

آموزش

رگرسیون خطی ساده

با نرم افزار SPSS

تدوین: رامین کریمی

www.kharazmi-Statistics.ir

مقدمه

تحلیل رگرسیون مرحله ای است بعد از همبستگی؛ تحلیل رگرسیون هنگامی استفاده می شود که بخواهیم مقادیر یک متغیر را از روی مقادیر متغیر دیگر پیش بینی کنیم. در این مورد، متغیری که ما از بهره می گیریم تا مقدار متغیر دیگر را پیش بینی کنیم، متغیر مستقل (یا پیش بین) نام دارد. متغیری را هم که می خواهیم پیش بینی کنیم متغیر وابسته (یا ملاک) نام دارد.

مثال

یک فروشنده ی یک برند معروف ماشین، علاقه مند شده است تا بررسی کند که بین درآمد شخصی افراد و هزینه ای که آنان برای یک ماشین می پردازند رابطه ای وجود دارد یا خیر. آن ها می خواهند از این نتایج برای پیشنهاد ماشین در مناطق جدیدی که میانگین درآمد آنان مشخص است استفاده کنند.

متغیر مستقل: درآمد افراد (Income)

متغیر وابسته: قیمت ماشین (Price)

مفروضات

جهت استفاده از این آزمون پیش رض های زیر می بایست محقق باشد:

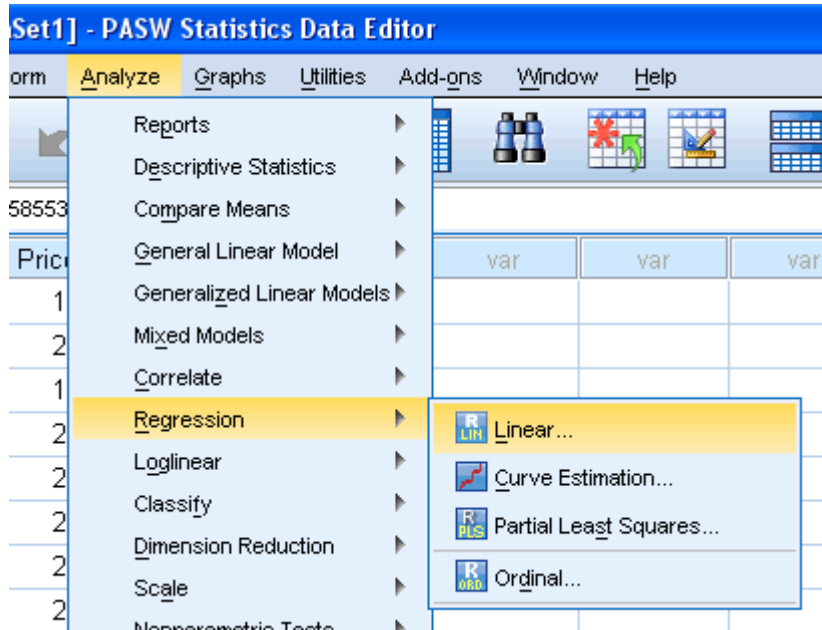
- متغیرها در سطح سنجش فاصله ای یا نسبی باشند.
- توزیع متغیرها نرمال یا نزدیک به نرمال باشد.
- بین دو متغیر رابطه خطی وجود داشته باشد.

جهت سنجش نرمال بودن توزیع متغیرها از آزمون Shapiro-Wilk و یا Kolmogorov-Smirnov استفاده می کنیم.

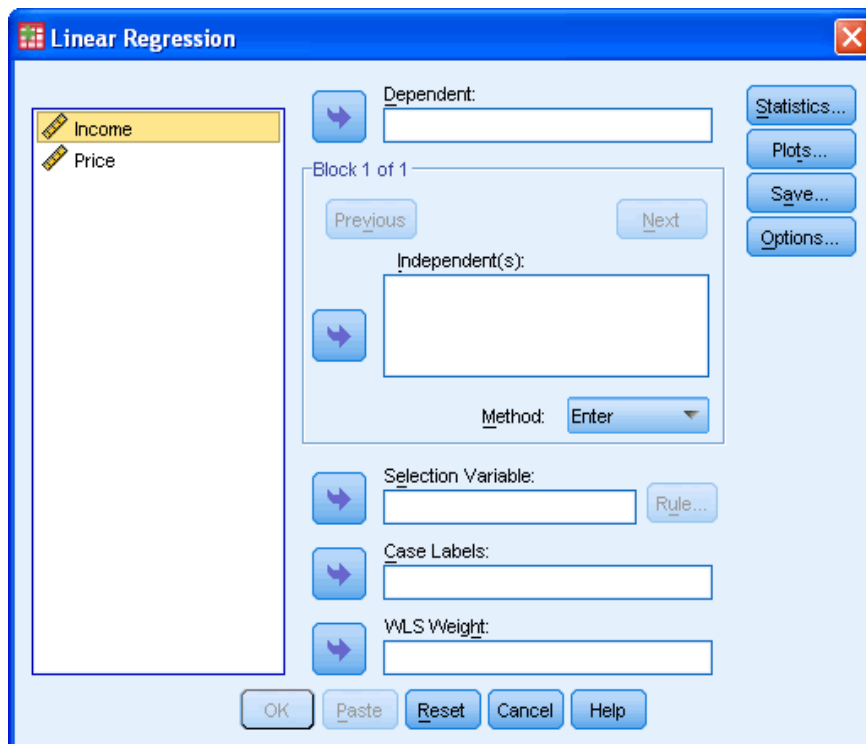
جهت سنجش وجود رابطه خطی بین دو متغیر نیز از آزمون همبستگی پیرسون استفاده می کنیم.


