



Regresyon Analizi

Yaşar Tonta
H.Ü. BBY

tonta@hacettepe.edu.tr

yunus.hacettepe.edu.tr/~tonta/courses/fall2008/sb5002/



- Not: Sunuş slaytları G.A. Morgan, O.V. Griego ve G.W. Gloeckner'in *SPSS for Windows: An Introduction to Use and Interpretation in Research* (Lawrence Erlbaum Associates, 2001) adlı eserinin 11. Bölümü ile Andy Field'in *Discovering Statistics Using SPSS* (Sage, 2005) adlı eserinin 5. ve 6. bölümlerinden ve kitapta kullanılan veri setlerinden yararlanılarak hazırlanmıştır.



- İki veya daha fazla değişken arasındaki ilişkileri ölçmek için kullanılır.
- Hem tanımlayıcı hem de çıkarımsal istatistik sağlar.
- Şehir nüfusu ile suç oranı arasındaki ilişki
- Beden eğitimi derslerinde öğretmen etkinliği
- $F = b_0 + b_1I + b_2x_1 + b_3x_2 + b_4x_3 + e$
- F = öğrenci son notu, b = regresyon ağırlığı, I = Başlangıç notu, x_1 =rehberlik ve destek uygulama, x_2 =içerik bilgisi, x_3 =işle ilgili bilgi, e =kalan ya da analiz edilen mevcut değişkenlerle açıklanamayan varyans.



- Basit doğrusal regresyon bize normal dağılmış, hakkında aralıklı/oranlı ölçekle veri toplanmış iki değişken arasında doğrusal ilişki olup olmadığını test etme olanağı verir. Değişkenlerden biri tahmin, biri sonuç değişkenidir. Örneğin, hsb2turkce veri dosyasını kullanarak yazma ve okuma puanları arasındaki ilişkiye bakalım. Başka bir deyişle öğrencilerin okuma puanlarından yazma puanlarını tahmin etmeye çalışalım.

Önce hipotez kuralım



- Boş Hipotez (H_0): “Öğrencilerin okuma ve yazma puanları arasında doğrusal bir ilişki yoktur.
- Araştırma Hipotezi (H_1): “Öğrencilerin okuma ve yazma puanları arasında doğrusal bir ilişki vardır.” (çift kuyruk testi).
 - $H_0 : \mu = \mu_0$
 - $H_1 : \mu \neq \mu_0$ (çift kuyruk testi)
- Boş hipotezleri büyüktür/küçüktür diye de kurabilirsiniz. O zaman tek kuyruk (büyükse sol, küçükse sağ) test yapılır.
- Örneğin, H_0 : “Öğrencilerin okuma puanları yüksekse yazma puanları da yüksektir.”
- H_1 : “Öğrencilerin okuma puanları yüksekse yazma puanları düşüktür.”
 - $H_0 : \mu > \mu_0$
 - $H_1 : \mu < \mu_0$ (sol kuyruk testi)



Mönüden:

- Analyze -> regression-> linear'ı seçin
- Yazma puanını bağımlı, okuma puanını bağımsız değişken olarak seçin.
- OK'e tıklayın

Basit doğrusal regresyon test sonucu



Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	okuma puani ^a	.	Enter

- a. All requested variables entered.
b. Dependent Variable: yazma puani

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,597 ^a	,356	,353	7,625

- a. Predictors: (Constant), okuma puani

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6367,421	1	6367,421	109,521	,000 ^a
	Residual	11511,454	198	58,139		
	Total	17878,875	199			

- a. Predictors: (Constant), okuma puani
b. Dependent Variable: yazma puani

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	23,959	2,806		8,539	,000
	okuma puani	,552	,053	,597	10,465	,000

- a. Dependent Variable: yazma puani



- Yazma puanıyla okuma puanı arasında pozitif (0,552) bir ilişki var. t- değerinden bu ilişkinin istatistiksel açıdan anlamlı olduğunu görüyoruz (t = 10,47, p =0,000).
- Okuma ile yazma arasında istatistiksel açıdan anlamlı pozitif doğrusal bir ilişki vardır.
- Boş hipotez reddedilir
- Bu ilişki için basit doğrusal regresyon formülü:

$$\text{Yazma puanı} = 23,959 + 0,552 * \text{okuma puanı}$$

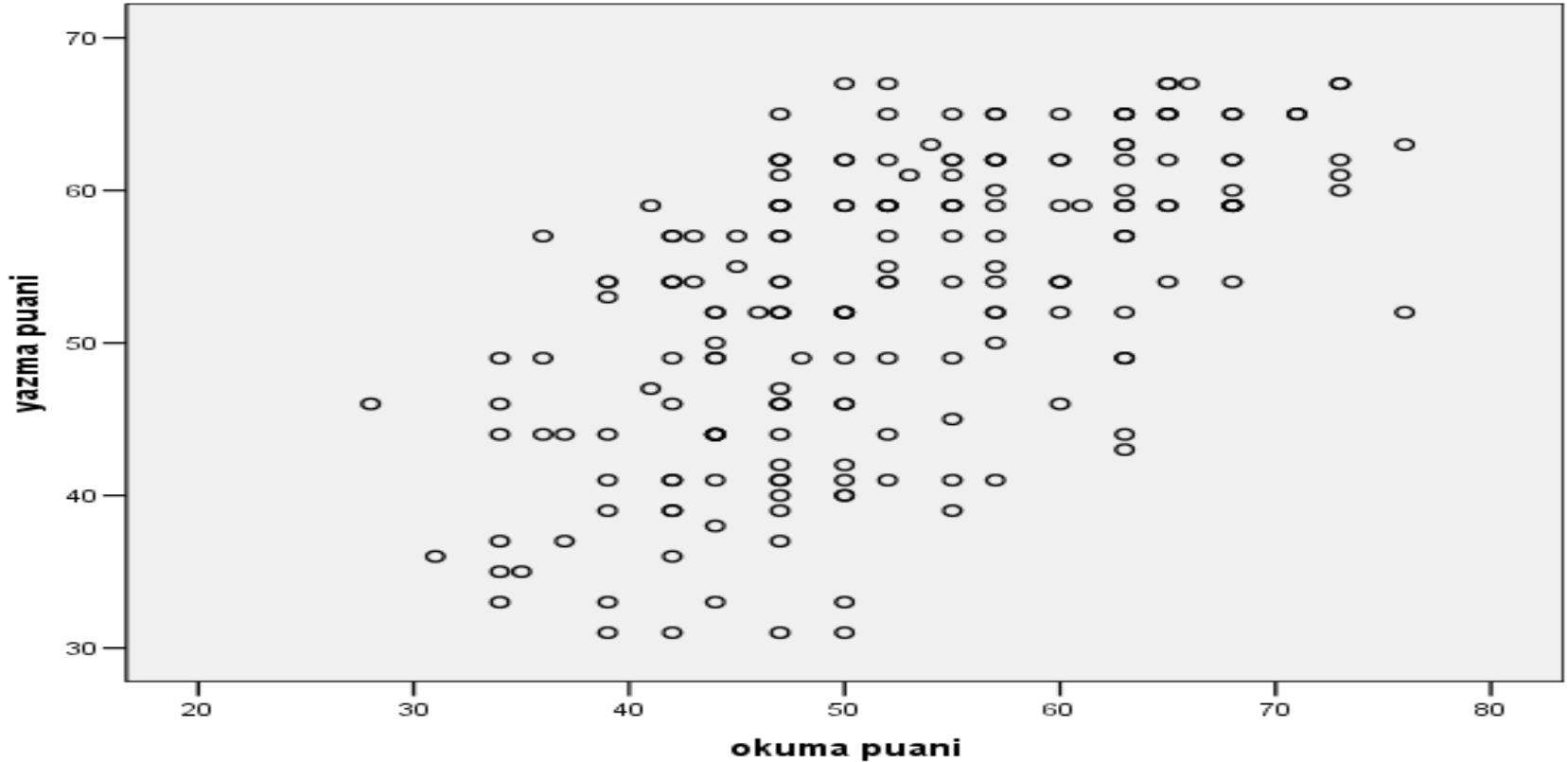
Saçılım grafiği



Nitekim bu pozitif doğrusal ilişkiyi

Graphs → Scatterplot → Simple Scatter'ı seçip

x eksenine okuma puanı, y eksenine yazma puanını atayarak aşağıdaki saçılım grafiğinde görebilirsiniz.



