmann estimate of tau = 3.117				Least-squares S = 2.502

ROBUST (SAĞLAM) REGRESYON

Çıktı incelendiğinde orijinal değerlerin **C1-C5**, **C6-C10** sütunlarına sıralama puanları olarak yazıldığı görülür. Sıralama puanlarına uygulanan regresyon analizinde model;

$$C9 = 0.69 + 0.108 C6 + 1.52 C7 - 0.787 C8 + 0.0222 C1$$

olarak belirlenir.

LOJİSTİK REGRESYON

- ✓ Lojistik regresyon; cevap değişkenin kategorik, ikili (binary, dichotomous), üçlü ve çoklu kategorilerde gözlemlendiği durumlarda açıklayıcı değişkenlerle neden sonuç ilişkisini belirlemede yararlanılan bir yöntemdir.
- ✓ Açıklayıcı değişkenlere göre cevap değişkenin beklenen değerleri olasılık olarak elde edildiği bir regresyon yöntemidir.
- ✓ Basit ve çoklu regresyon yönteminde bağımlı değişkenin normal dağılım göstermesi, bağımsız değişkenlerin normal dağılım göstermesi ve hata varyansının $\varepsilon \cong N(0, \sigma^2)$ parametrelili normal dağılım göstermesi gerekmektedir. Bu koşulları içermeyen veri setlerine basit ya da çoklu regresyon analizleri uygulanamaz.

LOJİSTİK REGRESYON

Lojistik regresyon analizi, sınıflama ve atama işlemi yapmaya yardımcı olan bir regresyon yöntemidir. Normal dağılım varsayımı, süreklilik varsayımı ön koşulu yoktur.

Bağımlı değişken üzerinde açıklayıcı değişkenlerin etkileri olasılık olarak elde edilerek risk faktörlerinin olasılık olarak belirlenmesi sağlanır.

LOJİSTİK REGRESYON

Doğada gözlenen fenomenlerin bazıları var-yok, başarılı-başarısız gibi ikili biçimde sonuçlanırlar. Bazı sonuçlar ise yok-orta-çok, hiç-az-çok, olumsuz-olumlu-çok olumlu biçiminde üçlü gözlem sonuçları olarak belirlenirler. Bazı sonuçlar ise çok sınıflı kategorik ya da sıralı ölçekli değerler olarak belirlenebilirler.

Bu sonuçların ortaya çıkmasında bir çok etken (faktör) rol oynar. Acaba faktörlerin değişimleri ve farklı kombinasyonları, sonucun görülmesi ya da görülmemesinde, oluşumun derecelendirilmesinde nasıl etkide bulunmaktadır? Normal dağılım varsayımı kurulamayan durumlarda sonucun ortaya çıkması-çıkarmaması, hafif-orta-ağır olarak belirlenmesine açıklayıcı değişkenlerin etkileri nasıl ortaya konabilir?

LOJİSTİK REGRESYON

Toplumda bazı kişilerde kalp hastalığı görülürken bazılarında görülmemektedir. Toplumda birçok yönden benzer özellik gösteren bireylerin bazılarında X hastalığı görülürken diğerlerinde görülmemektedir. Niçin? Hangi etken ya da etkenler ne düzeyde bu sonuçların ortaya çıkmasına etki etmektedirler? Bir olayın ortaya çıkmasında bu etkenlerin bir risk faktörü olduğu ve bu etkenlerden hangilerinin önemli risk faktörleri olduğu nasıl belirlenebilir?

LOJİSTİK REGRESYON

Yukarıda sayılan sorulara cevap vermek için verilerin Lojistik Regresyon Analizi ile analiz edilmesi gerekir.

Lojistik regresyon, bağımlı değişkenin tahmini değerlerini olasılık olarak hesaplayarak, olasılık kurallarına uygun sınıflama yapma imkanı veren bir istatistiksel yöntemdir. Lojistik regresyon tablolaştırılmış ya da ham veri setlerini analiz eden bir yöntemdir.

LOJİSTİK REGRESYON

Veri yapılarına göre kurulan lojistik modeller aşağıdaki gibi belirlenir.

İki değişkenli lojistik regresyon modeli;

$$P(Y) = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X)}}$$

Çokdeğişkenli lojistik regresyon modeli;

$$P(Y) = \frac{e^Z}{1 + e^Z}$$

Burada Z, bağımsız değişkenlerin doğrusal kombinasyonudur.

$$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

LOJİSTİK REGRESYON

Regresyon katsayılarının hesaplanması aşağıdaki gibi yapılır.

$$\ln\left(\frac{P(Y)}{Q(Y)}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$$

$$\frac{P(Y)}{Q(Y)} = e^{\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p} = e^{\beta_0} e^{\beta_1 X_1} e^{\beta_2 X_2} \dots e^{\beta_p X_p}$$

Burada $Q(Y)$, $Q(Y)=1-P(Y)$ olarak hesaplanır. Odds Ratio'nun $P(Y)/Q(Y)$ olarak hesaplandığını hatırlayacak olursak her bir parametrenin $\text{Exp}(\beta)$ değerleri OR değerleri olarak ele alınırlar. Böylece $\text{Exp}(\beta_p)$, Y değişkeninin X_p değişkeninin etkisi ile kaç kat daha fazla ya da yüzde kaç oranda fazla gözlenme olasılığına sahip olduğunu belirtir. β_p katsayısının önemliliği aynı zamanda $\text{OR}_p = \text{Exp}(\beta_p)$ 'nın da önemliliği olarak değerlendirilir.

