

madsg.com  
مادسیج

## انتشارت الکترونیکی مادسیج

برای دسترسی به دانلود هر قسمت کلیک کنید

[کتاب الکترونیکی بانک اساتید](#) [اطلاعات و اخبار](#) [مقالات رایگان](#)

[مدیریت](#)

[کتاب الکترونیکی بانک اساتید](#) [اطلاعات و اخبار](#) [مقالات رایگان](#)

[روانشناسی](#)

[کتاب الکترونیکی بانک اساتید](#) [اطلاعات و اخبار](#) [مقالات رایگان](#)

[علوم تربیتی](#)

[کتاب الکترونیکی تافل](#) [آیلتس](#) [نرم افزار یادگیری زبان](#)

[آموزش زبان](#)

[کتاب الکترونیکی مقالات isi](#) [پروپوزال و پایان نامه](#) [پرسشنامه](#)

[روش تحقیق](#)

[ورود به شبکه مادسیج](#)

مادسیج منتظر کتاب های تالیفی شما هست!!  
ebook@madsg.com

**بررسی**  
**رگرسیون خطی و براورد منحنی**  
**در نرم افزار**  
**spss**

تهیه کننده:

مریم جعفرزاده

## رگرسیون

- تحلیل رگرسیون یک از روشهای پرکاربرد در مطالعات اجتماعی - اقتصادی است.

- این روش ارتباط تنگاتنگی با ضریب همبستگی داشته و عموماً به طور همزمان در مطالعات مورد استفاده قرار می‌گیرد.

## کاربرد رگرسیون

رگرسیون این امکان را برای محقق فراهم می‌کند تا تغییرات متغیر وابسته را از طریق متغیرهای مستقل پیش‌بینی و سهم هر یک از متغیرهای مستقل را در تبیین متغیر وابسته تعیین کند.

بررسی مقدار تاثیر متغیر  $X$  در متغیر  $Y$  را تبیین می‌نامند و هرچه قدرت تبیین  $Y$  توسط  $X$  بیشتر باشد انتظار می‌رود متغیر  $X$  بتواند در آینده با دقت بیشتری متغیر  $Y$  را پیش‌بینی کند.

## تاریخچه تحلیل رگرسیون

مفهوم رگرسیون برای اولین بار توسط فرانسیس گالتون در سال ۱۸۷۷ مورد استفاده قرار گرفت. او در مطالعه خود نشان داد که قد کودکان متولد شده از والدین بلند قامت گرایش به برگشت به متوسط قد افراد جامعه دارد. او فرآیند پیش‌بینی یک متغیر را از طریق متغیر دیگر را با مفهوم رگرسیون بیان کرد.

## تحلیل رگرسیون و ضریب همبستگی

رگرسیون رابطه نزدیکی با ضریب همبستگی دارد. بدین معنا که برای انجام رگرسیون باید ضریب همبستگی را محاسبه کرد. اگر میان متغیرهای مورد مطالعه، همبستگی وجود داشت، تنها در این صورت است که می‌توانیم از رگرسیون برای آزمون فرضیه‌های تحقیق استفاده نماییم.

## مفروضات داده‌ای

- فاصله ای بودن سطح سنجش داده ها

در صورت عدم تحقق این شرط، باید حتما آنها را به متغیر تصنعی (dummy variable) تبدیل نمود.

- نرمال بودن توزیع متغیرهای وابسته

### رگرسیون خطی

- یکی از تکنیک های پیچیده آماری که معمولا در سطح سنجش فاصله ای است.

### رگرسیون خطی ساده

- پیش بینی مقدار یک متغیر وابسته براساس مقدار یک متغیر مستقل
- رگرسیون چند متغیره
- تحلیل مشارکت جمعی و فردی دویا چند متغیر مستقل در تغییرات یک متغیر وابسته

### روش های رگرسیون خطی

- روش همزمان (Enter method)
- روش گام به گام (Stepwise method)
- روش حذف (Remove method)
- روش پس رونده (Backward method)
- روش پیش رونده (Forward method)

### روش همزمان (Enter method)

در این روش کلیه متغیرهای مستقل به طور همزمان وارد مدل می شوند تا تاثیر کلیه متغیرهای مهم و غیر مهم برمتغیر وابسته مشخص گردد.

نکته 

مشکل این روش آنست که چون تمامی متغیرها بدون توجه به ضریب همبستگی شان با متغیر وابسته وارد معادله می شوند، بنابراین متغیرهایی هم که حضورشان در معادله معنی دار نیست، باقی می ماند.

### روش گام به گام (Stepwise method)

- این روش مانند روش Forward متغیرها را یک به یک وارد مدل می کند
- ابتدا متغیری که بیشترین همبستگی را با متغیر وابسته دارد انتخاب می کند.
- دومین متغیری که وارد تحلیل می شود، متغیری است که پس از تفکیک متغیر مقدم بر آن، موجب بیشترین افزایش در مقدار ضریب تعیین ( $R^2$ ) می شود.
- در این روش، ورود متغیرها به مدل را یک به یک و تازمانی انجام می دهیم که معنی داری متغیر به ۹۵ درصد برسد؛ سپس عملیات متوقف می شود.
- در هر مرحله متغیر یا متغیرهایی باقی می ماند که پس از تفکیک مشارکت متغیرهای مقدم بر آن ها دارای بیشترین مجذور همبستگی ( $R$ ) باشند.
- در این روش کارایی هر مرحله با آزمون  $F$  سنجیده می شود.

نکته 

° فرق این روش با Forward در آن است که در روش Forward متغیرهای وارد شده در تحلیل معادله باقی می ماند، ولی در روش Stepwise، با ورود متغیر جدید متغیرهایی که قبلا وارد معادله شده اند از نو آزموده می شوند تا مشخص شود آیا هنوز هم حضور آن ها در مدل، به موفقیت آن کمک می کند یا نه.

### روش حذف (Remove method)

- با این روش می توان متغیرهای یک بلوک راز مدل رگرسیونی حذف کرد.

- این روش را نمی توان به عنوان روش اولین بلوک به کاربرد. زیرا می بایست متغیرها در یکی از بلوک های قبلی وارد مدل شوند و سپس در بلوکهای بعدی با انتخاب این روش آن ها را حذف نمود.

#### نکته

این روش کاربرد چندانی در رگرسیون چند متغیره ندارد، چون تحلیل واریانس را انجام نمی دهد.

#### روش پس رونده یا غربالی (Backward method)

- در این روش همانند روش Enter، ابتدا کلیه متغیرهای مستقل وارد معادله شده و اثر کلیه متغیرها بر روی متغیر وابسته سنجیده می شود.
- برخلاف روش Enter، در این روش به مرور متغیرهای ضعیف تر و کم اثرتر یکی پس از دیگری از معادله خارج شده و در نهایت این مرحله تا زمانی ادامه می یابد که خطای آزمون معنی داری به ۱۰ درصد برسد.