Çoklu Regresyon Analizi



- Basit regresyona çok benzer. Çoklu regresyon denkleminde birden fazla tahmin değişkeni vardır.
- Formül

$$Y_i = (b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n) + e_i$$

biçimini alır.

Y bağımlı değişken, b_0 regresyon eğrisinin y eksenini kesim noktası, b_1 ilk tahmin değişkeninin X_1 katsayısı, b_2 ikinci tahmin değişkeninin X_2 katsayısı, ... e_i ise i'inci denek için Y'nin tahmin edilen değeriyle gözlenen değeri arasındaki farktır.



- Birçok tahmin değişkeni olan karmaşık bir model kurmak istiyorsak hangi tahmin değişkenlerini seçeceğimize nasıl karar vereceğiz?
- Hangi tahmin değişkenlerinin seçildiği ve bunların modele nasıl girildiği önemli.
- Rastgele çok sayıda tahmin değişkeni seçmek doğru değil.
- Peki tahmin değişkenlerini modele nasıl gireceğiz?

Tahmin deęişkenlerini regresyon modeline ekleme yolları



- Hiyerarşik ekleme (Blockwise entry): (Önceki çalışmaların sonuçlarına dayanarak) en önemli tahmin deęişkeni önce girilir.
- Zorla ekleme (Enter): Tüm tahmin deęişkenleri eş zamanlı olarak modele girilir
- Adım adım ekleme: Tamamen matematiksel ölçütlere göre girilir. Bilgisayar baęımlı deęişkendeki deęişimi en fazla açıklayan tahmin deęişkenini bulur, sonra geri kalan deęişimi en fazla açıklayan tahmin deęişkenini bulur, vs. (adım adım eklemede backward yöntemini seçmek daha uygun).

Regresyon modelim ne kadar doğru?



- İki önemli soru:
 - Model gözlenen verilere ne kadar iyi uyuyor veya model az sayıda vakadan etkileniyor mu?
 - Model diğer örneklere genellenebilir mi?
- İlk sorunun yanıtı için uç değerlere (outliers) ve model tarafından tahmin edilen değerlerle gözlenen değerler arasındaki farka (residuals) bakılmalı.
- Uç değerler regresyon eğrisini gerçekte olduğundan farklı gösterir.

Sayıtlılar (assumptions)



- Tüm tahmin değişkenleri aralıklı/oranlı (nicel) veya kategorik olmalı (iki kategori), çıktı değişkeni nicel, sürekli ve sınırsız olmalı (çıktı 1-10 arasında değişiyorsa ama toplanan veri 3-7 arasındaysa veri sınırlı demektir).
- Tahmin değişkenlerinin varyansı 0 (sıfır) olmamalı.
- Tahmin değişkenleri arasında mükemmel doğrusal ilişkiler olmamalı (o zaman aralarında doğrusal ilişki olan değişkenler için b değeri aynı olur; R'nin büyüklüğünü sınırlar; hangi değişkenin önemli olduğunu söyleyemeyiz; regresyon katsayıları –b değerleri-örneklemeden örnekleme değişir)
- Hatalar normal dağılmalı (yani modelle gözlenen veriler arasındaki farklar sıfır ya da sıfıra yakın olmalı)
- İlişki doğrusal olmalı



- Modeldeki her tahmin değişkeni için en az 10-15 ölçüm (veri) olmalı. Yani beş değişken varsa 50-75 ölçüm olmalı
- Örneklem ne kadar büyükse o kadar iyi.
- R, tahmin değişkeni sayısına (k) ve örneklem büyüklüğüne (N) bağlı ($R = k / (N - 1)$). Örneğin, $R = 6 / (21 - 1) = 0,3$ (orta düzeyde etki); $R = 6 / (100 - 1) = 0,06$.



- Bir plak şirketi yöneticisi bir albüme harcanan reklam bütçesiyle o albüm piyasaya çıkmadan önceki hafta albümün radyoda çalınma sayısının ve albümü yapan grubun çekiciliğinin albüm satışları üzerindeki etkisini bilmek istiyor.
- $\text{Albüm satışları}_i = (b_0 + b_1 \text{Reklam bütçesi}_i + b_2 \text{Radyoda çalınma sayısı}_i + b_3 \text{Grubun çekiciliği}_i) + e_i$

SPSS'te Çoklu Regresyon



The screenshot shows the SPSS Data Editor window with a data table and the Linear Regression dialog box open. The data table has columns for 'adverts', 'sales', 'airplay', and 'attract'. The Linear Regression dialog box is configured with 'Record Sales (thousand)' as the dependent variable and 'Advertising Budget (thousand)' as the independent variable. The method is set to 'Enter'.

	adverts	sales	airplay	attract	var	var	var	var	var	var	var	var	var	var	var	var
1	10,26	330,00	43,00	10,00												
2	985,69	120,00	28,00	7,00												
3	1445,56	360,00	35,00	7,00												
4	1188,19	270,00	33,00	7,00												
5	574,51	220,00	44,00	5,00												
6	568,95	170,00	19,00	5,00												
7	471,81	70,00	20,00	1,00												
8	537,35	210,00	22,00	9,00												
9	514,07	200,00	21,00	7,00												
10	174,09	300,00	40,00	7,00												
11	1720,81	290,00	32,00	7,00												
12	611,48	70,00	20,00	2,00												
13	251,19	150,00	24,00	8,00												
14	97,97	190,00	38,00	6,00												
15	406,81	240,00	24,00	7,00												
16	265,40	100,00	25,00	5,00												
17	1323,29	250,00	35,00	5,00												
18	196,65	210,00	36,00	8,00												
19	1326,60	280,00	27,00	8,00												
20	1380,69	230,00	33,00	8,00												
21	792,35	210,00	33,00	7,00												
22	957,17	230,00	28,00	6,00												
23	1789,66	320,00	30,00	9,00												
24	656,14	210,00	34,00	7,00												
25	613,70	230,00	49,00	7,00												

Analyze->Regression->Linear'ı seçin; Yöntem olarak tüm tahmin değişkenlerinin eş zamanlı olarak modele girildiği varsayılan seçenek Enter'ı seçin. Geçmiş araştırmalar reklam bütçesiyle albüm satışları arasında ilişki olduğunu gösterdiğinden bağımsız değişkene reklam bütçesini yerleştirin. İlk bloğu tanımladıktan sonra "Next"e tıklayın.