LOJİSTİK REGRESYON

Lojistik regresyon analizinde üç temel yöntem vardır.

- ✓ İkili Lojistik Regresyon (BLOGREG, Binary Logistic Regression)
- ✓ Ordinal Lojistik Regresyon (OLOGREG, Ordinal Logistic Regression)
- ✓ İsimsel Lojistik Regresyon (NLOGREG, Nominal Logistic Regression)

İkili Lojistik Regresyon (BLOGREG) Analizi

İkili cevap içeren bağımlı değişkenlerle yapılan lojistik regresyon analizidir. Bir ya da daha fazla açıklayıcı değişken ile ikili cevap değişken arasındaki bağıntıyı ortaya koyar.

Açıklayıcı değişkenler ya faktör değişkenlerdir ya da ortak değişkendir (covariate). Faktör değişkenler kategorik isimsel ölçeklidirler, ortak değişkenler ise sürekli değişken olmalıdır.

İkili Lojistik Regresyon (BLOGREG) Analizi

Örnek: Yenidoğanın doğum ağırlığının <3.0 kg (doğum ağırlığı=1) ve $3.0+$ kg (doğum ağırlığı=2) olmasında annenin yaşı, boyu, kilosuna, sigara içip içmemesinin (içiyorsa sigara=1, içmiyorsa sigara=0) rolü araştırılmaktadır.

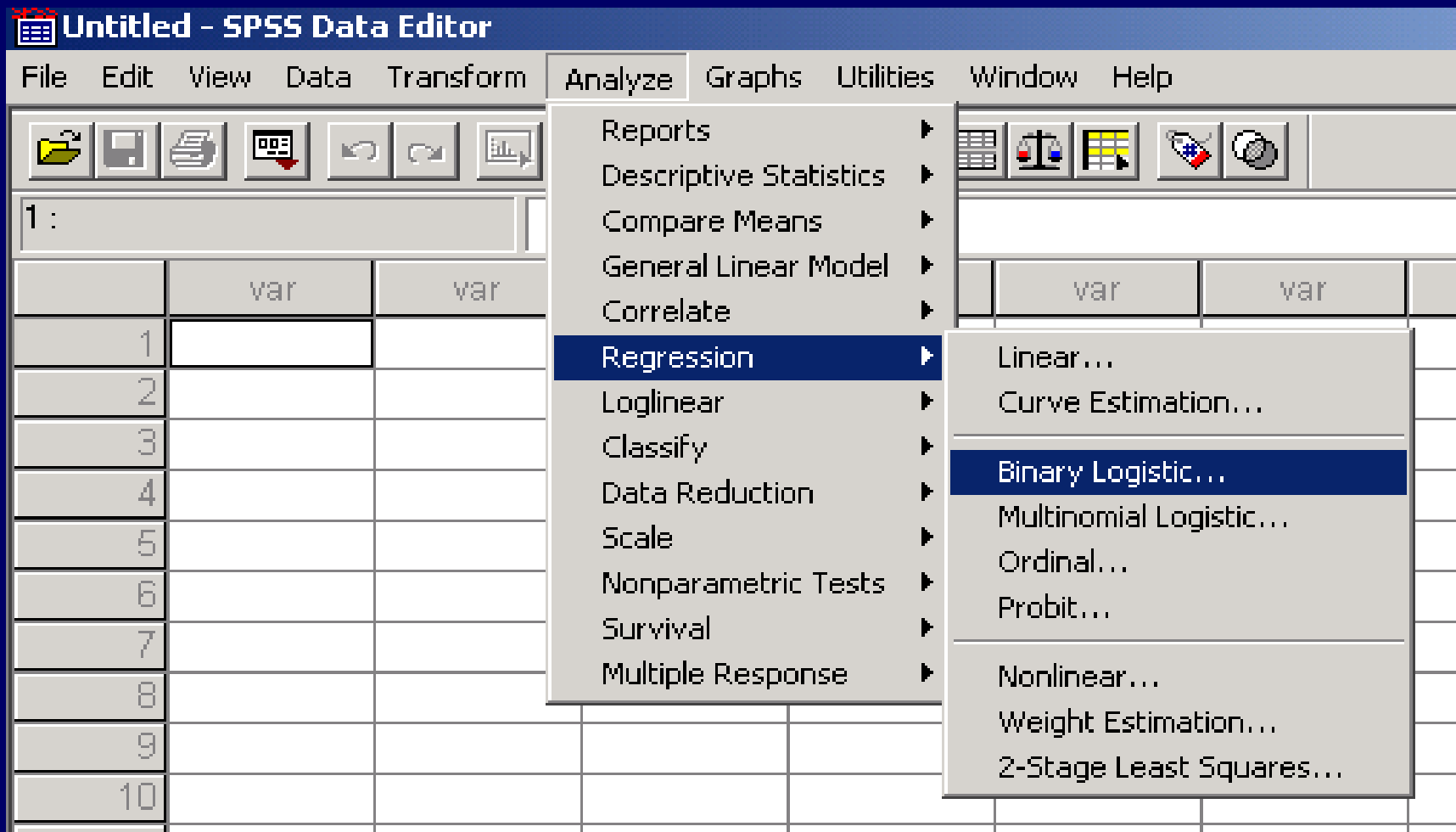
Bu amaçla rasgele seçilen 30 hamile kadınla ilgili bilgiler tabloda verilmiştir. Bebeklerin doğum ağırlığının Düşük (<3.0 kg) ya da Normal ve üstü ($3+$ kg) olmasında annenin yaşının, boyunun, kilosunun ve sigara alışkanlığının rolü var mıdır?