#### روش پیشرو (Forward method)

- در این حالت متغیرهای مهم شناخته می شوند و در معادله نهایی باقی می مانند روش پیش رونده
- این روش ابتدا همبستگی ساده بین هریک از متغیرهای مستقل را با متغیر وابسته محاسبه می کند. سپس متغیر مستقلی که بیشترین همبستگی را با متغیر وابسته دارد و به عبارتی بیشترین مقدار واریانس آن را تبیین می کند، وارد تحلیل می کند.
- دومین متغیری که وارد تحلیل می شود، متغیری است که پس از تفکیک متغیر اول بیشترین ضریب همبستگی را با متغیر وابسته دارد.
- این روش تا زمانی ادامه دارد که خطای آزمون به ۵ درصد برسد. در اینجا انجام رگرسیون متوقف می شود.

نکته 

نتایج به دست آمده همانند روش Stepwise می باشد، بنابراین نحوه تفسیر خروجی این دو یکسان می باشد.

### رگرسیون خطی ساده / دو متغیره (Simple Linear Regression)

- در این نوع رگرسیون، مقادیر یک متغیر (وابسته یا  $Y$ ) از روی مقادیر متغیر دیگر (مستقل یا  $X$ ) به کمک یک معادله خطی برآورد می شود.
- این معادله شکل کلی زیر را دارد:

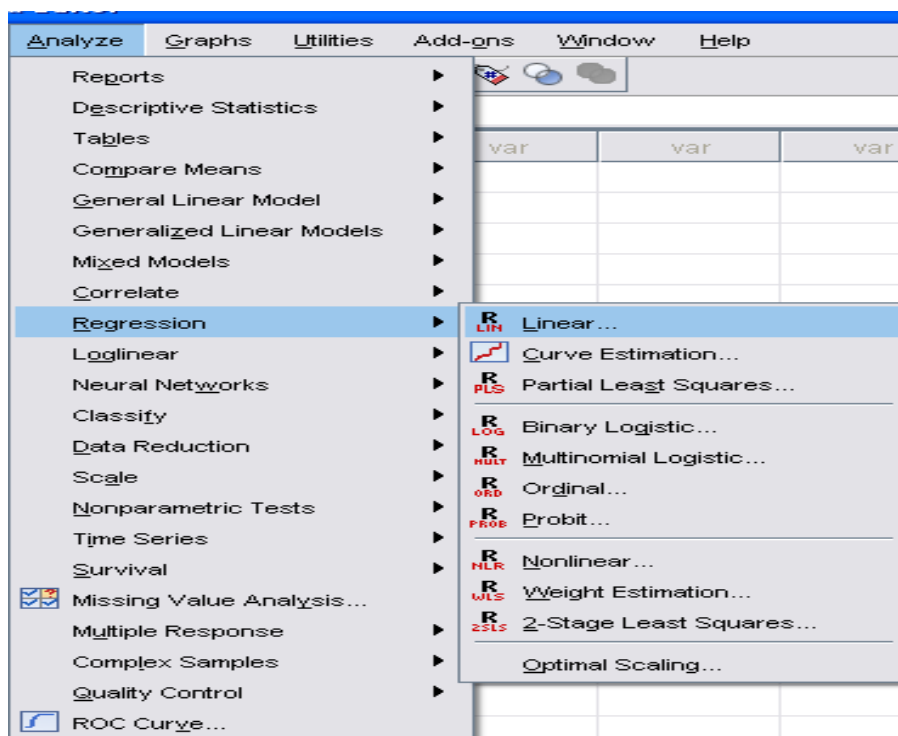
$$Y' = b_0 + b_1(x_1)$$

که  $Y'$  مقدار برآورد شده  $y$ ،  $b_0$  عرض از مبدا (مقدار ثابت رگرسیون)،

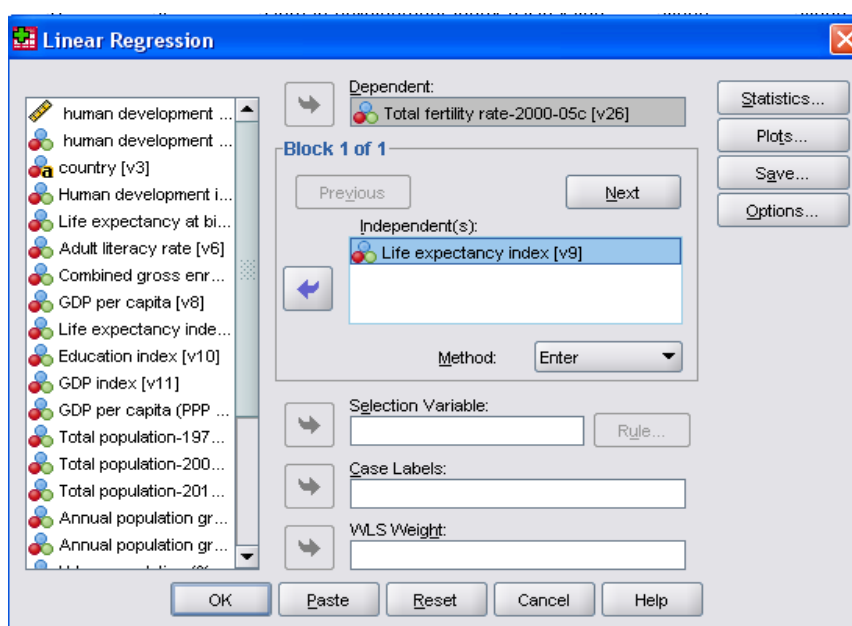
$b_1$  ضریب زاویه (ضریب رگرسیون) و  $x_1$  متغیر اول است.

## نحوه اجرای رگرسیون

۱- ابتدا دستور Analyze>Regression>linear... را اجرا می‌کنیم.

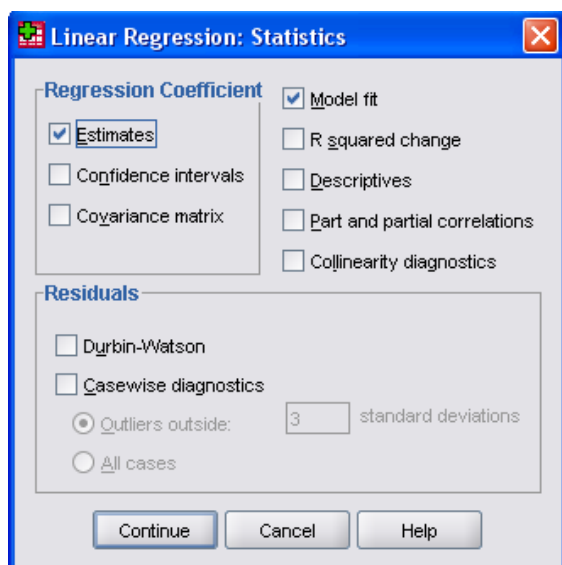


۲- متغیر وابسته را وارد کادر Dependent و متغیر مستقل را وارد کادر Independent می‌کنیم.