Record2 - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : adverts 10,256

	adverts	sales	airplay	attract	var	var	var	var	var	var	var	var	var	var	var	var	var
1	10,26	330,00	43,00	10,00													
2	985,69	120,00	28,00	7,00													
3	1445,56	360,00	35,00	7,00													
4	1188,19	270,00	33,00	7,00													
5	574,51	220,00	44,00	5,00													
6	568,95	170,00	19,00	5,00													
7	471,81	70,00	20,00	1,00													
8	537,35	210,00	22,00	9,00													
9	514,07	200,00	21,00	7,00													
10	174,09	300,00	40,00	7,00													
11	1720,81	290,00	32,00	7,00													
12	611,48	70,00	20,00	2,00													
13	251,19	150,00	24,00	8,00													
14	97,97	190,00	38,00	6,00													
15	406,81	240,00	24,00	7,00													
16	265,40	100,00	25,00	5,00													
17	1323,29	250,00	35,00	5,00													
18	196,65	210,00	36,00	8,00													
19	1326,60	280,00	27,00	8,00													
20	1380,69	230,00	33,00	8,00													
21	792,35	210,00	33,00	7,00													
22	957,17	230,00	28,00	6,00													
23	1789,66	320,00	30,00	9,00													
24	656,14	210,00	34,00	7,00													
25	613,70	230,00	49,00	7,00													
26	313,36	250,00	40,00	8,00													
27	336,51	60,00	20,00	4,00													
28	1544,00	330,00	43,00	7,00													
29																	
30																	
31																	
32	377,93	80,00	19,00	8,00													

Linear Regression

Dependent: Record Sales (thousand)

Independent(s): No. of plays on Radio 1 per week, Attractiveness of Band (attract)

Method: Enter

Selection Variable:

Case Labels:

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

SPSS Processor is ready

start Adobe... Chapt... 17 I... 2 Mi... 3 Mi... Adobe... Windo... 4 Vol... 2 Mi... Micros... 3 SP... 18:55

İkinci modelde radyoda çalınma sayısı ile grubun çekiciliğini birlikte girin



The screenshot shows the SPSS Data Editor interface with a data table and a Linear Regression dialog box. The data table has columns for 'adverts', 'sales', 'airplay', and 'attract'. The Linear Regression dialog box is open, and the 'Linear Regression: Statistics' sub-dialog is also open, showing various options for regression analysis.

	adverts	sales	airplay	attract	var	var	var	var	var	var	var	var	var	var	var	var
1	10,26	330,00	43,00	10,00												
2	985,69	120,00	28,00	7,00												
3	1445,56	360,00	35,00	7,00												
4	1188,19	270,00	33,00	7,00												
5	574,51	220,00	44,00	5,00												
6	568,95	170,00	19,00	5,00												
7	471,81	70,00	20,00	1,00												
8	537,35	210,00	22,00	9,00												
9	514,07	200,00	21,00	7,00												
10	174,09	300,00	40,00	7,00												
11	1720,81	290,00	32,00	7,00												
12	611,48	70,00	20,00	2,00												
13	251,19	150,00	24,00	8,00												
14	97,97	190,00	38,00	6,00												
15	406,81	240,00	24,00	7,00												
16	265,40	100,00	25,00	5,00												
17	1323,29	250,00	35,00	5,00												
18	196,65	210,00	36,00	8,00												
19	1326,60	280,00	27,00	8,00												
20	1380,69	230,00	33,00	8,00												
21	792,35	210,00	33,00	7,00												
22	957,17	230,00	28,00	6,00												
23	1789,66	320,00	30,00	9,00												
24	656,14	210,00	34,00	7,00												
25	613,70	230,00	49,00	7,00												
26	313,36	250,00	40,00	8,00												
27	336,51	60,00	20,00	4,00												
28	1544,90	330,00	43,00	7,00												
29	68,95	150,00														
30	785,69	150,00														
31	125,63	180,00														
32	377,93	80,00	19,00	8,00												

Linear Regression dialog box options:

- Dependent: Advertising Budget (thou)
- Linear Regression: Statistics sub-dialog options:
 - Regression Coefficients:
 - Estimates
 - Confidence intervals
 - Covariance matrix
 - Model fit:
 - Model fit
 - R squared change
 - Descriptives
 - Part and partial correlations
 - Collinearity diagnostics
 - Residuals:
 - Durbin-Watson
 - Casewise diagnostics
 - Outliers outside: 2 standard deviations
 - All cases

Statistics'e tıklayıp yukarıdaki seçenekleri işaretleyin

Regresyon grafiđi



The screenshot shows the SPSS Data Editor window with a data table and the Linear Regression: Plots dialog box open. The data table has columns for 'adverts', 'sales', 'airplay', and 'attract'. The dialog box is titled 'Linear Regression' and 'Dependent: Advertising Budget (tho)'. It shows the 'Linear Regression: Plots' sub-dialog with the following options:

- DEPENDENT: *ZPRED
- *ZRESID
- *DRESID
- *ADJPRED
- *SRESID
- *SDRESID
- Scatter 1 of 1: Y: *ZRESID, X: *ZPRED
- Standardized Residual Plots: Histogram, Normal probability plot
- Produce all partial plots:

The dialog box also has buttons for 'OK', 'Paste', 'Reset', 'Cancel', 'Help', 'Continue', 'Cancel', 'Help', 'WLS >>', 'Statistics...', 'Plots...', 'Save...', and 'Options...'.

Plots'a tıklayıp yukarıdaki seçenekleri işaretleyin. Bu iki seçenek modelin Açıklayamadığı kısmın grafiđini verir (x ekseninde modele göre tahmin edilen deđer, y ekseninde gözlenen deđerle modelin tahmin ettiđi deđer arasındaki fark gösterilir.

Regresyon tanılama



Record2 - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : adverts 10,256

	adverts	sales	airplay	attract	var	var	var	var	var	var	var	var	var	var	var	var	var
1	10,26	330,00	43,00	10,00													
2	985,69	120,00	28,00	7,00													
3	1445,56	360,00	35,00	7,00													
4	1188,19	270,00	33,00	7,00													
5	574,51	220,00	44,00	5,00													
6	568,95	170,00	19,00	5,00													
7	471,81	70,00	20,00	1,00													
8	537,35	210,00	22,00	9,00													
9	514,07	200,00	21,00	7,00													
10	174,09	300,00	40,00	7,00													
11	1720,81	290,00	32,00	7,00													
12	611,48	70,00	20,00	2,00													
13	251,19	150,00	24,00	8,00													
14	97,97	190,00	38,00	6,00													
15	406,81	240,00	24,00	7,00													
16	265,40	100,00	25,00	5,00													
17	1323,29	250,00	35,00	5,00													
18	196,65	210,00	36,00	8,00													
19	1326,60	280,00	27,00	8,00													
20	1380,69	230,00	33,00	8,00													
21	792,35	210,00	33,00	7,00													
22	957,17	230,00	28,00	6,00													
23	1789,66	320,00	30,00	9,00													
24	656,14	210,00	34,00	7,00													
25	613,70	230,00	49,00	7,00													
26	313,36	250,00	40,00	8,00													
27	336,51	60,00	20,00	4,00													
28	1544,90	330,00	42,00	7,00													
29	68,95	150,00	35,00	8,00													
30	785,69	150,00	8,00	6,00													
31	125,63	180,00	49,00	7,00													
32	377,93	80,00	19,00	8,00													

Linear Regression: Save

Predicted Values

- Unstandardized
- Standardized
- Adjusted
- S.E. of mean predictions

Distances

- Mahalanobis
- Cook's
- Leverage values

Prediction Intervals

- Mean Individual

Confidence Interval: 95 %

Save to New File

- Coefficient statistics: File...