İkili Lojistik Regresyon (BLOGREG) Analizi

Birey	DOG_AGR	YAS	SIGARA	BOY	KILO
1	1	33	1	168	58.78
2	1	19	1	170	57.79
3	1	29	0	164	56.80
4	2	27	0	151	41.42
5	2	30	1	166	65.15
6	1	18	0	168	55.83
7	2	21	1	157	56.72
8	1	13	1	166	55.09
9	2	33	0	170	60.84
10	1	28	1	157	60.79
11	1	32	1	165	66.68
12	2	28	0	157	49.58
13	2	23	0	162	57.15
14	2	32	1	165	58.49
15	1	28	0	177	78.48
16	2	24	1	170	62.59
17	1	28	0	172	61.98
18	1	24	0	159	66.21
19	2	24	0	155	58.47
20	2	34	1	164	66.79
21	2	24	0	165	51.70
22	2	30	1	166	59.11
23	1	30	1	164	55.02
24	2	26	1	161	55.58
25	2	28	0	161	67.98
26	1	11	0	162	64.88
27	1	24	0	171	61.20
28	2	28	0	163	60.92
29	2	26	1	165	56.74
30	1	34	0	160	56.90

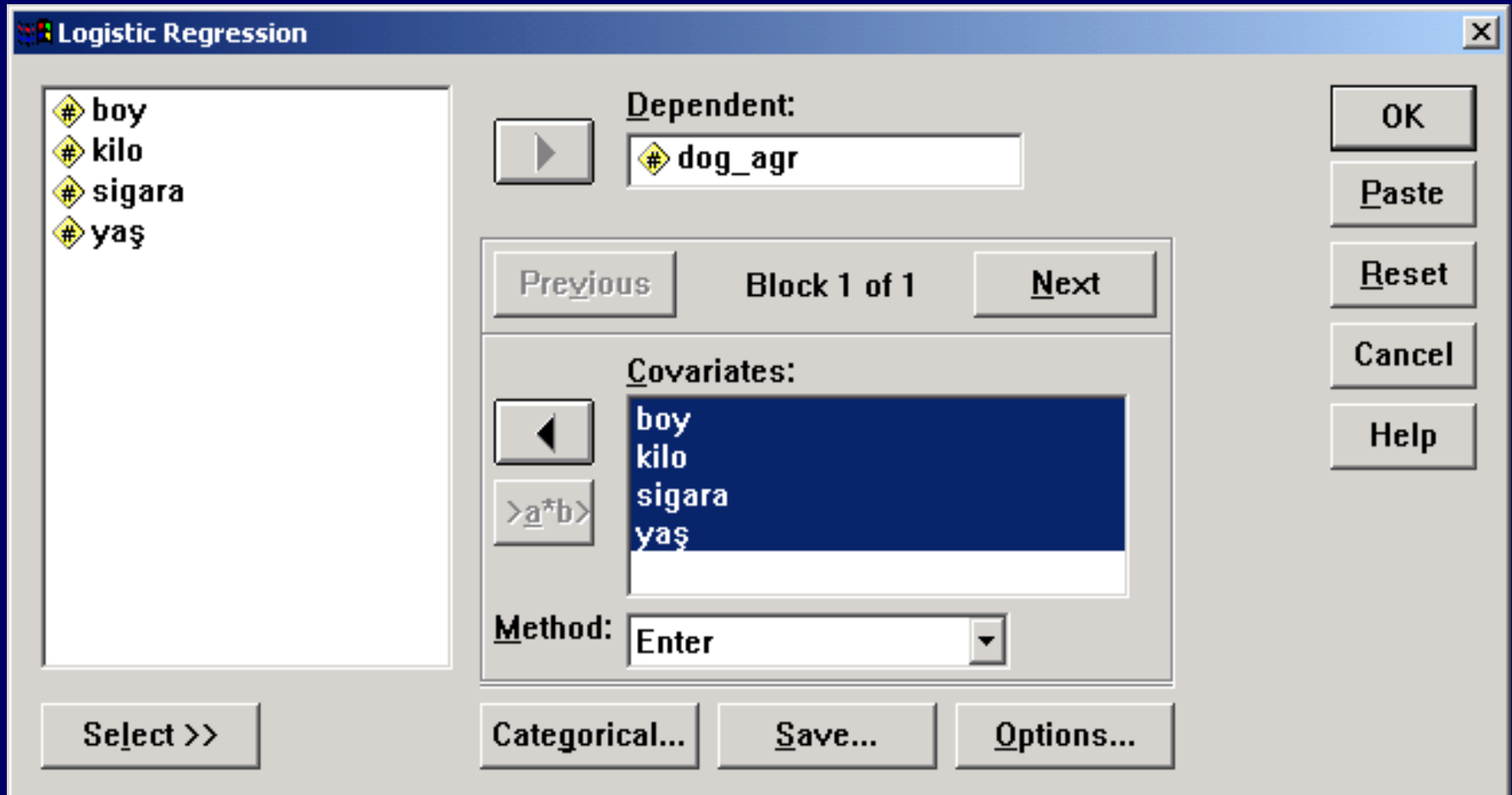
İkili Lojistik Regresyon (BLOGREG) Analizi