۳- دکمه Statistic را کلیک کرده و گزینه های Estimate و Model fit را در قسمت Regression coefficient انتخاب و سپس بر روی دکمه continue کلیک می کنیم (همچنین در این قسمت می توانیم با انتخاب descriptive، آماره های توصیفی و میزان همبستگی بین دو متغیر را به دست آوریم).



## linear Regression :staistics

Regression coefficients گزینه های این کادر عبارتند از:

Estimates- این گزینه برای تمام متغیر های مستقل در مدل، ضرایب استاندارد شده، خطای معیار، مقدار  $t$  و سطح معنی داری دودامنه برای مقدار  $t$  را محاسبه می کند. در مورد متغیر های مستقلی که در مدل نیستند موارد زیر محاسبه می شود ضریب استاندارد شده آن را در صورتی که در مدل می بود، مقدار  $t$  سطح معنی داری  $t$ ، همبستگی نسبی متغیر وابسته با کنترل متغیر مستقل در معادله، و مقدار حداقل تحمل (  $\text{minmum tolerance}$ ). این آماره ها به صورت پیش گزیده محاسبه می شوند. در صورتی که نمی خواهید محاسبه بشوند این گزینه را غیر فعال کنید

Confidence intervals- این گزینه برای تمام ضرایب رگرسیون استاندارد نشده فاصله اطمینان ۰.۹۵ را ارائه

دهد می دهد

Model fit - با این گزینه اماره  $R^2$ ,  $R^2$ ,  $R^2$  تعدیل شده و خطای معیار برآورد به اضافه جدول Anova ارائه می شود. این اماره ها به صورت پیش گزیده ارائه می شوند. در صورتی که نمی خواهید ارائه بشوند این گزینه را غیر فعال کنید.

R squared change - این گزینه میانگین تغییر  $R^2$  را که با اضافه کردن یا کم کردن یک متغیر مستقل از معادله رگرسیون به وجود می آید، نمایش می دهد.

Descriptives - این گزینه میانگین و انحراف معیار تمام متغیرهای انتخاب شده و ماتریس همبستگی را ارائه می دهد

Part and partial correlation - این گزینه انواعی از ضرایب همبستگی را محاسبه می کند که همبستگی بین دو متغیر وابسته و مستقل را با توجه به اثر سایر متغیرهای مستقل اندازه گیری می کند

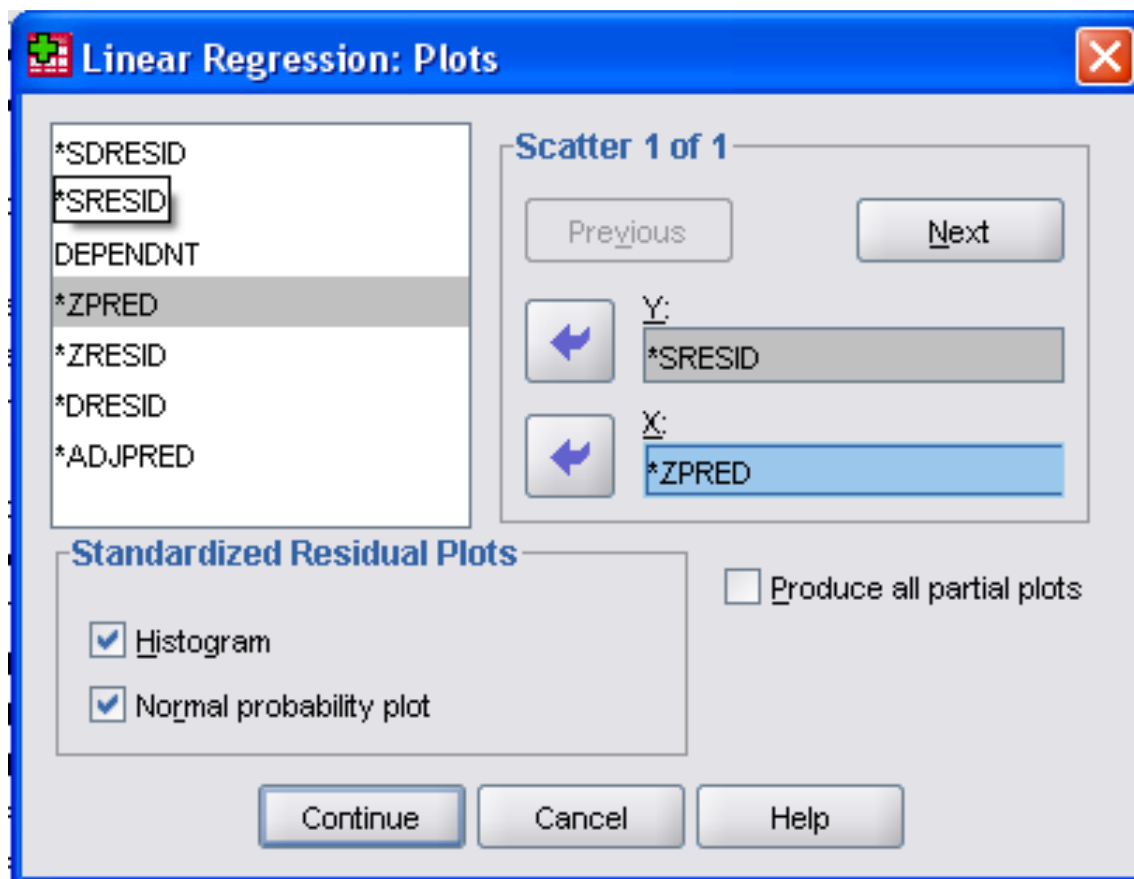
Collinearity diagnostics - این گزینه در رگرسیون چند گانه چندین اماره و آزمون را برای بررسی collinearity بین متغیرهای مستقل ارائه می کند.

Durbin-Watson این گزینه اماره های توصیفی برای باقیمانده های استاندارد شده و نشده و مقادیر پیش بینی شده ارائه می کند

علاوه بر این اماره Durbin - Watson که همبستگی سریال بین باقیمانده ها را اندازه گیری می کند ارائه می دهد

Casewise diagnostics این گزینه لیستی از مقادیر متغیرهای انتخاب شده در نمونه هایی را نشان می دهد که قدر مطلق مقدار باقیمانده استاندارد شده بزرگتر از یک نقطه قطع است

Residual plots انالیز باقیمانده



دکمه plots در کادر اصلی را کلیک کرده و SRESID را به عنوان متغیر وابسته (Y) و \*ZPRED را به عنوان متغیر مستقل (X) انتخاب می‌کنیم. در همین کادر، گزینه‌های Histogram و Normal probability plot را انتخاب و سپس بر روی دکمه continue کلیک می‌کنیم.

