

## Regresyon Analizi

Regresyon analizi bir ya da bir kaç değişkenin değerlerinden yararlanarak başka bir değişkenin değerlerini tahmin etmek amacıyla kullanılır. Örneğin bulutların yoğunluğu, havadaki nem miktarı gibi değişkenleri kullanarak ertesi gün yağmurun ne miktarda yapabileceğini tahmin etmek böyle bir analizle mümkündür. Kendi kendine frene basan bir araba yapmak istiyorsanız öndeki arabaya yaklaşma hızı bilgisini kullanarak frene ne kuvvetle basıldığında ne olduğunu (arabanın kaç metre önce durduğu veya duramayıp diğer arabaya çarptığı) tahmin edip arabanın hangi durumda ne kadar frene basması gerektiği arabanın bilgisayarına öğretilir. Öğrencilerin ders çalışmaya ayırdıkları süreyi kullanarak sınavdan alacakları puanlar tahmin edilebilir. Bir tedavi yönteminin kanda oluşturduğu değerler kullanılarak hastanın kaç gün içinde iyileşeceği tahmin edilebilir. Ya da örgüt kültürü boyutlarının mesleki güçlendirme üzerindeki varyansın ne kadarını açıkladığı tahmin edilebilir. Doğrusal, lojistik, polinom, adimsal, ridge, kement ve elastik ağ regresyon analizi gibi bir çok regresyon analizi türü ve bu türlerin de kendi içinde alt türleri bulunmakta olup bunlar arasında en sık kullanılanı doğrusal (basit veya çoklu) analizi modelidir. Regresyon analizi için öncelikle tahmin edilmek istenen (bağımlı) değişken ve tahmin edici (bağımsız) değişkenin ölçülmüş olması gerekir. Ardından elde edilen veriler kullanılarak regresyon denklemi kurulur.

$$R: Y=a+(b_1*x_1)+(b_2*x_2)+(b_3*x_3)+(b_4*x_4)$$

Bu denklemde R değeri işlem sonucu elde edilecek regresyon katsayısını yani tahmin edilen değeri vermektedir. Bu değer de Y değerine eşittir. Y değeri yordanan (tahmin edilen/bağımlı) değişkeni temsil etmektedir. a değeri ise sabit (Constant/ Yordayan değişken 0 olduğunda yordanan değişkenin alacağı değerdir) katsayıdır.  $x_1$ ,  $x_2$ ,  $x_3$  ve  $x_4$  değerleri bağımsız (tahmin edici/yordayan) değişkenin faktörlerini göstermektedir.  $b_1$ ,  $b_2$ ,  $b_3$  ve  $b_4$  değerleri ise bağımsız (tahmin edici/yordayan) değişkenlerin bağımlı (tahmin edilen/yordanan) değişkeni açıklama miktarını gösteren beta katsayılarıdır. Başka bir ifade ile “b” değerleri yordayan değişken 1 birim arttığında yordanan değişkendeki değişim miktarını göstermektedir. Özetlemek gerekirse bu denklem ile 4 faktörlü bir yapının tahmin edilmek istenen değişkenin % kaçını açıkladığını veya bağımlı değişkendeki her 1 birimlik değişimin % kaçının bağımsız değişkenden kaynaklandığını hesaplamaya çalışırız. Regresyon analizi ile her bağımsız değişkenin her bağımlı değişkeni tahmin edebileceğini düşünmek büyük bir yanılgıdır. Çoğu durumda bağımsız değişken ile bağımlı değişken arasında regresyonel (tahmin edici) bir ilişki ortaya çıkmayabilir. Bu nedenle regresyon analizi öncelikle bir değişkenden (bağımsız) başka bir değişkenin (bağımlı) tahmin edilip edilemeyeceğinin belirlendiği, sonrasında ise denklemin üretildiği bir analizdir. Eğer ilk aşamada regresyon yapılamayacağı anlaşılırsa ikinci aşama olan denklem üretmeye geçilmez. Bu sebeple regresyon analizine geçilmeden önce veri setinin normallik, sıfır ortalama, sabit varyans, otokorelasyon olmaması, bağımsız değişkenin tesadüfi değişken olmaması, çoklu doğrusal bağlılık olmaması (çoklu doğrusal regresyonda),  $n > k$  oranının (n gözlem sayısı, k parametre sayısı) (çoklu doğrusal regresyonda) olması gibi bazı temel regresyon varsayımlarına sahip olup olmadığı kontrol edilir.

Aşağıda ana hatları ile açıklanan regresyon analizine yönelik bir uygulama gerçekleştirmek için SPSS de çoklu doğrusal analizi örneklendirilmiştir.