Örnek verilerine SPSS’de BLOGREG analizi uygulamak için *Analyze>Regression>Binary-Logistic* seçenekleri aşağıdaki ekrandaki gibi seçilir.



İkili Lojistik Regresyon (BLOGREG) Analizi

Seçimler yapıldığında Lojistik regresyon işlem penceresi görüntülenir. **Dependent** alanına **dog_agr** girilir. Diğer açıklayıcı değişkenler **Covariates** alanına taşınırlar. **Categorical** seçeneği tıklanır.



The screenshot shows the "Logistic Regression" dialog box. On the left, a list of variables includes "# boy", "# kilo", "# sigara", and "# yaş". The "Dependent:" field contains "# dog_agr". The "Covariates:" field contains "boy", "kilo", "sigara", and "yaş". The "Method:" dropdown is set to "Enter". The "Block 1 of 1" indicator is visible. Buttons for "Previous", "Next", "OK", "Paste", "Reset", "Cancel", and "Help" are present. At the bottom, there are buttons for "Select >>", "Categorical...", "Save...", and "Options...".

Logistic Regression

boy
kilo
sigara
yaş

Dependent:
dog_agr

Previous Block 1 of 1 Next

Covariates:
boy
kilo
sigara
yaş

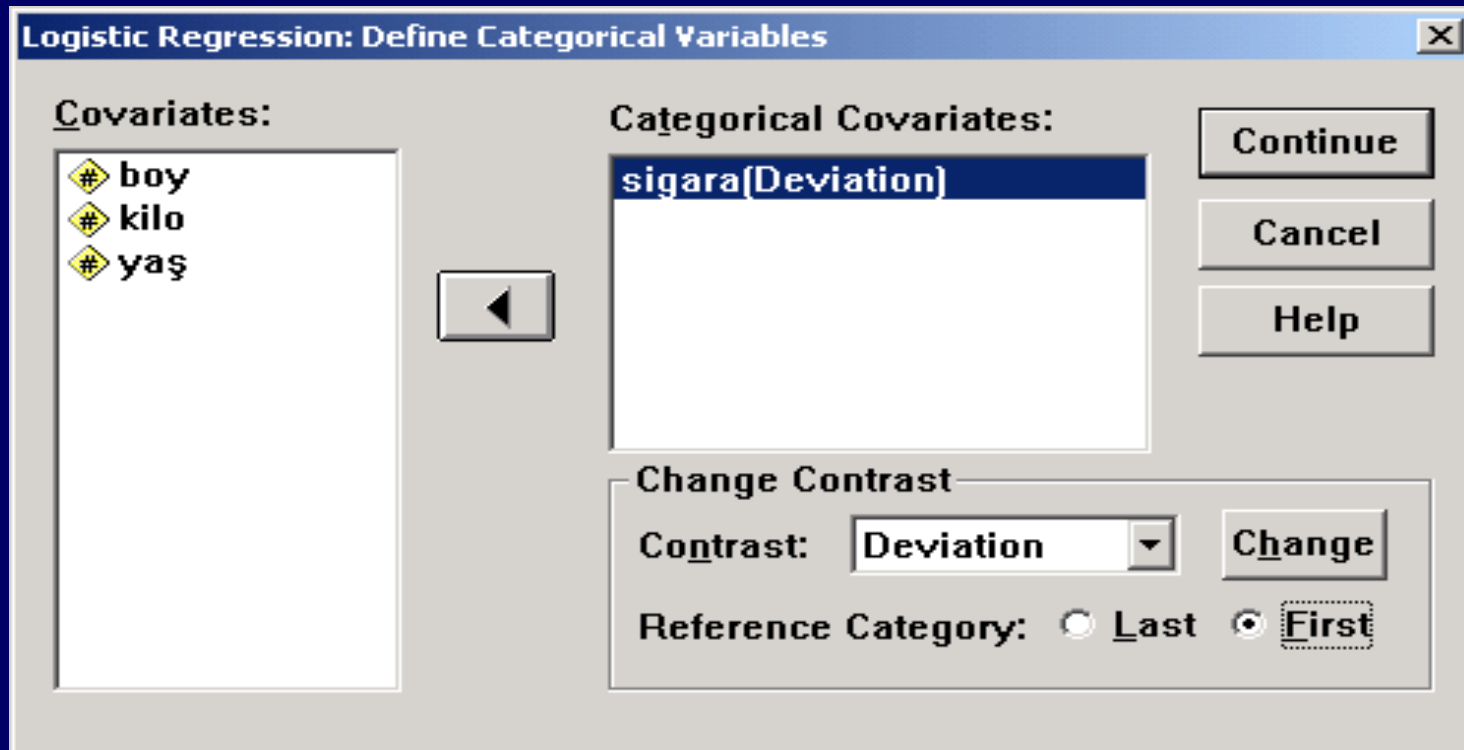
Method: Enter

Select >> Categorical... Save... Options...