برای این نمودارها متغیرهای زیر در دسترس هستند

Zpred - مقادیر پیش‌بینی شده استاندارد شده برای متغیر وابسته

Zresid - باقیمانده استاندارد شده

Dresid - باقیمانده پاک شده، باقیمانده یک نمونه هنگامی که از محاسبه رگرسیون کنار گذاشته شده است

Adjpred - مقادیر پیش بینی شده تعدیل شده شده، مقادیر پیش بینی شده نمونه هنگامی که از محاسبه رگرسیون کنار گذاشته شده است

Sresid - باقیمانده استاندارد شده

Dresid باقیمانده پاک شده استاندارد شده

سایر گزینه های کادر گفتگو linear Regression plots عبارتند از :

Produce all partial plots - این گزینه نمودارهایی را ایجاد می کند که به عنوان نمودار تشخیص در آنالیز رگرسیون به کار می روند . در معادله هایی که حد اقل دومتغیر مستقل دارند برای هر یک از این متغیرها نموداری تشکیل می دهد

Standardized Residual plots - در این کادر نمودار های زیر در دسترس است:

Histogram - این گزینه هیستوگرامی از مقادیر استاندارد شده رسم می کند که جهت بررسی نرمال بودن توزیع به کار می رود

Normal probability plot - این گزینه نمودار  $p-p$  از توزیع باقیمانده های استاندارد شده در مقایسه با یک توزیع نرمال استاندارد رسم می کند

## linear Regression save

۵- دکمه save در کادر اصلی را کلیک نموده و گزینه Standard و Standardized همچنین گزینه های cooks و Leverage values را انتخاب و سپس، روی دکمه continue و سپس ok کلیک می کنیم.

برای به دست آوردن نمودار پراکنش بین هر کدام از آنها، یکی از آنها را انتخاب کرده به کادر Y منتقل کرده و سپس متغیر دیگر را انتخاب کرده و به کادر X منتقل نمایید

برای به دست آوردن نمودارهای دیگر بر روی Nex کلیک کرده و جفت دیگری را انتخاب نمایید برای مرور کردن یا تغییر دادن نمودارهای قبلی بر روی دکمه previous کلیک کنید تا به نمودار مورد نظرتان برسید سپس می توانید متغیر X یا Y انتخاب کرده وان را از کادر خارج کنید و در صورت تمایل متغیر دیگر را جایگزین کنید

**Linear Regression: Save**

**Predicted Values**

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

**Residuals**

- Unstandardized
- Standardized
- Studentized
- Deleted
- Studentized deleted

**Distances**

- Mahalanobis
- Cook's
- Leverage values

**Influence Statistics**

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

**Prediction Intervals**

- Mean
- Individual

Confidence Interval:  %

**Coefficient statistics**

- Create coefficient statistics
- Create a new dataset  
Dataset name:
- Write a new data file

**Export model information to XML file**

- Include the covariance matrix

تفسیر خروجی های رگرسیون خطی ساده:

### Variables Entered/Removed<sup>b</sup>

Model	Variables Entered	Variables Removed	Method
1	Life expectancy index <sup>a</sup>		Enter

a. All requested variables entered.

b. Dependent Variable: Total fertility rate-2000-05c

جدول فوق ه عنوان اولین خروجی، تعداد مدل، متغیرهای وارد شده و متغیرهای خارج شده به مدل و همچنین روش. ورداستفاده برای ورود متغیرها به مدل را نشان می دهد

جدول بعدی خلاصه آماره های مربوط به برازش مدل را نشان می دهد:

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.796 <sup>a</sup>	.634	.632	1.04404

a. Predictors: (Constant), Life expectancy index

b. Dependent Variable: Total fertility rate-2000-05c

- Model: تعداد مدل های تشکیل شده را نشان می دهد که در این مثال برابر با یک مدل است.
  - R: به ضریب همبستگی چندگانه معروف است و میزان همبستگی چندگانه بین مجموعه متغیرهای مستقل و وابسته را نشان میدهد. مقدار این ضریب بین (۰) تا (۱) نوسان دارد. هرچه مقدار این ضریب به (۱) نزدیکتر باشد، همبستگی بین متغیرهای مستقل و متغیر وابسته قوی تر و هرچه به صفر نزدیکتر باشد ضعیف تر است.
  - R Square: به مجذور ضریب همبستگی چندگانه یا ضریب تعیین معروف بوده و با علامت  $R^2$  نوشته می شود. این ضریب میزان تبیین واریانس و تغییرات متغیر وابسته توسط مجموعه متغیرهای مستقل را نشان می دهد. مقدار این ضریب نیز بین (۰) تا (۱) نوسان دارد. هرچه مقدار این ضریب به (۱) نزدیکتر باشد، نشان از آن دارد که متغیرهای مستقل توانسته اند میزان زیادی از متغیر وابسته را تبیین کنند و برعکس، هرچه مقدار ضریب تعیین به (۰) نزدیکتر باشد، دلالت بر نقش کم تر متغیرهای مستقل در تبیین واریانس متغیر وابسته دارد.
- اشکال ضریب تعیین این است که میزان موفقیت مدل را بیش از اندازه برآورد می کند؛ همچنین ضریب تعیین تعداد درجات آزادی را به حساب نمی آورد.

#### 📖 نکته

همیشه مدلی مناسب تر است که از یک طرف مقدار ضریب تعیین آن بالاتر باشد و از طرف دیگر متغیرهای زیادی را شامل نشود.

نسبت تعداد متغیرهای مستقل بر حجم نمونه ( $\frac{k}{N}$ ) یکی از عوامل موثر بر افزایش  $R^2$  است. این نسبت باید از  $\frac{1}{30}$  کم تر باشد. یعنی به ازای هر ۳۰ نفر پاسخگو (آزمودنی)، حداقل یک متغیر مستقل لازم است.

- Adjusted R square: ضریب تعیین تعدیل شده نام دارد و آن را با  $R^2_{adj}$  نشان می دهند. برای تفسیر ضریب تعیین، معمولاً از این مقدار استفاده می شود. چون در این ضریب، مقدار ضریب تعیین با درجات آزادی تعدیل شده است.

در جدول بالا مقدار ضریب تعیین شده ۰.۶۳۲ است یعنی متغیر مستقل ۶۳.۲ درصد از تغییرات متغیر وابسته را تبیین می کند. و مابقی این تغییرات (۳۶.۸)، که به مجذور کمیت خطا ( $e^2$ ) معروف است؛ تحت تاثیر متغیرهای خارج از مدل می باشد.