Export model information to XML file

Browse

Residuals

- Unstandardized
- Standardized
- Studentized
- Deleted
- Studentized deleted

Influence Statistics

- DfBeta(s)
- Standardized DfBeta(s)
- DfFit
- Standardized DfFit
- Covariance ratio

Continue Cancel Help

Save'e tıklayıp yukarıdaki seçenekleri işaretleyin. Bu seçenekler modeli daha iyi tanımamıza yardımcı olur. SPSS her seçenek için veri dosyasında ayrı bir sütun yaratır

Options



Record2 - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : adverts 10,256

	adverts	sales	airplay	attract	var	var	var	var	var	var	var	var	var	var	var	var
1	10,26	330,00	43,00	10,00												
2	985,69	120,00	28,00	7,00												
3	1445,56	360,00	35,00	7,00												
4	1188,19	270,00	33,00	7,00												
5	574,51	220,00	44,00	5,00												
6	568,95	170,00	19,00	5,00												
7	471,81	70,00	20,00	1,00												
8	537,35	210,00	22,00	9,00												
9	514,07	200,00	21,00	7,00												
10	174,09	300,00	40,00	7,00												
11	1720,81	290,00	32,00	7,00												
12	611,48	70,00	20,00	2,00												
13	251,19	150,00	24,00	8,00												
14	97,97	190,00	38,00	6,00												
15	406,81	240,00	24,00	7,00												
16	265,40	100,00	25,00	5,00												
17	1323,29	250,00	35,00	5,00												
18	196,65	210,00	36,00	8,00												
19	1326,60	280,00	27,00	8,00												
20	1380,69	230,00	33,00	8,00												
21	792,35	210,00	33,00	7,00												
22	957,17	230,00	28,00	6,00												
23	1789,66	320,00	30,00	9,00												
24	656,14	210,00	34,00	7,00												
25	613,70	230,00	49,00	7,00												
26	313,36	250,00	40,00	8,00												
27	336,51	60,00	20,00	4,00												
28	1544,90	330,00	42,00	7,00												
29	68,95	150,00	35,00	8,00												
30	785,69	150,00	8,00	6,00												
31	125,63	180,00	49,00	7,00												
32	377,93	80,00	19,00	8,00												

Linear Regression

Dependent:

Linear Regression: Options

Stepping Method Criteria

Use probability of F

Entry: .05 Removal: .10

Use F value

Entry: 3,84 Removal: 2,71

Include constant in equation

Missing Values

Exclude cases listwise

Exclude cases pairwise

Replace with mean

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

OK Paste Reset Cancel Help

Options'a tıklayın.

Tanımlayıcı istatistikler



Descriptive Statistics

	Mean	Std. Deviation	N
Record Sales (thousands)	193,2000	80,69896	200
Advertsing Budget (thousands of pounds)	614,4123	485,65521	200
No. of plays on Radio 1 per week	27,5000	12,26958	200
Attractiveness of Band	6,7700	1,39529	200

Ortalama ve standart sapmalar

Correlations

		Record Sales (thousands)	Advertsing Budget (thousands of pounds)	No. of plays on Radio 1 per week	Attractiveness of Band
Pearson Correlation	Record Sales (thousands)	1,000	,578	,599	,326
	Advertsing Budget (thousands of pounds)	,578	1,000	,102	,081
	No. of plays on Radio 1 per week	,599	,102	1,000	,182
	Attractiveness of Band	,326	,081	,182	1,000
Sig. (1-tailed)	Record Sales (thousands)	.	,000	,000	,000
	Advertsing Budget (thousands of pounds)	,000	.	,076	,128
	No. of plays on Radio 1 per week	,000	,076	.	,005
	Attractiveness of Band	,000	,128	,005	.
N	Record Sales (thousands)	200	200	200	200
	Advertsing Budget (thousands of pounds)	200	200	200	200
	No. of plays on Radio 1 per week	200	200	200	200
	Attractiveness of Band	200	200	200	200

Korelasyon katsayıları

Anlamlılık

Regresyon Modeli Özeti



Model Summary^c

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	,578 ^a	,335	,331	65,99144	,335	99,587	1	198	,000	
2	,815 ^b	,665	,660	47,08734	,330	96,447	2	196	,000	1,950

a. Predictors: (Constant), Advertising Budget (thousands of pounds)

b. Predictors: (Constant), Advertising Budget (thousands of pounds), Attractiveness of Band, No. of plays on Radio 1 per week

c. Dependent Variable: Record Sales (thousands)

Bağımlı değişken

Tahmin değişkenleri

İlk model sadece reklam bütçesiyle albüm satışları arasındaki ilişkiyi, 2. model ise reklam bütçesi, radyoda yayın sayısı ve grubun çekiciliğinin albüm satışları üzerindeki etkisini gösteriyor. İlk model (reklam bütçesi) varyansın %33,5'ini açıklarken, radyoda yayın sayısı ve grubun çekiciliği de eklendiğinde varyansın %66,5', açıklanıyor. Yani sonradan eklenen iki değişken varyansın toplam %30'unu daha açıklıyor.

Düzeltilmiş R^2 modelin genellenebilirliğini gösteriyor. Yani model örneklem yerine evrenden üretilmiş olsaydı toplam varyansın %66'sını açıklıyor olacaktı.



ANOVA^c

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	433687,833	1	433687,833	99,587	,000 ^a
	Residual	862264,167	198	4354,870		
	Total	1295952,000	199			
2	Regression	861377,418	3	287125,806	129,498	,000 ^b
	Residual	434574,582	196	2217,217		
	Total	1295952,000	199			

- a. Predictors: (Constant), Adverting Budget (thousands of pounds)
- b. Predictors: (Constant), Adverting Budget (thousands of pounds), Attractiveness of Band, No. of plays on Radio 1 per week
- c. Dependent Variable: Record Sales (thousands)

ANOVA sonuçlarını nasıl yorumlayacağımızı biliyoruz.
Her iki sonuç için de F istatistiksel açıdan anlamlı.
Yani modeller en iyi tahmin olarak ortalamaları kullanmaktan daha iyi.

Model parametreleri



Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	134,140	7,537		17,799	,000	119,278	149,002					
	Advertsing Budget (thousands of pounds)	,096	,010	,578	9,979	,000	,077	,115	,578	,578	,578	1,000	1,000
2	(Constant)	-26,613	17,350		-1,534	,127	-60,830	7,604					
	Advertsing Budget (thousands of pounds)	,085	,007	,511	12,261	,000	,071	,099	,578	,659	,507	,986	1,015
	No. of plays on Radio 1 per week	3,367	,278	,512	12,123	,000	2,820	3,915	,599	,655	,501	,959	1,043
	Attractiveness of Band	11,086	2,438	,192	4,548	,000	6,279	15,894	,326	,300	,188	,963	1,038

a. Dependent Variable: Record Sales (thousands)

İlk modelde 1000 birimlik bir reklam harcamasıyla fazladan 96 albüm satılacağı tahmin ediliyor. Reklam harcamasının albüm satışlarının sadece üçte birini açıkladığını hatırlayın. Regresyon formülü:

$$\text{Albüm satışları} = 134,14 + (0,09612 \times \text{Reklam bütçesi})$$

2. modelde regresyon formülü:

$$\text{Albüm satışları} = -26,61 + ((0,08 \times \text{Reklam bütçesi}) + (3,37 \times \text{Radyoda çalınma sayısı}) + (11,09 \times \text{Grubun çekiciliği}))$$

Model parametreleri (devamla)