OK
Paste
Reset
Cancel
Help

İkili Lojistik Regresyon (BLOGREG) Analizi

SPSS, faktörleri kategorik ortak değişken olarak modele alır.



Değişkenler içinde sadece sigara alışkanlığı kategorik olduğu için sigara değişkeni **Categorical Variables** alanına alınır. Değişkenin **Contrast** tipi belirlenir. Kurulu seçenek **Deviation**'dur.

İkili Lojistik Regresyon (BLOGREG) Analizi

Çıktıda yer alması istenilen bilgilerin belirlenmesi için **Options** seçeneği tıklanır ve aşağıdaki ekran görüntülenir. Bu ekranda uygun seçimler yapılır.

Logistic Regression: Options

Statistics and Plots

- ☐ Classification plots
- ☐ Hosmer-Lemeshow goodness-of-fit
- ☐ Casewise listing of residuals
- ☒ Correlations of estimates
- ☐ Iteration history
- ☐ CI for exp(B): 95 %
- ☒ Outliers outside 2 std. dev.
- ☐ All cases

Display

- ☐ At each step
- ☒ At last step

Probability for Stepwise

Entry: .05 Removal: .10

Classification cutoff: .5

Maximum Iterations: 20

☒ Include constant in model

Continue
Cancel
Help

İkili Lojistik Regresyon (BLOGREG) Analizi

SPSS lojistik regresyon analizinde hesaplanan ve daha sonraki analizlerde kullanılmak üzere veri sayfasına kaydedilmesi istenilen istatistikleri belirlemek için *Save* seçeneği tıklanır ve aşağıdaki ekran görüntülenir. Bu ekranda dosyaya kaydedilmesi istenilen değerler belirlenir.

Logistic Regression: Save New Variables

Predicted Values

- ☒ Probabilities
- ☐ Group membership

Influence

- ☐ Cook's
- ☐ Leverage values
- ☐ DfBeta(s)

Residuals

- ☐ Unstandardized
- ☐ Logit
- ☐ Studentized
- ☐ Standardized
- ☐ Deviance

Continue

Cancel

Help

İkili Lojistik Regresyon (BLOGREG) Analizi

İlgili tüm tanımlamalar yapıldıktan sonra *OK* tıklanır. Özetlenmiş ve düzenlenmiş sonuçlar aşağıdaki gibi elde edilir.

Dependent Variable Encoding

Original Value	Internal Value
1,00	0
2,00	1

Case Processing Summary

Unweighted Cases ^a		N	Percent
Selected Cases	Included in Analysis	30	100,0
	Missing Cases	0	,0
	Total	30	100,0
Unselected Cases		0	,0
Total		30	100,0

a. If weight is in effect, see classification table for the total number of cases.

İkili Lojistik Regresyon (BLOGREG) Analizi

Categorical Variables Coding^a

	Frequency	Parameter
		(1)
SIGARA ,00	16	1,000
1,00	14	-1,000

- a. Categorical variable(s) with only 0, and 1 values have been recoded using the above coding scheme. Parameter estimates are not the same as for indicator (0,1) variables.

Classification Table^{a,b}

Observed			Predicted		
			DOG AGR		Percentage Correct
			1,00	2,00	
Step 0	DOG_AGR	1,00	0	14	,0
		2,00	0	16	100,0
Overall Percentage					53,3

a. Constant is included in the model.

b. The cut value is ,500

İkili Lojistik Regresyon (BLOGREG) Analizi

Variables in the Equation

	B	S. E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	,134	,366	,133	1	,715	1,143

Omnibus Tests of Model Coefficients

	Chi-square	df	Sig.
Step 1 Step	4,315	4	,365
Block	4,315	4	,365
Model	4,315	4	,365

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	37,140	,134	,179

Variables not in the Equation

	Score	df	Sig.
Step 0 Variables			
YAS	,332	1	,565
SIGARA(1)	3,274	1	,070
BOY	,699	1	,403
KILO	,758	1	,384
Overall Statistics	4,126	4	,389

Classification Table^a

		Predicted		
		DOG_AGR		Percentage Correct
		1,00	2,00	
Step 1	Observed			
	DOG_AGR 1,00	9	5	64,3
	DOG_AGR 2,00	6	10	62,5
	Overall Percentage			63,3

a. The cut value is ,500

Variables in the Equation

	B	S. E.	Wald	df	Sig.	Exp(B)
Step 1 ^a YAS	-,019	,071	,074	1	,786	,981
SIGARA(1)	,681	,398	2,921	1	,087	1,975
BOY	-,027	,084	,103	1	,748	,973
KILO	-,042	,074	,326	1	,568	,958
Constant	7,564	12,479	,367	1	,544	1926,737

a. Variable(s) entered on step 1: YAS, SIGARA, BOY, KILO.