#### نکته

برخی معتقدند  $R^2$  تعدیل شده میزان  $R^2$  را که برای نمونه محاسبه شده است، اصلاح و تعدیل می کند و به کل جامعه تعمیم می دهد. بنابراین  $R^2$  برای نمونه آماری و  $R^2$  تعدیل شده برای جامعه آماری است. زمانی که می خواهیم دقت مدل را ارزیابی کنیم از  $R^2$  استفاده می کنیم اما اگر هدف مقایسه ۲ یا چند مدل باشد از رگرسیون تعدیل شده استفاده می کنیم

- Std. Error of the Estimate: خطای استاندارد برآورد نشان دهنده میزان قدرت پیش بینی معادله رگرسیون است. و شاخصی است که تغییر پذیری اندازه های پیش بینی شده  $y$  را در اطراف خط رگرسیون نشان می دهد. هرچه اشتباه استاندارد برآورد کوچکتر باشد پیش بینی دقیق تر و ضریب همبستگی قوی تر است

#### ANOVA<sup>b</sup>

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	321.304	1	321.304	294.766	.000 <sup>a</sup>
	Residual	185.305	170	1.090		
	Total	506.609	171			

a. Predictors: (Constant), Life expectancy index



جدول فوق نتایج تحلیل واریانس را نشان می دهد. در این جدول منبع تغییرات متغیر وابسته در دو منبع رگرسیون (Regression) و باقیمانده (Residual) نشان داده شده است و برای هر یک از این منابع، مجموع مجذورات، درجه آزادی و میانگین مجذورات آمده است.

- منبع رگرسیون: اطلاعات مربوط به میزان تغییرات متغیر وابسته را که در نتیجه مدل تحقیق ماست، نشان می دهد.
- منبع باقی مانده: اطلاعات مربوط به میزان تغییرات متغیر وابسته را که در نتیجه مدل تحقیق ماست نشان میدهد.

بنابراین هرچه مقدار مجموع مجذورات باقیمانده کوچکتر از مجموع مجذورات رگرسیون باشد، نشان دهند قدرت تبیین گیری بالای مدل در توضیح تغییرات متغیر وابسته است. برعکس، هرچه مقدار باقیمانده به میزان بیشتری از رگرسیون بزرگتر باشد، نشان میدهد که نقش مدل در تبیین تغییرات متغیر وابسته ضعیف است. در چنین حالتی، باید متغیرهای مستقل ضعیف تر را از مدل حذف کرد و متغیرهای مستقل دیگری را که نقش بیش تری در تبیین تغییرات متغیر وابسته دارند، وارد مدل می شوند.

- ردیف مربوط به کل (Total) کل اطلاعات مربوط به هر دو منبع تغییر رگرسیون و باقیمانده را نشان می دهد.
- در ستون چهارم درجه آزادی مربوط به دو منبع رگرسیون و باقیمانده نشان داده شده است:  
 ۱) درجه آزادی رگرسیون ( $DF_{reg}$ ) برابر است با تعداد متغیرهای مستقلی ( $K$ ) که در معادله باقی مانده اند. در این مثال، مقدار این درجه آزادی برابر است با (۱) چون تنها یک متغیر مستقل (شاخص امید به زندگی) داریم.

$$DF_{reg} = K$$

- ۲) درجه آزادی باقیمانده ( $DF_{res}$ ) برابر است با تعداد نمونه منهای تعداد متغیر مستقل باقیمانده در معادله منهای یک.

$$DF_{res} = \mathcal{N} \cdot \mathcal{K} - 1$$

در جدول Anova می‌توانیم معنی دار بودن مدلها را بررسی کنیم.

مقدار F نشان دهنده آن است که آیا مدل رگرسیونی تحقیق مدل مناسبی است یا خیر. به عبارتی، آیا متغیرهای مستقل قادر به خوبی تغییرات متغیر وابسته را توضیح دهند یا خیر. تشخیص این موضوع با معنی دار بودن مقدار F در سطح خطای کوچکتر یا بزرگتر از ۰.۰۵ امکان پذیر است مقدار F از تقسیم میانگین مجزورات رگرسیون (MSR) بر میانگین مجزورات باقیمانده به دست می‌آید

### Coefficients<sup>a</sup>

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	7.688	.273		28.152	.000
	Life expectancy index	-6.621	.386	-.796	-17.169	.000

a. Dependent Variable: Total fertility rate-

2000-05c

جدول فوق میزان تاثیر گذاری متغیر یا متغیرهای وابسته، معنی دار بودن متغیر مستقل در مدل ونیز معادله رگرسیون را نشان می‌دهد

اولین اماری که در این جدول به چشم می‌خورد عدد ثابت است که همان عرض از مبدا می‌باشد و میزان دخالت متغیرهای مستقل (یعنی زمانی که میزان متغیرهای مستقل صفر است) نشان می‌دهد

۱) ضرایب تاثیر رگرسیون دو دسته اند ضریب تاثیر استاندارد نشد  $(B)^2$

۲) ضرایب تاثیر رگرسیون استاندارد شده یا بتا (Beta)

$(B)^2$  ضرایب رگرسیونی استاندارد نشده : مربوط به مدل رگرسیونی برآورد شده می باشد اما از آنجا که در تحلیل رگرسیون مقیاس اغلب متغیرهای مستقل از واحدهای متفاوتی تشکیل شده بنابراین به راحتی نمی توان به مقایسه سهم هر متغیر مستقل در تبیین تغییرات یا واریانس متغیر وابسته پرداخت از همین رو ضرایب رگرسیونی استاندارد شده (Beta) استفاده می شود

ضرایب رگرسیونی استاندارد شده (Beta) : برای تعیین اهمیت و نقش متغیر یا متغیرهای مستقل در پیشگویی معادله رگرسیون از مقادیر بتا استفاده می شود. از طریق آن می توان در مورد اهمیت نسبی متغیرها قضاوت کرد بزرگ بودن مقدار بتا نشان دهنده اهمیت نسبی و نقش آن در پیشگویی متغیر وابسته می باشد لذا در تفسیر نتایج رگرسیون بهتر است به ضرایب رگرسیونی استاندارد شده اشاره کنیم تا غیر استاندارد.

ضریب رگرسیون استاندارد شده براساس مقادیر انحراف استاندارد سنجیده می شود که در تفسیر آن باید رعایت کنیم.

اماره اهمیت نسبی حضور هر متغیر مستقل در مدل را نشان می دهد. برای اینکه تشخیص دهیم کدام متغیرها تاثیر آماري معنی داری بر متغیر وابسته داشتند می توانیم به مقدار آنگاه کنیم در ضمن، در هنگام تفسیر مقدار آبه علامت مثبت یا منفی آن توجه نکرده و تنها قدر مطلق آن را گزارش می دهیم

تفاوت همبستگی تفکیکی (partial correlation) و همبستگی نیمه تفکیکی (partial correlation)