- b değerleri albüm satışlarıyla her tahmin değişkeni arasındaki ilişkiyi gösteriyor (pozitif veya negatif).
- Yani reklam bütçesi arttıkça albüm satışları artıyor; radyoda çalınma sayısı için de aynı şey geçerli; grubun çekiciliği de albüm satışlarına olumlu etki yapıyor.
- Dahası b değerleri bize tüm diğer tahmin değerlerinin etkisi sabit tutulduğunda her tahmin değişkeninin bağımlı değişkeni ne derece etkilediğini gösterir (reklam bütçesi 0,085, radyoda çalınma sayısı 3,367, grubun çekiciliği 11,086)

Model parametreleri (devamla)



- b değerleri önemli, ama standartlaştırılmış b değerlerini yorumlamak daha kolay. Standart b değerleri tahmin değişkenindeki bir standart sapma değişikliğin (tanımlayıcı istatistik tablosuna bakınız) bağımlı değişkende ne kadar değişiklik yaratacağını söyler (örneğin reklam bütçesiyle radyoda çalınma sayısı hemen hemen aynı etkiyi gösteriyor --0,512 ve 0,511--).
- Örneğin, reklam bütçesi 1 SS artarsa (yani 485.655 birim) albüm satışları 0,511 SS artacak (albüm satışlarının SS'si 80.699, bunun 0,511'i 41.240 albüme karşılık geliyor). Bu yorum sadece radyoda çalınma sayısı ve grubun çekiciliği sabit tutulduğunda geçerli.
- İyi bir modelin %95 güven aralıkları küçüktür (reklam bütçesi ve radyoda çalınma sayısınıninkilerle grubun çekiciliğini karşılaştırın).

Descriptive Statistics

	Mean	Std. Deviation	N
Record Sales (thousands)	193,2000	80,69896	200
Advertsing Budget (thousands of pounds)	614,4123	485,65521	200
No. of plays on Radio 1 per week	27,5000	12,26958	200
Attractiveness of Band	6,7700	1,39529	200

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	134,140	7,537		17,799	,000	119,278	149,002					
	Advertsing Budget (thousands of pounds)	,096	,010	,578	9,979	,000	,077	,115	,578	,578	,578	1,000	1,000
2	(Constant)	-26,613	17,350		-1,534	,127	-60,830	7,604					
	Advertsing Budget (thousands of pounds)	,085	,009	,511	9,221	,000	,071	,099	,578	,659	,507	,986	1,015
	No. of plays on Radio 1 per week	3,367	,278	,512	12,113	,000	2,820	3,915	,599	,655	,501	,959	1,043
	Attractiveness of Band	11,086	2,438	,192	4,548	,000	6,279	15,894	,326	,309	,188	,963	1,038

a. Dependent Variable: Record Sales (thousands)

Korelasyonlar



- Zero-order korelasyonlar basit Pearson korelasyon katsayılarıdır.
- Kısmi korelasyonlar diğer iki değişkenin etkilerini kontrol ederek her tahmin değişkeniyle bağımlı değişken arasındaki ilişkiyi temsil eder.

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		Correlations			Collinearity Statistics		
	B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF	
1	(Constant)	134,140	7,537		17,799	,000	119,278	149,002					
	Advertsing Budget (thousands of pounds)	,096	,010	,578	9,979	,000	,077	,115	,578	,578	,578	1,000	1,000
2	(Constant)	-26,613	17,350		-1,534	,127	-60,830	7,604					
	Advertsing Budget (thousands of pounds)	,085	,007	,511	12,261	,000	,071	,099	,578	,659	,507	,986	1,015
	No. of plays on Radio 1 per week	3,367	,278	,512	12,123	,000	2,820	3,915	,599	,655	,501	,959	1,043
	Attractiveness of Band	11,086	2,438	,192	4,548	,000	6,279	15,894	,326	,309	,188	,963	1,038

a. Dependent Variable: Record Sales (thousands)

Tahmin deęişkenleri arasındaki korelasyon (collinearity)



- Varyans Enflasyon Faktörü (VIF) 10'un altında, tolerans istatistikleri de 0,2'nin üstünde olduğundan sorun yok. Yani tahmin deęişkenleri arasında mükemmel doğrusal ilişkiden söz edilemez.
- Nitekim bir sonraki slaytta her tahmin deęişkeninin varyansın farklı boyutunu açıkladığını görüyoruz.

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	134,140	7,537		17,799	,000	119,278	149,002					
	Advertsing Budget (thousands of pounds)	,096	,010	,578	9,979	,000	,077	,115	,578	,578	,578	1,000	1,000
2	(Constant)	-26,613	17,350		-1,534	,127	-60,830	7,604					
	Advertsing Budget (thousands of pounds)	,085	,007	,511	12,261	,000	,071	,099	,578	,659	,507	,986	1,015
	No. of plays on Radio 1 per week	3,367	,278	,512	12,123	,000	2,820	3,915	,599	,655	,50	,959	1,043
	Attractiveness of Band	11,086	2,438	,192	4,548	,000	6,279	15,894	,326	,309	,188	,963	1,038

a. Dependent Variable: Record Sales (thousands)

Collinearity



Collinearity Diagnostics³

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions			
				(Constant)	Advertising Budget (thousands of pounds)	No. of plays on Radio 1 per week	Attractiveness of Band
1	1	1,785	1,000	,11	,11		
	2	,215	2,883	,89	,89		
2	1	3,562	1,000	,00	,02	,01	,00
	2	,308	3,401	,01	,96	,05	,01
	3	,109	5,704	,05	,02	,93	,07
	4	,020	13,219	,94	,00	,00	,92

a. Dependent Variable: Record Sales (thousands)

Ortalamadan ± 2 SS farklı olan değerler



Casewise Diagnostics^a

Case Number	Std. Residual	Record Sales (thousands)	Predicted Value	Residual
1	2,125	330,00	229,9203	100,0797
2	-2,314	120,00	228,9490	-108,9490
10	2,114	300,00	200,4662	99,5338
47	-2,442	40,00	154,9698	-114,9698
52	2,069	190,00	92,5973	97,4027
55	-2,424	190,00	304,1231	-114,1231
61	2,098	300,00	201,1897	98,8103
68	-2,345	70,00	180,4156	-110,4156
100	2,066	250,00	152,7133	97,2867
164	-2,577	120,00	241,3240	-121,3240
169	3,061	360,00	215,8675	144,1325
200	-2,064	110,00	207,2061	-97,2061

Ortalamadan
2 SS uzakta
olan vakalar

a. Dependent Variable: Record Sales (thousands)

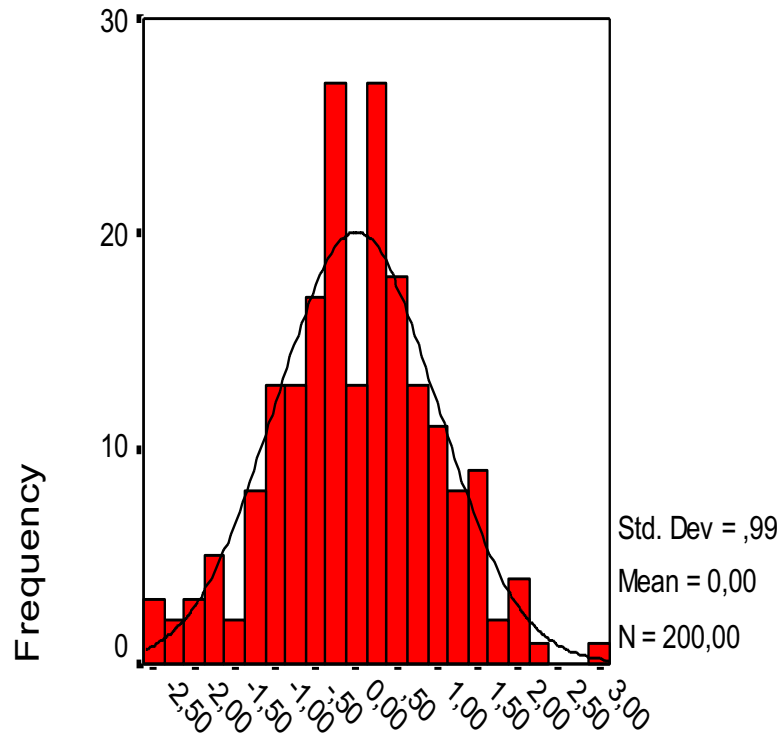
Alameda bir örnekleme ölçümlerin %95'inin modelin tahmin ettiği değerle gözlenen değer arasındaki farkı (residual) ortalamadan en çok 2 SS uzaktır. Bu örnekleme 200 ölçümün 12'si ortalamadan ± 2 SS'den daha büyüktür (yani %6, ki normal).