اجرا

۱. مسیر Analyze > Regression > Linear... را دنبال می کنیم.

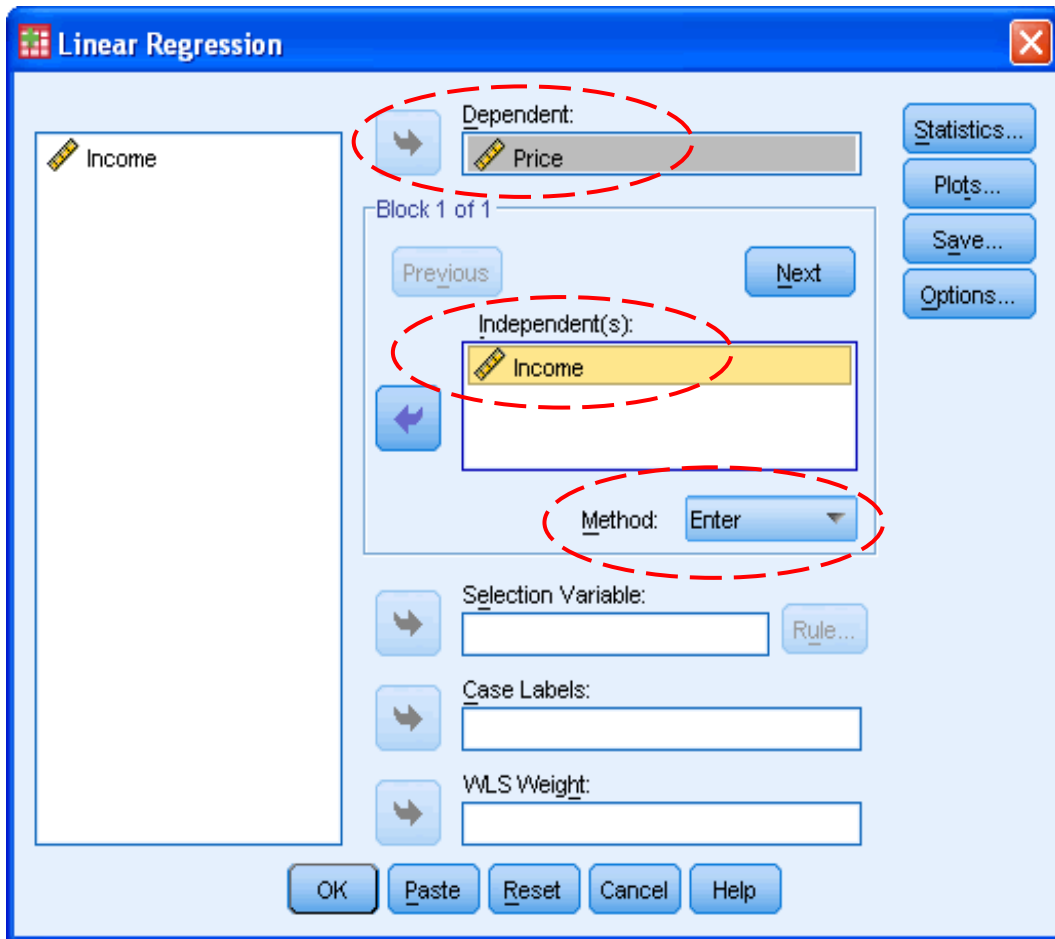


حالا شما کادر زیر را مشاهده می کنید:



۲. متغیر مستقل (پیش بین) یا همان درآمد را به داخل کادر Independent(s) انتقال می دهیم. متغیر وابسته یا همان قیمت خودرو را وارد کادر Dependent می کنیم. این کار را با انتخاب متغیر و کلیک بر روی آیکون  انجام می دهیم.

روش (Method) اجرای رگرسیون را نیز مطابق پیش فرض رگرسیون، یعنی همزمان (Enter) انتخاب می کنیم.



3. بر روی  کلیک می کنیم

تفسیر نتایج

نرم افزار SPSS چند جدول را جهت تحلیل ارائه می دهد. در این مبحث ما به مهم ترین این جداول می پردازیم. اولین جدول Model Summary به معنای خلاصه مدل است. این جدول مقادیر R و R^2 را نشان می دهد. مقدار R برابر است با 0.873، اشاره دارد به همبستگی ساده بین دو متغیر و به عبارتی شدت همبستگی بین دو متغیر را نشان می دهد. همان طور که از مقدار R (همبستگی پیرسون بین دو متغیر) نمایان است، بین دو متغیر درآمد و قیمت ماشین همبستگی در حد خیلی قوی وجود دارد.