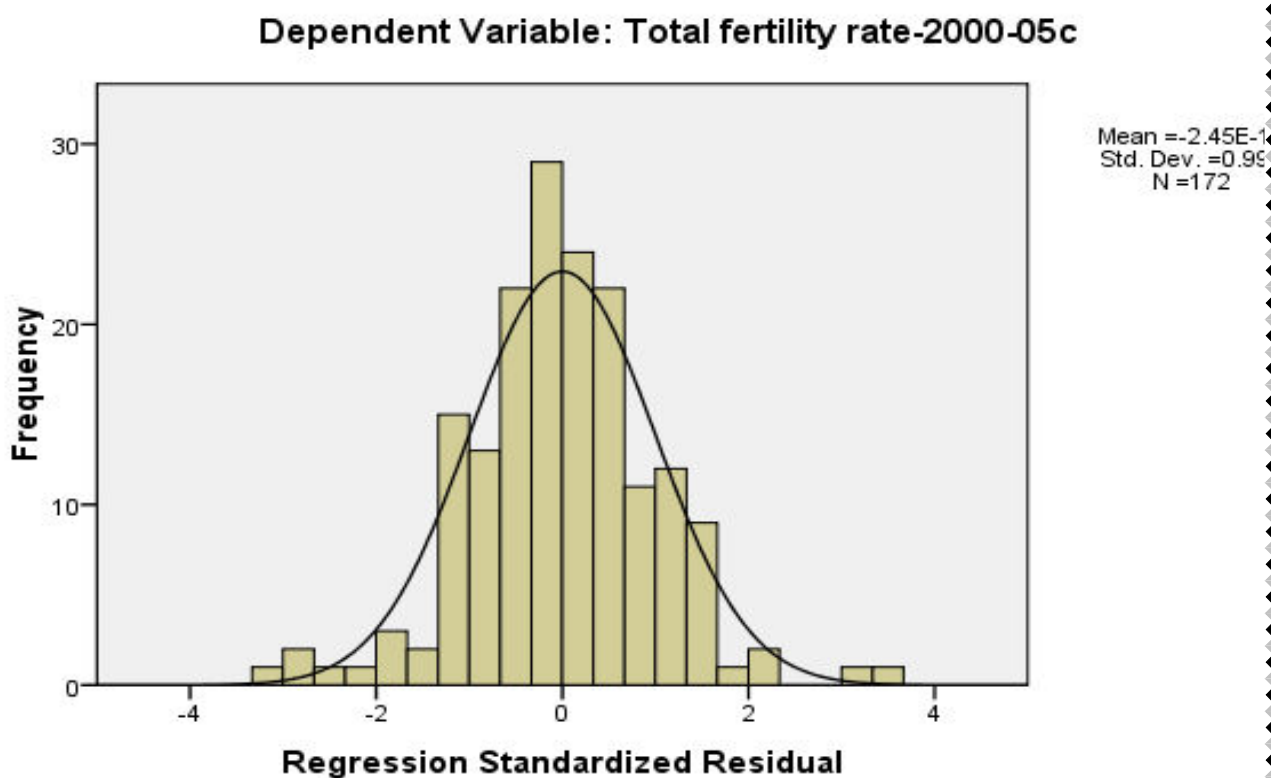
همبستگی تفکیکی یک شاخص کنترل آماری است که به وسیله آن اعتبار روابط بیرونی متغیرها مورد آزمون قرار می گیرد. همبستگی تفکیکی برآورد دقیقی از همبستگی میان متغیر وابسته و متغیر مستقل مربوط را پس از تفکیک اثر متغیر یا متغیرهای دیگر را بر روی آن دو نشان می دهد.

درهمبستگی تفکیکی اثر متغیر یا متغیرهای کنترلی، روی متغیرهای وابسته و مستقل ثابت نگهداشته می شود ولی در همبستگی نیمه تفکیکی فقط اثر متغیر یا متغیرها روی متغیر مستقل مورد نظر کنترل می شود

## اماره Tolerance

برای بررسی هم خطی بودن مدل از این اماره استفاده می شود هر چه مقدار این اماره به صفر نزدیکتر شود نشان دهنده عدم هم خطی بودن مدل می باشد

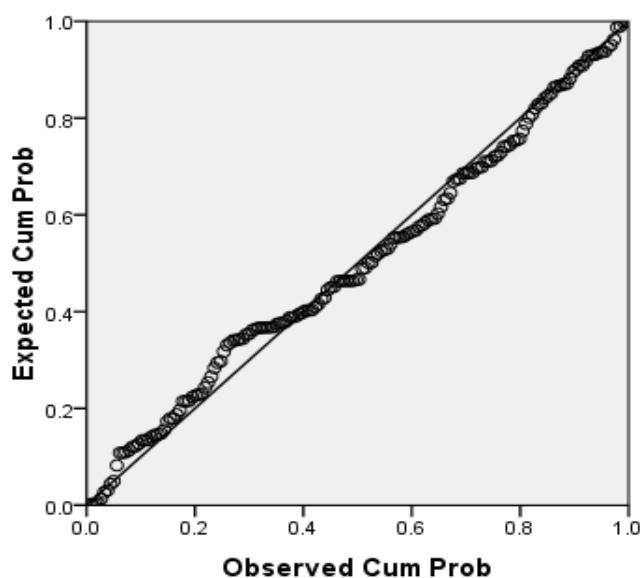
### Histogram



Normal p-p plot در نمودار زیر نقاطی که روی قطر قرار می گیرند، نشان می دهند که احتمال تجمعی شده با احتمال تجمعی مورد نظر یکسان است در واقع هر قدر تجمع نقاط حول قطر بیشتر باشد دقیق تر می توان متغیر وابسته را پیش بینی کرد. و همچنین از این نمودار برای سنجش نرمال بودن متغیرها استفاده می شود هرچه نقاط به حول قطر بیشتر باشند و به قطر نزدیکتر باشند نشان دهنده نرمال بودن متغیرها می باشد

### Normal P-P Plot of Regression Standardized Residual

Dependent Variable: Total fertility rate-2000-05c



# Curve Stimulation

رگرسیون برآورد منحنی از خانواده تحلیل رگرسیون غیر خطی است. این نوع رگرسیون زمانی مورد استفاده قرار می‌گیرد که رابطه بین متغیر وابسته و متغیر مستقل به صورت غیر خطی است و بنابراین، نمی‌توانیم از رگرسیون نوع خطی استفاده کنیم. برای اینکه پی ببریم رابطه بین متغیر وابسته و متغیرهای مستقل از نوع خطی است یا غیر خطی، می‌توانیم از یک نمودار پراکنش استفاده کنیم.

رگرسیون برآورد منحنی شامل ۱۱ نوع رگرسیون غیر خطی است که با اجرای آن می‌توانیم بهترین مدل رگرسیونی برای مشاهدات مان را انتخاب کنیم. این ۱۱ نوع مدل رگرسیونی در نرم افزار SPSS عبارتند از:

۱) رگرسیون خطی

۲) رگرسیون لگاریتمی

۳) رگرسیون معکوس

۴) رگرسیون سهمی با درجه ۲

۵) رگرسیون درجه ۳

۶) رگرسیون توانی

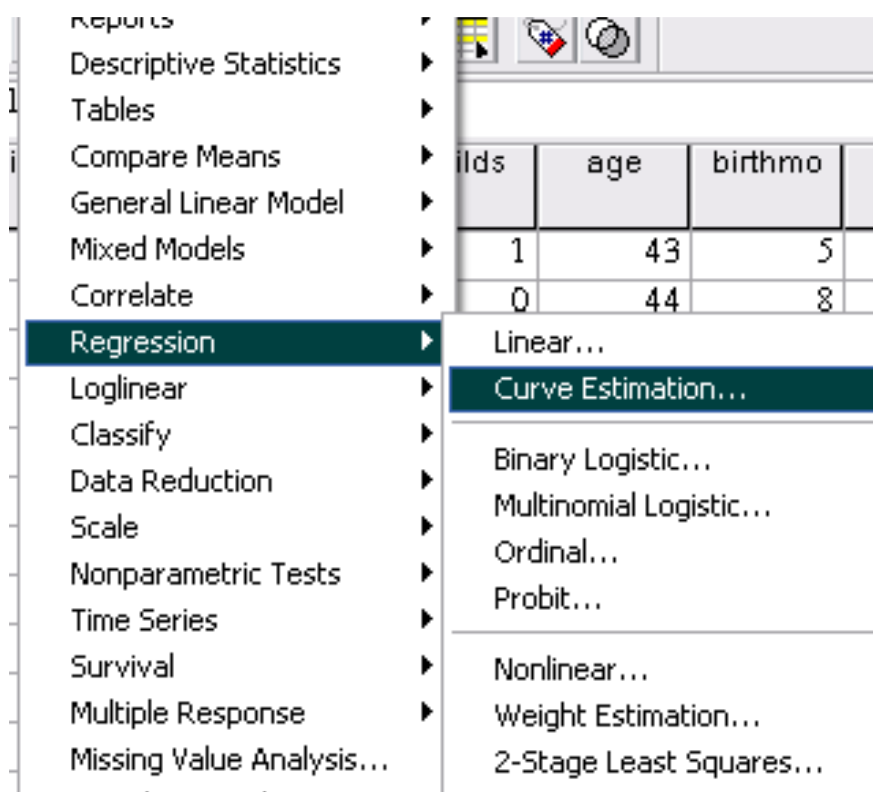
۷) رگرسیون مرکب

۸) رگرسیون منحنی S

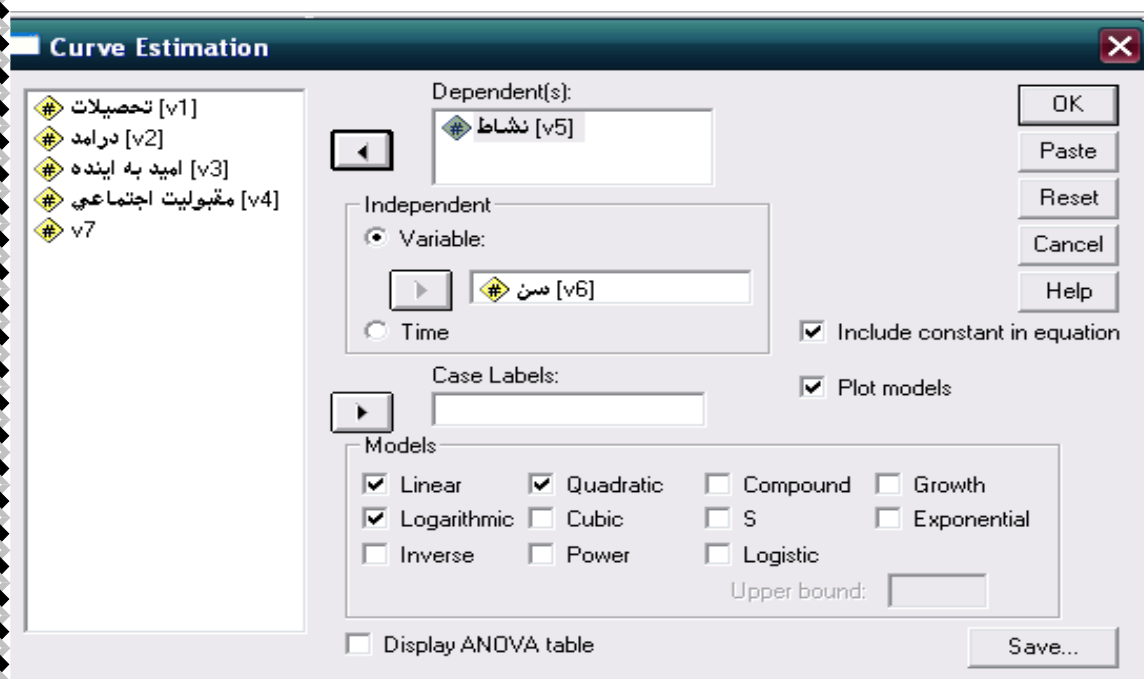
۹) رگرسیون لجستیک

مثال به نظر می‌رسد سن شهروندان برنشاط آنها موثر است

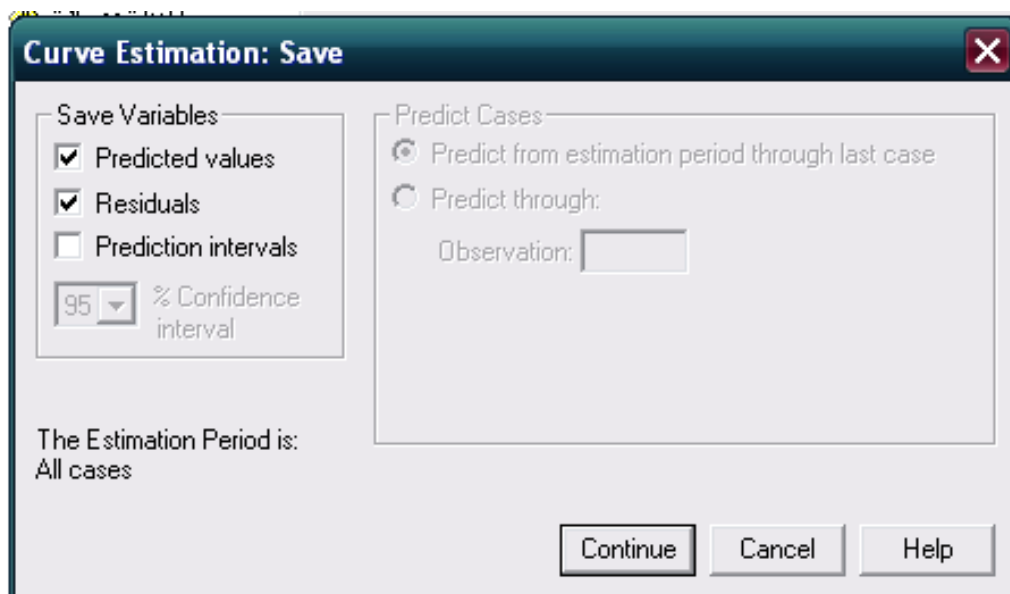
۱) دستور Analyze>Regression>Curve Estimation را اجرا می‌کنیم



متغیر وابسته نشاط را وارد کادر dependent و متغیر مستقل (سن=Age) را وارد کادر independent می‌کنیم. سپس مدل‌های (خطی) Linear، (لجستیک) logarithmic، (سه‌می) quadratic را به عنوان مدل رگرسیونی انتخاب می‌کنیم



۳) دکمه save را کلیک کرده و گزینه Residuals و predicted values را انتخاب می کنیم.