Residuals normal dağılmış



Histogram

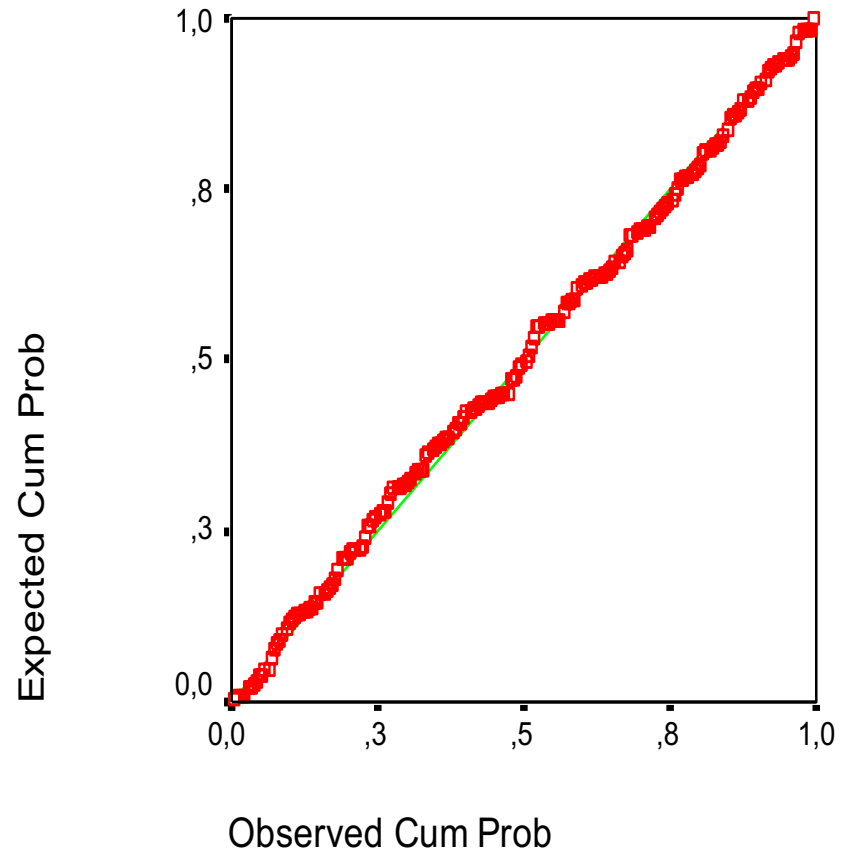
Dependent Variable: Record Sales



Regression Standardized Residual

Normal P-P Plot of Regression ϵ

Dependent Variable: Record Sal

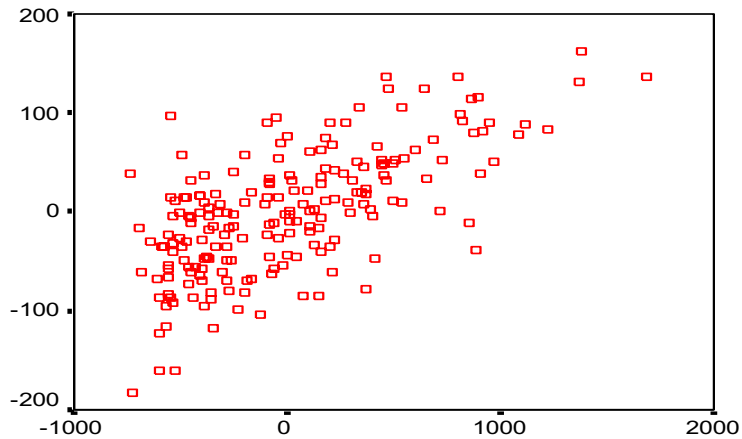


Kısmi regresyon grafikleri



Partial Regression Plot

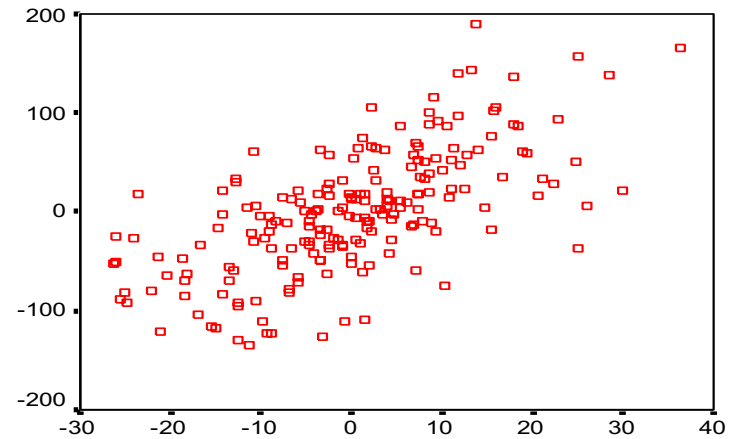
Dependent Variable: Record Sales



Advertising Budget (thousands of pounds)

Partial Regression Plot

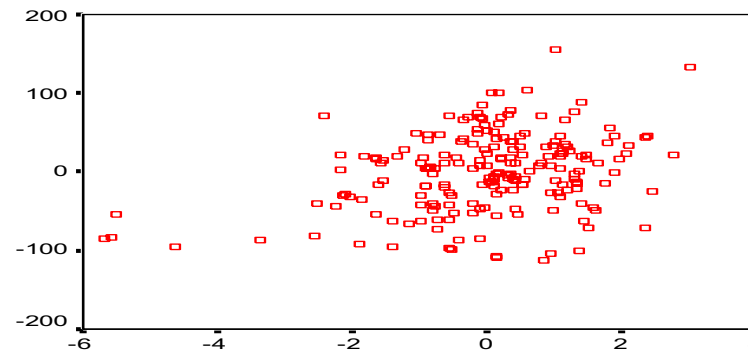
Dependent Variable: Record Sales



No. of plays on Radio 1 per week

Partial Regression Plot

Dependent Variable: Record Sales



Attractiveness of Band



- Önce değişkenlerle ilgili tanımlayıcı istatistikler verilir ve yorumlanır (önceki slaytlardan yararlanılabilir).