İkili Lojistik Regresyon (BLOGREG) Analizi

Çıktı incelendiğinde bütün açıklayıcı değişkenlerin *dog_agr* üzerindeki etkisinin önemsiz olduğu, düşük doğum ağırlığını belirleyici bir risk faktörü olmadıkları gözlenmektedir.

Sigaranın düşük doğum ağırlıklı bebek doğum olasılığını 1.9755 kat artırdığı bulunmuş fakat bu önemli bir risk faktörü olarak sayılmasına yetmemiştir. (OR=1.9755, P=0.0874^{ns}).

P(*dog_agr*)'nin tahmini için model sabit ve açıklayıcı değişkenlerin regresyon katsayıları kullanılarak aşağıdaki gibi yazılır.

$$P(Y)=1/(1+e^{-Z})$$

$$Z=7.564-0.0269 \text{ boy}-0.0424 \text{ kilo}+0.6808 \text{ sigara}-0.0193 \text{ yaş}$$

Model önemli olarak Y'nin tahminine katkıda bulunmaktadır. Risk faktörlerinin (açıklayıcı değişkenlerin) tek başlarına önemli rolleri yok ise de ortak olarak Y'nin değişimini belirleme etkinliğine sahip görülmektedir.

Sıralı Lojistik Regresyon (OLOGREG) Analizi

Sıralı Lojistik regresyon (OLOGREG) yöntemi, cevap değişkeninin üç ve daha fazla kategori içerdiği ve değerlerin sıralı ölçekle elde edildiği durumlarda; cevap değişkeni ile açıklayıcı değişkenler arasındaki neden sonuç ilişkilerini ortaya koymaya yönelik bir yöntemdir.

Cevap değişkeninin sıralı ölçekli olduğu durumlarda uygulanan bir yöntemdir. Sıralı ölçekli cevap değişkeni, en az üç kategoride gözlenen değerler içermelidir.

Sıralı Lojistik Regresyon (OLOGREG) Analizi

Sıralı ölçekli veriler kodlanırken ya da isimsel olarak kategorileri belirlendiğinde cevapların doğal sıralama yapısında olması gerekir. Örneğin, hastalık şiddeti söz konusu ise, hafif<orta<ağır olarak kategoriler belirlenmelidir. Hasta bireyin hastalık şiddeti bu kategori yapısı içinde doğru olarak değerlendirilmelidir.

Bir oluşuma karşı beğeni sıralaması sözkonusu ise; kategoriler, beğenmedim<az beğendim<beğendim<çok beğendim biçiminde sıralanmalıdır.

Bu isimsel değerlerin kod değerleri de aynı büyüklük sıralamasını izlemesi gerekir (1<2<3<4 gibi). OLOGREG analizi isimsel kategoriler yerine kod değerleri de işlemektedir.

İsimsel Lojistik Regresyon (NLOGREG) Analizi

İsimsel Lojistik regresyon (NLOGREG) yöntemi, cevap değişkeninin üç ve daha fazla kategori içerdiği ve değerlerin isimsel ölçekle elde edildiği durumlarda; cevap değişkeni ile açıklayıcı değişkenler arasındaki neden sonuç ilişkilerini ortaya koymaya yönelik bir yöntemdir.

Örneğin bir meslek dalları tercihlerinde sınıflar; Mühendislik, Bankacılık, Tıp, Turizm, vb. isimsel olarak belirlenebilirler.