Correlations						
		GÜÇLENDİRME	GÖREV	DESTEK	BÜROKRATİK	BAŞARI
Pearson Correlation	GÜÇLENDİRME	1.000	.321	.215	-.021	.235
	GÖREV	.321	1.000	.286	.141	.368
	DESTEK	.215	.286	1.000	-.031	.686
	BÜROKRATİK	-.021	.141	-.031	1.000	-.019
	BAŞARI	.235	.368	.686	-.019	1.000
Sig. (1-tailed)	GÜÇLENDİRME	.	.000	.000	.357	.000
	GÖREV	.000	.	.000	.006	.000
	DESTEK	.000	.000	.	.293	.000
	BÜROKRATİK	.357	.006	.293	.	.369
	BAŞARI	.000	.000	.000	.369	.
N	GÜÇLENDİRME	313	313	313	313	313
	GÖREV	313	313	313	313	313
	DESTEK	313	313	313	313	313
	BÜROKRATİK	313	313	313	313	313
	BAŞARI	313	313	313	313	313

Otokorelasyon (çoklu doğrusallık) olup olmadığı burada kontrol edilir. Yukarıdaki tabloya bakıldığında bağımlı değişkinin faktörleri arasında otokorelasyon olmadığı görülecektir. Yordayan değişkenler arasında .70 veya .90 üzerinde bir ilişkinin olmaması gerekir.

Model Summary <sup>b</sup>										
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	.355a	.126	.114	.39446	.126	11.079	4	308	.000	1.734

a Predictors: (Constant), BAŞARI, BÜROKRATİK, GÖREV, DESTEK  
b Dependent Variable: GÜÇLENDİRME

Model Summary; görev, destek, bürokratik ve başarı kültürü değişkenlerinin bir arada güçlendirme değişkeninin ne kadarını açıkladığını veya güçlendirme değişkenini ne kadar tahmin ettiğini görmek için kullanılır. Bunun için tablodaki “R Square” değerine bakılır. Elimizdeki örneğe göre görev, destek, bürokratik ve başarı kültürü değişkenleri bir arada güçlendirme değişkeninin %13’ünü açıkladığı söylenebilir. Bu tablodaki “Durbin-Watson” değeri de regresyon sayılıtlarından biri olan otokorelasyon olup olmadığını kontrol için kullanılır. “Durbin-Watson” değerinin 1.5 ile 2.5 arasında olması otokorelasyonun olmadığına yönelik işaretlerden biridir.

ANOVA <sup>a</sup>						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	6.896	4	1.724	11.079	.000b
	Residual	47.924	308	.156		
	Total	54.820	312			

a Dependent Variable: GÜÇLENDİRME  
b Predictors: (Constant), BAŞARI, BÜROKRATİK, GÖREV, DESTEK

ANOVA tablosu ise yukarıda kurduğumuz ve  $R=a+(b*x)+(c*y)+(d*z)+(e*t)$  formülü ile ifade edilen regresyon modelin anlamlılığını test etmek için kullandığımız analiz sonucuna yönelik tablodur. Bu tabloda “Sig.” Değeri .005’ten küçük olduğu için kurulan regresyon modelinin anlamlı olduğu söylenebilir.

Coefficientsa											
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	2.983	.250		11.914	.000					
	GÖREV	.239	.050	.278	4.781	.000	.321	.263	.255	.840	1.190
	DESTEK	.066	.058	.082	1.124	.262	.215	.064	.060	.527	1.898
	BÜROKRATİK	-.044	.042	-.056	-1.037	.300	-.021	-.059	-.055	.973	1.027
	BAŞARI	.060	.061	.075	.991	.322	.235	.056	.053	.496	2.016

Coefficients tablosundaki Tolerance ve VIF değerleri ile Collinearity Diagnostics'teki Condition Index değerleri bağımsız değişkenler arasındaki çoklu bağlantılılığı (otokorelasyon) incelemek için kullanılan bir diğer araçtır. Tolerance<0.2 , VIF>10 , CI>30 ise, değişkenler arasında çoklu bağlantılılık olduğu görülür. Yukarıdaki tablo incelendiğinde bir kez daha otokorelasyon olmadığı sonucuna varılabilir. Bu tablonun asıl önemli olan kısmı regresyon modelindeki a, b1, b2, b3 ve b4 değerini vermesidir. Burada modele standardize edilmemiş değerler yazılır. Buna göre a=2.983, b1=.239, b2=.066, b3=-.044 ve b4=.060 olarak bulunmuştur denilebilir.

Collinearity Diagnosticsa									
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions					
				(Constant)	GÖREV	DESTEK	BÜROKRATİK	BAŞARI	
1	1	4.946	1.000	.00	.00	.00	.00	.00	
	2	.031	12.609	.00	.00	.05	.65	.04	
	3	.012	20.634	.04	.59	.18	.21	.05	
	4	.006	28.151	.84	.25	.02	.13	.23	
	5	.005	30.188	.13	.16	.75	.01	.68	

Collinearity diagnostics tablosu .005 anlamlılık düzeyi ve -2,+2 standart sapmanın dışında kalan değerli gösterir. Bu tabloda beş değer olması .005 in altında olduğunu gösterir. Bu da modeli bozacak kadar uç değer olmadığını bir göstergesi olarak sunulabilir.

Sonuç olarak yukarıdaki örneğin regresyon modeli formülü şudur:

$$R: \text{Güçlendirme} = 2.983 + (.239 * \text{Görev}) + (.066 * \text{Destek}) + (-.044 * \text{Bürokratik}) + (.060 * \text{Başarı})$$