Çoklu Regresyon Sonuçlarını Rapor Etme



APA stiline göre:

	B	SH	β
Model 1			
Sabit	134,14	7,54	
Reklam bütçesi	0,1	0,01	0,58*
Model 2			
Sabit	-26,61	17,35	
Reklam bütçesi	0,09	0,01	0,51*
Radyoda çalınma sayısı	3,37	0,28	0,51*
Grubun cazibesi	11,09	2,44	0,19*

Model		Unstandardized Coefficients		Standardized Coefficients
		B	Std. Error	Beta
1	(Constant)	134,140	7,537	
	Advertsing Budget (thousands of pounds)	,096	,010	,578
2	(Constant)	-26,613	17,350	
	Advertsing Budget (thousands of pounds)	,085	,007	,511
	No. of plays on Radio 1 per week	3,367	,278	,512
	Attractiveness of Band	11,086	2,438	,192

a. Dependent Variable: Record Sales (thousands)

Not: Model 1 için $R^2 = 0,34$ ($p < 0,001$); Model 2 için $\Delta R^2 = 0,33$ ($p < 0,001$). * $p < 0,001$

Tablo daha önceki slaytlarda yer alan sözlü yorumlarla zenginleştirilebilir.



- Çoklu regresyon analizinde bağımlı değişken aralıklı/oranlı, sürekli ve sınırsız
- Logistik regresyonda bağımlı değişken kategorik
- Çoklu regresyonda tahmin değişkenleri regresyon katsayılarıyla çarpılarak bağımlı değişken tahmin edilir
- Logistik regresyonda tahmin değişkenlerinin değerlerini bildiğimizde bağımlı değişkenin meydana gelme olasılığı tahmin edilmeye çalışılır
- Logistik regresyon formülü:

$$P(Y) = 1 / 1 + e^{-(b_0 + b_1 X_1 + b_2 X_2 + \dots + b_n X_n) + e_i}$$

Logistik Regresyon Analizi (devamla)



- Doğrusal regresyonun sayıltılarından biri ilişkinin doğrusal olması
- Ama bağımlı değişkenin kategorik olduğu durumlarda bu sayıltı ihlal edilir
- Bunu çözenin bir yolu verileri logaritmik dönüştürüme tabi tutmaktır
- Bu işlem doğrusal olmayan bir ilişkiyi koruyarak ilişkinin formunu doğrusal hale getirir
- Yani doğrusallık sayıltısı ihlal edilmeden çoklu doğrusal regresyon logaritmik terimlerle açıklanmış olur
- Sonuçta bağımlı değişkenin değeri 0 ile 1 arasında çıkar. Değer sıfıra yakınsa bağımlı değişkenin olma olasılığı düşük, 1'e yakınsa yüksektir.



- Gelişme psikolojisinden bir test örneği (geçti-kaldı). Örnek, çocukların kendilerine verilen bir armağanı beğenmediklerinde armağanı veren kişinin duygularını incitip (ağlama) incitmemeleriyle (nazikçe gülümseyerek teşekkür etme) ilgili. Bu test çocukların karşıdaki kişinin ne düşündüğünü anlayacak bilişsel yeteneğe (akıl kuramına) sahip olup olmadıklarını ölçer.



- Bağımlı değişken: Çocuğun testi geçip geçmediği (Evet / Hayır)
- Tahmin değişkeni: akıl kuramına sahip olup olmama (Evet / Hayır)
- Tahmin değişkeni: Çocuğun kaç aylık olduğu

SPSS'te logistik regresyon



Logistic Regression

Dependent: Display Rule understar

Block 1 of 1

Covariates: age, fb, age*fb

Method: Forward:LR

Method options: Enter, Forward:Conditional, Forward:LR, Forward:Wald, Backward:Conditional, Backward:LR, Backward:Wald

	age	fb	display	pre_1	pgr_1	coo_1	lev_1	zre_1	dfb0_1	dfb1_1	pre_2	pgr_2	coo_2	lev_2	zre_2	dfb0_2	dfb1_2
1	24,00	,00	,00														
2	26,00	,00	,00														
3	30,00	,00	,00														
4	31,00	,00	,00														
5	36,00	,00	,00														
6	29,00	,00	1,00														
7	28,00	,00	,00														
8	36,00	,00	,00														
9	34,00	,00	1,00														
10	31,00	,00	,00														
11	32,00	,00	,00														
12	30,00	1,00	1,00														
13	26,00	,00	,00														
14	35,00	1,00	,00														
15	36,00	,00	,00														
16	36,00	,00	,00														
17	25,00	,00	1,00														
18	24,00	,00	,00														
19	24,00	1,00	1,00	,80487	1,00	,00606	,02439										
20	26,00	,00	,00	,20690	,00	,00932	,03448										
21	29,00	,00	,00	,20690	,00	,00932	,03448										
22	34,00	1,00	1,00	,80487	1,00	,00606	,02439	,49237	,00000	,03106	,80487	1,00	,00606	,02439	,49237	,00000	,031
23	31,00	1,00	1,00	,80487	1,00	,00606	,02439	,49237	,00000	,03106	,80487	1,00	,00606	,02439	,49237	,00000	,031
24	40,00	1,00	,00	,80487	1,00	,10312	,02439	-2,03097	,00000	-,12812	,80487	1,00	,10312	,02439	-2,03097	,00000	-,12E
25	45,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E

Analyze->Regression->Binary logistics'i seçin; bağımlı değişkene Display Rule (davranış testi), kovaryansa yaş (age) ve yanlış inanç görevini (fb) geçip geçmediği ve aralarındaki etkileşim (age*fb) girin. Yöntem olarak Forward:LR'yi girin.

Kategorik tahmin deęişkenleri



Logistik regresyonda hangi tahmin deęişken(ler)inin kategorik olduęu belirtilir

The screenshot displays the SPSS interface with a data view containing 32 rows and 18 columns. The columns are: age, fb, display, pre_1, pgr_1, coo_1, lev_1, zre_1, dfb0_1, dfb1_1, pre_2, pgr_2, coo_2, lev_2, zre_2, dfb0_2, and dfb1_2. The data values range from 0.00 to 1.00 for categorical variables and various numerical values for continuous variables.

Overlaid on the data view are two dialog boxes:

- Logistic Regression**: The 'Dependent' variable is set to 'Age in years [age]'. The 'Covariates' list is empty.
- Logistic Regression: Define Categorical Variables**: This sub-dialog is open for the 'Age in years [age]' variable. The 'Categorical Covariates' list contains 'fb(Indicator)'. The 'Change Contrast' section shows 'Contrast: Indicator' and 'Reference Category: First'.

The Windows taskbar at the bottom shows the start button, several open applications (Adobe, Chapt..., 17 I..., 2 Mi..., 3 Mi..., Adobe..., Windo..., 4 Vol..., 4 Mi..., Micros..., Displa...), and the system clock showing 00:21.

Residuals



Logistic Regression

Dependent: Age in years [age]

Logistic Regression: Save New Variables

Probabilities

Group membership

Cook's

Leverage values

DiBeta(s)

Unstandardized

Logit

Studentized

Standardized

Deviance

OK

Paste

Reset

Cancel

Help

Continue

Cancel

Help

Select >>

Categorical...

Save...