مقدار R^2 نشان می دهد که چه مقدار از متغیر وابسته یعنی قیمت ماشین، می تواند توسط متغیر مستقل یعنی درآمد، تبیین شود. در این مثال متغیر درآمد می تواند 76.2 درصد از تغییرات متغیر قیمت ماشین را تبیین کند، که در واقع مقدار چشمگیری است.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.873 ^a	.762	.749	874.779

a. Predictors: (Constant), Income

جدول بعدی ANOVA نام دارد. این جدول نشان می دهد که آیا مدل رگرسیون می تواند به طور معناداری (و مناسبی) تغییرات متغیر وابسته را پیش بینی کند. برای بررسی معناداری به ستون آخر جدول (sig) نگاه می کنیم. این ستون معناداری آماری مدل رگرسیون را نشان می دهد که چنانچه میزان بدست آمده کمتر از 0.05 باشد نتیجه می گیریم که مدل به کار رفته، پیش بینی کننده خوبی برای متغیر قیمت ماشین است. میزان معناداری در مثال ما (P < 0.01)، کمتر از میزان 0.05 است که بیانگر این است که مدل رگرسیونی معنادار است.

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	4.418E7	1	4.418E7	57.737	.000 ^a
	Residual	1.377E7	18	765238.393		
	Total	5.796E7	19			

a. Predictors: (Constant), Income

b. Dependent Variable: Price

جدول زیر Coefficients، اطلاعاتی را در مورد متغیرهای پیش بین به ما می دهد. این جدول اطلاعات ضروری برای پیش بین متغیر وابسته را در اختیار ما قرار می دهد. مشاهده می کنیم که مقدار ثابت (constant) و متغیر درآمد هر دو در مدل معنادار شده اند (به ستون Sig) نگاه می کنیم. پس از تعیین معنادار بودن مقدار ثابت و متغیر درآمد، ستون Standardized Coefficients (Beta) بیانگر ضریب رگرسیونی استاندارد شده یا مقدار بتا است. ضریب رگرسیونی استاندارد شده یا Beta در مثال ما برابر شده است با ۰.۸۷۳. که نشان گر میزان تاثیر متغیر مستقل (درآمد) بر وابسته (قیمت ماشین) است.

نکته: میزان Beta زمانی که تنها یک متغیر مستقل در مدل داریم، با میزان همبستگی پیرسون بین دو متغیر (R) دقیقا برابر است. اما زمانی که بیش از یک متغیر مستقل داشته باشیم، میزان Beta با ضریب همبستگی بین متغیرهای مستقل و وابسته تفاوت خواهد داشت.

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	8286.786	1852.256		4.474	.000
	Income	.564	.074	.873	7.598	.000

a. Dependent Variable: Price

نکته: کاربرد ضریب Beta هنگامی است که بیش از یک متغیر مستقل در مدل داشته باشیم. در این صورت مقدار Beta به ما کمک می کند که سهم نسبی هر متغیر مستقل را در پیش بینی متغیر وابسته مقایسه کنیم و به عبارتی بتوانیم تعیین کنیم که کدام متغیرها بیش ترین تاثیر را بر متغیر وابسته دارند.

جهت ایجاد معادله رگرسیونی از میزان ضریب رگرسیونی استاندارد نشده (B) استفاده می کنیم. معادله رگرسیونی جهت پیش بینی دقیق مقادیر متغیر وابسته مورد استفاده قرار می گیرد و معادله آن به صورت زیر است.

$$Y = a + bX$$

Y = مقدار پیش بینی شده ی متغیر وابسته (قیمت ماشین)

a = عرض از مبدأ نقطه تقاطع خط رگرسیون با محور Y (در جدول: مقدار ثابت یا constant)

b = شیب خط (ضریب رگرسیونی استاندارد نشده یا B)

X = مقادیر مختلف متغیر مستقل

$$\text{Price} = 8287 + 0.564 (\text{Income})$$

معادله رگرسیونی مثال اجرا شده:

در معادله بالا با قرار دادن مقادیر درآمد افراد، می توان قیمت ماشینی که آنان تمایل به خریدنش را دارند پیش بینی کرد. توجه داشته باشیم که در این معادله مقادیر a و b ثابت است و تنها مقادیر متغیر درآمد تغییر می کند

مرکز خدمات آماری خوارزمی

انجام تحلیل آماری پایان نامه کارشناسی ارشد و دکترا و مقالات ISI

با نرم افزارهای SPSS – LISREL – AMOS – PLS – Eviews و شبکه های عصبی با Matlab

ایمیل: RKarimi777@yahoo.com

سایت: www.Kharazmi-Statistics.ir

www.SPSS100.ir

رامین کریمی: ۰۹۱۲۷۶۹۴۰۶۶

مؤلف کتاب "راهنمای آسان تحلیل آماری با SPSS"