İsimsel Lojistik Regresyon (NLOGREG) Analizi

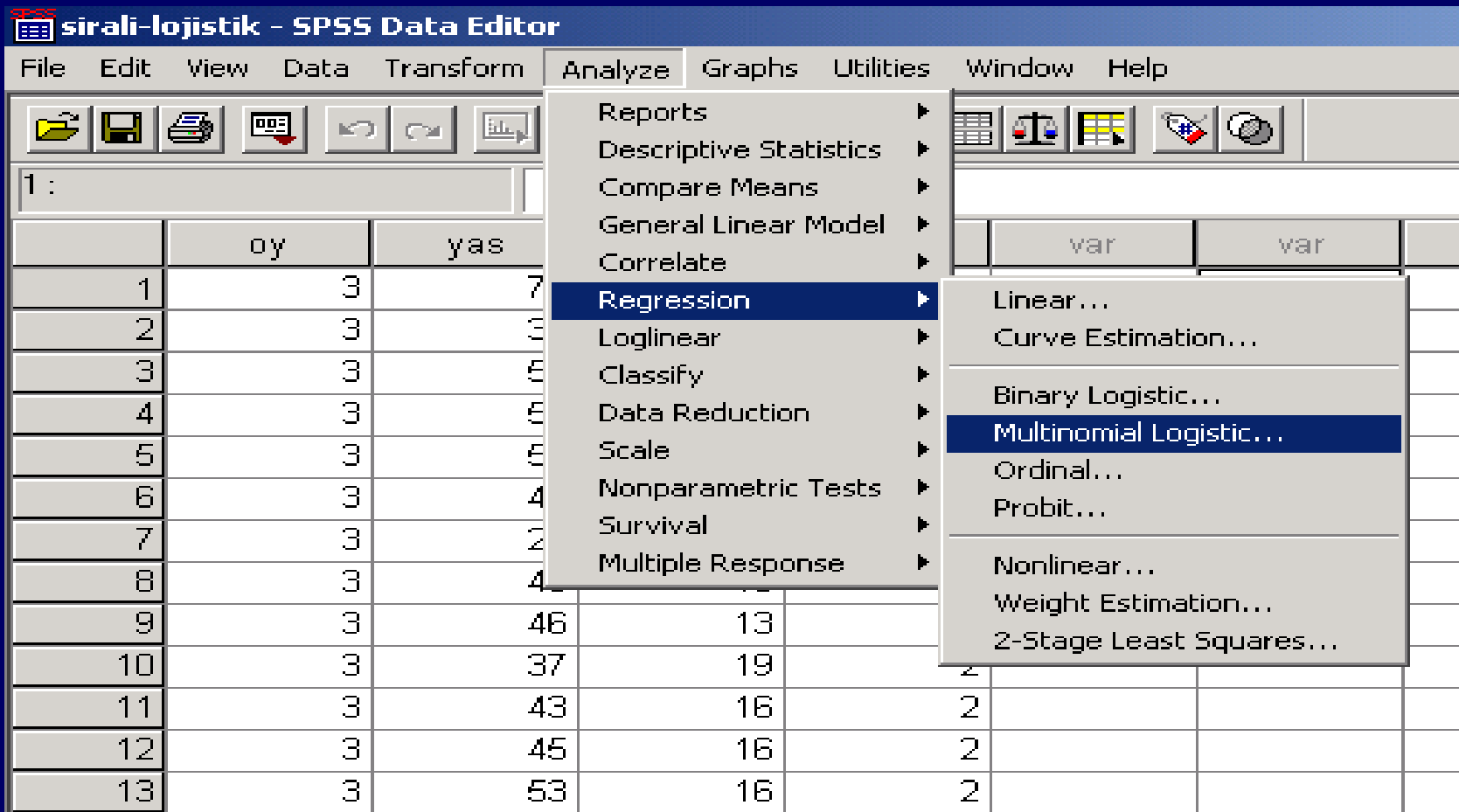
Örnek: ABD’de 1992 yılındaki başkanlık seçimleri için yapılan kamuoyu araştırmasında 1847 kişiye; Bush, Perot ve Clinton için oyları, yaşları, eğitim süreleri (yıl olarak) ve cinsiyetleri sorulmuştur. Elde edilen verilerin bir kısmı aşağıdaki tabloda verilmiştir. Burada oy sütunundaki verilerde “1 Bush” için, “2 Perot” ve “3 Clinton” için oy verileceğini göstermektedir. Cinsiyet sütununda ise “1 erkek”, “2 kadın”ı göstermektedir.

İsimsel Lojistik Regresyon (NLOGREG) Analizi

Birey	Oyu	Yaşı	Eğitim Süresi	Cinsiyeti
1	3	79	12	1
2	3	32	17	1
3	3	50	6	2
4	3	56	8	2
5	3	51	17	2
6	3	48	12	1
7	3	29	13	2
8	3	40	13	2
9	3	46	13	2
10	3	37	19	2
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
1847	3	33	12	2