Options...

	age	fb	display	pre_1	pgr_1	coo_1	lev_1	zre_1	dfb0_1	dfb1_1	pre_2	pgr_2	coo_2	lev_2	zre_2	dfb0_2	dfb1_2
1	24,00	,00	,00														
2	26,00	,00	,00														
3	30,00	,00	,00														
4	31,00	,00	,00														
5	36,00	,00	,00														
6	29,00	,00	1,00														
7	28,00	,00	,00														
8	36,00	,00	,00														
9	34,00	,00	1,00														
10	31,00	,00	,00														
11	32,00	,00	,00														
12	30,00	1,00	1,00														
13	26,00	,00	,00														
14	35,00	1,00	,00														
15	36,00	,00	,00														
16	36,00	,00	,00														
17	25,00	,00	1,00														
18	24,00	,00	,00														
19	24,00	1,00	1,00	,80487	1,00	,00606	,02439	,49237	,00000	,03106	,80487	1,00	,00606	,02439	,49237	,00000	,031
20	26,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E
21	29,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E
22	34,00	1,00	1,00	,80487	1,00	,00606	,02439	,49237	,00000	,03106	,80487	1,00	,00606	,02439	,49237	,00000	,031
23	31,00	1,00	1,00	,80487	1,00	,00606	,02439	,49237	,00000	,03106	,80487	1,00	,00606	,02439	,49237	,00000	,031
24	40,00	1,00	,00	,80487	1,00	,10312	,02439	-,203097	,00000	-,12812	,80487	1,00	,10312	,02439	-,203097	,00000	-,12E
25	45,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E
26	30,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E
27	30,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E
28	36,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E
29	45,00	1,00	1,00	,80487	1,00	,00606	,02439	,49237	,00000	,03106	,80487	1,00	,00606	,02439	,49237	,00000	,031
30	40,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E
31	41,00	,00	1,00	,20690	,00	,13690	,03448	1,95787	,17262	,17262	,20690	,00	,13690	,03448	1,95787	,17262	-,172
32	32,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E

Logistik regresyona özgü tekil özellikler tahmin edilen olasılıklarla tahmin edilen grup üyelikleridir.

Options



Display - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

20 : dfb0_1 -0,0450312206556068

	age	fb	display	pre_1	pgr_1	coo_1	lev_1	zre_1	dfb0_1	dfb1_1	pre_2	pgr_2	coo_2	lev_2	zre_2	dfb0_2	dfb1_2
1	24,00	,00	,00														
2	26,00	,00	,00														
3	30,00	,00	,00														
4	31,00	,00	,00														
5	36,00	,00	,00														
6	29,00	,00	1,00														
7	28,00	,00	,00														
8	36,00	,00	,00														
9	34,00	,00	1,00														
10	31,00	,00	,00														
11	32,00	,00	,00														
12	30,00	1,00	1,00														
13	26,00	,00	,00														
14	35,00	1,00	,00														
15	36,00	,00	,00														
16	36,00	,00	,00														
17	25,00	,00	1,00														
18	24,00	,00	,00														
19	24,00	1,00	1,00														
20	26,00	,00	,00														
21	29,00	,00	,00														
22	34,00	1,00	1,00														
23	31,00	1,00	1,00	,80487	1,00	,00606	,02439	,49237	,00000	,03106	,80487	1,00	,00606	,02439	,49237	,00000	,031
24	40,00	1,00	,00	,80487	1,00	,10312	,02439	-2,03097	,00000	-,12812	,80487	1,00	,10312	,02439	-2,03097	,00000	-,12E
25	45,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E
26	30,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E
27	30,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E
28	36,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E
29	45,00	1,00	1,00	,80487	1,00	,00606	,02439	,49237	,00000	,03106	,80487	1,00	,00606	,02439	,49237	,00000	,031
30	40,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E
31	41,00	,00	1,00	,20690	,00	,13690	,03448	1,95787	,17262	-,17262	,20690	,00	,13690	,03448	1,95787	,17262	-,172
32	32,00	,00	,00	,20690	,00	,00932	,03448	-,51076	-,04503	,04503	,20690	,00	,00932	,03448	-,51076	-,04503	,04E

Logistic Regression

Age in years [age] Dependent:

Logistic Regression: Options

Statistics and Plots

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

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

Data View Variable View

SPSS Processor is ready

start Adobe... Chapt... 17 I... 2 Mi... 3 Mi... Adobe... Windo... 4 Vol... 3 Mi... Micros... Displa... 00:30

Logistik regresyonu yorumlama



Dependent Variable Encoding

Original Value	Internal Value
No	0
Yes	1

Categorical Variables Codings

		Frequency	Parameter (1)
False Belief understanding	No	29	1,000
	Yes	41	,000

Kodlama bilgileri



Classification Table^{a,b}

Observed			Predicted		Percentage Correct
			Display Rule understanding		
			No	Yes	
Step 0	Display Rule understanding	No	0	31	,0
		Yes	0	39	100,0
		Overall Percentage			55,7

a. Constant is included in the model.

b. The cut value is ,500

SPSS tüm çocukların davranış kuralını anladıklarını varsayarak vakaların %56'sında doğru tahmin yapıyor

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	,230	,241	,910	1	,340	1,258

B sabiti 0,23.

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	AGE	15,956	1	,000
		FB(1)	24,617	1	,000
		AGE by FB(1)	18,556	1	,000
	Overall Statistics		26,257	3	,000

Modelde olmayan tahmin değişkenleri modelin tahmin gücünü istatistiksel açıdan anlamlı derecede artıracak. Overall Statistics olarak verilen Ki kare değeri ve anlamlı.



Omnibus Tests of Model Coefficients

	Chi-square	df	Sig.
Step 1 Step	26,083	1	,000
Block	26,083	1	,000
Model	26,083	1	,000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	70,042	,311	,417

Classification Table^a

Observed			Predicted		
			Display Rule understanding		Percentage Correct
			No	Yes	
Step 1	Display Rule understanding	No	23	8	74,2
		Yes	6	33	84,6
		Overall Percentage			80,0

a. The cut value is ,500

İlk adımda modele yanlış inanç görevi tahmin değişkeni ekleniyor. Yanlış inanç anlama sahibi 39 çocuk var. Model bu çocukların 33'ünün testi geçip geçmediklerini doğru tahmin ediyor. Model yanlış inanç anlama sahibi olmayan 31 çocuktan 23'ünü de doğru tahmin ediyor. Genel sınıflamada doğruluk %80'e çıktı. (Sadece sabit alındığında bu oran %56 idi.)

Denklemdaki değişkenler



Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95,0% C.I. for EXP(B)	
								Lower	Upper
Step	FB(1)	-2,761	,605	20,857		,000	,063	,019	,207
1	Constant	1,417	,394	12,930	1	,000	4,125		

a. Variable(s) entered on step 1: FB.

Bu tabloda modelde yer alan değişkenlerin katsayı tahminleri var. Wald istatistiği tahmin değişkeninin b katsayısının sıfırdan anlamlı derecede farklı olup olmadığını gösteriyor (hem sabit hem de fb için anlamlı).



Model if Term Removed

Variable	Model Log Likelihood	Change in -2 Log Likelihood	df	Sig. of the Change
Step 1 FB	-48,062	26,083	1	,000

Denklemdaki fb deęişkenini kaldırmak iyi bir fikir deęil, çünkü modelin tahmin gücünü istatistiksel açıdan anlamlı düzeyde azaltıyor.



Variables not in the Equation

			Score	df	Sig.
Step	Variables	AGE	2,313	1	,128
1		AGE by FB(1)	1,261	1	,262
	Overall Statistics		2,521	2	,284

Denkleimde olmayan diğer değişkenler (AGE ve AGE*FB) modelin tahmin gücünü istatistiksel açıdan anlamlı düzeyde artırmıyor. Bu nedenle denkleme başka değişken eklemek gereksiz.



- Seyrek kullanıldığı için nasıl rapor edileceği konusunda somut talimatlar yok.
- APA stiline göre beta değerleri ve standart hataları ile anlamlılık düzeyleri belirtilmeli.
- Modelin R^2 değeri ya da uyum iyiliği testi
- Anlamli olmayan tahmin deęişkenleri de belirtilebilir.



Regresyon Analizi

Yaşar Tonta
H.Ü. BBY

tonta@hacettepe.edu.tr

yunus.hacettepe.edu.tr/~tonta/courses/fall2008/sb5002/