۴) در کادر اصلی دستور بر روی دکمه OK کلیک می کنیم

جدول زیر نتایج کلی سه مدل رگرسیونی خطی (Linear) لگاریتمی را نشان می دهد

این جدول همچنین نام متغیر وابسته (نشاط) و متغیر مستقل (سن) آمده است

Model Name		MOD_1
Dependent Variable	1	نشاط
Equation	1	Linear
	2	Logarithmic
	3	Quadratic
Independent Variable		سن
Constant		Included
Variable Whose Values Label Observations in Plots		Unspecified
Tolerance for Entering Terms in Equations		.0001



جدول بعدی نشان می دهد که تعداد کل پاسخگویان برابر با ۲۰۰۰ نفر است. اما در هنگام تحلیل ۲۳ نفر آن به دلیل داشتن مقادیر گمشده از تحلیل خارج شده اند

### Case Processing Summary

	N
Total Cases	2000
Excluded Cases <sup>a</sup>	23
Forecasted Cases	0
Newly Created Cases	0

a. Cases with a missing value in any variable are excluded from the analysis.

نتیجه جدول زیر به طور دقیق تر نشان می دهد که ۲۳ نفر از پاسخگويانی که دارای مقدار گمشده بوده اند در متغیر مستقل (سن) بوده است و هیچ پاسخگویی در متغیر وابسته (نشاط) دارای مقادیر گمشده نمی باشد

### Variable Processing Summary

	Variables	
	Dependent	Independent
	نشاط	سن
Number of Positive Values	2000	1977
Number of Zeros	0	0
Number of Negative Values	0	0
Number of Missing Values	User-Missing	0
	System-Missing	0

جدول بعد و نمودار مربوط به آن، مهمترین جدول در تفسیر نتایج دستور رگرسیون برآورد منحنی است. جدول زیر میزان برازش هر سه مدل را از مومن می کند. در این جدول سطح معنی دار از مومن F برای هر سه مدل خطی

لگاریتمی و سهمی کوچکتر از ۰.۰۵ است در چنین شرایطی، برای اینکه بهترین مدل را انتخاب کنیم، باید مدلی را برگزینیم که مقدار اماره F در آن بیشترین مقدار را دراد

همانطور که در جدول ملاحظه می شود مقدار اماره F برای مدل لگاریتمی بیشتر از دو مدل دیگر است بنابراین می توان گفت که مدل رگرسیون لگاریتمی برازنده ترین مدل برای پیش بینی تغییرات متغیر وابسته نشاط برحسب متغیر سن است

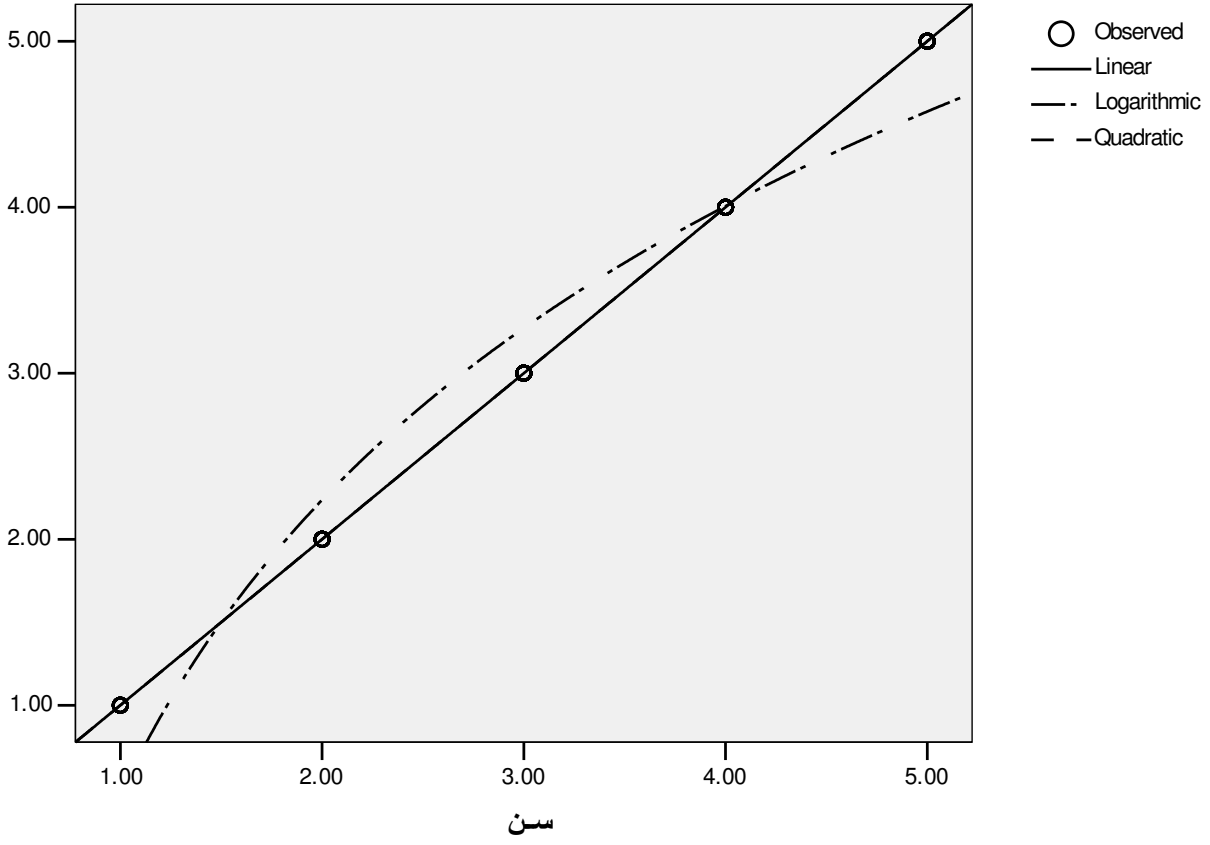
### Model Summary and Parameter Estimates

Dependent Variable: نشاط

Equation	Model Summary					Parameter Estimates		
	R Square	F	df1	df2	Sig.	Constant	b1	b2
Linear	.022	43.509	1	1975	.000	34.921	-.061	
Logarithmic	.026	52.173	1	1975	.000	40.532	-2.230	
Quadratic	.026	26.551	2	1974	.000	37.368	-.212	.002

The independent variable is سن.

نشاط



## منابع

- ۱- معادل سازی معدلات ساختاری در تحقیقات اجتماعی اقتصادی، خلیل کلانتر فرهنگ صبا ۱۳۸۸
- ۲- حبیب پور، کرم و صفری، رضا، (۱۳۸۸) راهنمای جامع کاربرد **spss** در تحقیقات پیمایشی، تهران: لویه، متفکران.
- ۳- استفاده از **Help** نرم افزار **spss**
- ۴- نقل قولهای دکتر جباری
- ۵- استفاده از اسلایدهای دکتر حیدری چروده
- ۶- **Help spss**