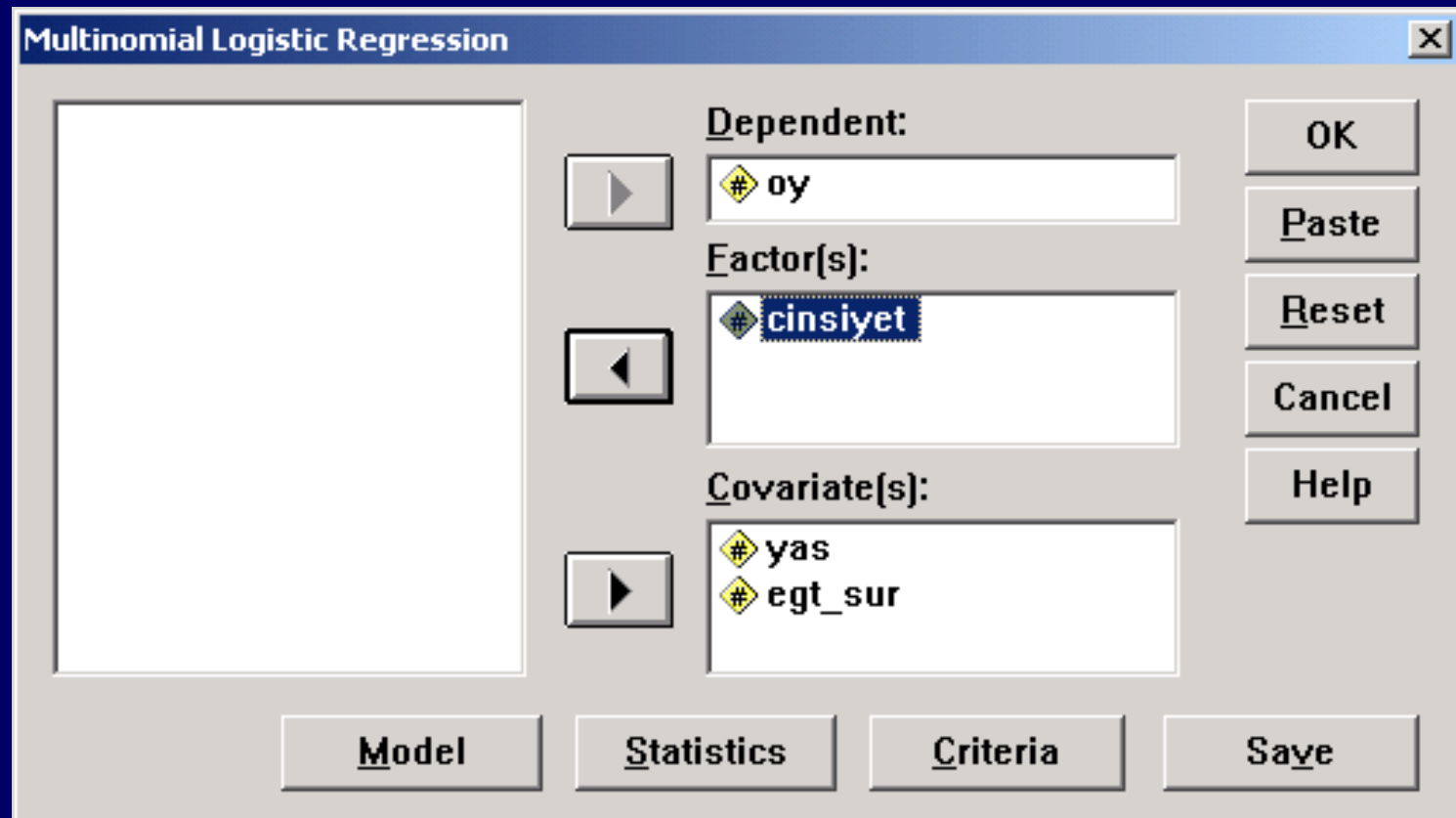
İsimsel Lojistik Regresyon (NLOGREG) Analizi

Veriler SPSS paket programında ayrı ayrı sütunlara girilir. Analyze > Regression > Multinomial Logistic Regression tıklanır.



İsimsel Lojistik Regresyon (NLOGREG) Analizi

Dependent alanına bağımlı değişken olan *oy* taşınır. Factor(s) alanına ise kategorik veri olan *cinsiyet* ve Covariate(s) alanına ise sürekli değişken olan *yaş* ve *eğitim süresi* taşınır. OK tıklanır.



İsimsel Lojistik Regresyon (NLOGREG) Analizi

Parameter Estimates

OY		B	Std. Error	Wald	df	Sig.	Exp(B)
Bush	Intercept	-,536	,346	2,398	1	,122	
	EGT_SUR	1,041E-02	,019	,311	1	,577	1,010
	YAS	-2,19E-03	,003	,450	1	,502	,998
	[CINSIYET=1]	,429	,104	17,013	1	,000	1,536
	[CINSIYET=2]	0 ^a	,	,	0	,	,
Perot	Intercept	,270	,475	,322	1	,570	
	EGT_SUR	-1,63E-02	,027	,377	1	,539	,984
	YAS	-3,47E-02	,005	48,075	1	,000	,966
	[CINSIYET=1]	,742	,141	27,618	1	,000	2,100
	[CINSIYET=2]	0 ^a	,	,	0	,	,

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

İsimsel Lojistik Regresyon (NLOGREG) Analizi

Çıktı sonucunda; eğitim süresinin, Clinton yerine Bush ve ya Perot'un seçilmesinde önemli bir etkisi yoktur (Bush için $p=0,577>0,05$, Perot için $p=0,539>0,05$).

Yaşın ise Clinton'nın yerine Bush'un seçilmesinde önemsiz fakat Clinton'nın yerine Perot'un seçilmesinde önemli bir etkisinin olduğu ortaya çıkmıştır (Bush için $p=0,502>0,05$, Perot için $p=0,000<0,001$).

Yaş arttıkça Clinto'nun yerine Perot'un seçilme olasılığı azalmaktadır. Hesaplanan ODDS ratio oranı 0,966'dır. Yaş bir birim arttıkça Clinton'nun yerine Perot'un seçilme şansı 0,966 kat artmaktadır.

İsimsel Lojistik Regresyon (NLOGREG) Analizi

Cinsiyetin ise Clinton'nın yerine Bush'un ve ya Perot'un seçilmesinde önemli bir etkisinin olduğu ortaya çıkmıştır (Bush için $p=0,000<0,001$ ve Perot için $p=0,000<0,001$).

Bush için cinsiyet değişkeninin ODDS ratio değeri 1,536'dır. Buda erkeklerin Clinton'nun yerine Bush'u seçme şansı kadınlarinkinden 1,536 kat fazladır.

Yine aynı şekilde Perot için Cinsiyet değişkeninin ODDS ratio değeri 2,1'dir. Buda erkeklerin Clinton'un yerine Perot'u seçme şansı kadınlara oranla 2,1 kat daha fazladır.