



فهرست دروس

آموزش نرم افزار آماری Minitab «» رگرسیون

امکانات دانلود: PDF Web Zip

کار با Minitab	
محاسبات در Minitab	
انجام کار های آماری در Minitab	
نمودارهای کنترل	
قابلیت فرآیند	
آنالیز واریانس	
طراحی آزمایشات	
رگرسیون	
قابلیت اطمینان	
کار با نمودارها	

محاسبه رگرسیون	
رگرسیون	
بدست آوردن رگرسیون خطی	
رسم نمودار باقیمانده	

محاسبه رگرسیون

برای آشنایی محاسبه معده رگرسیون در Minitab ابتدا مثالی را بصورت زیر مورد بررسی قرار می دهیم و سپس برای این مثال معادله رگرسیونی را تهیه می نمائیم .

مثال : در سال ۱۹۲۰ برای بدست آوردن ارزش یک هکتار از زمینهای کشاورزی شهر (آبوا) متغیرهای زیر در نظر گرفته شده اند :

متوسط ارزش هر هکتار بر حسب دلار : Y

متوسط برداشت سالانه در ۱۰ سال قبل : X1

درصد دانه گندم به کشتزار : X2

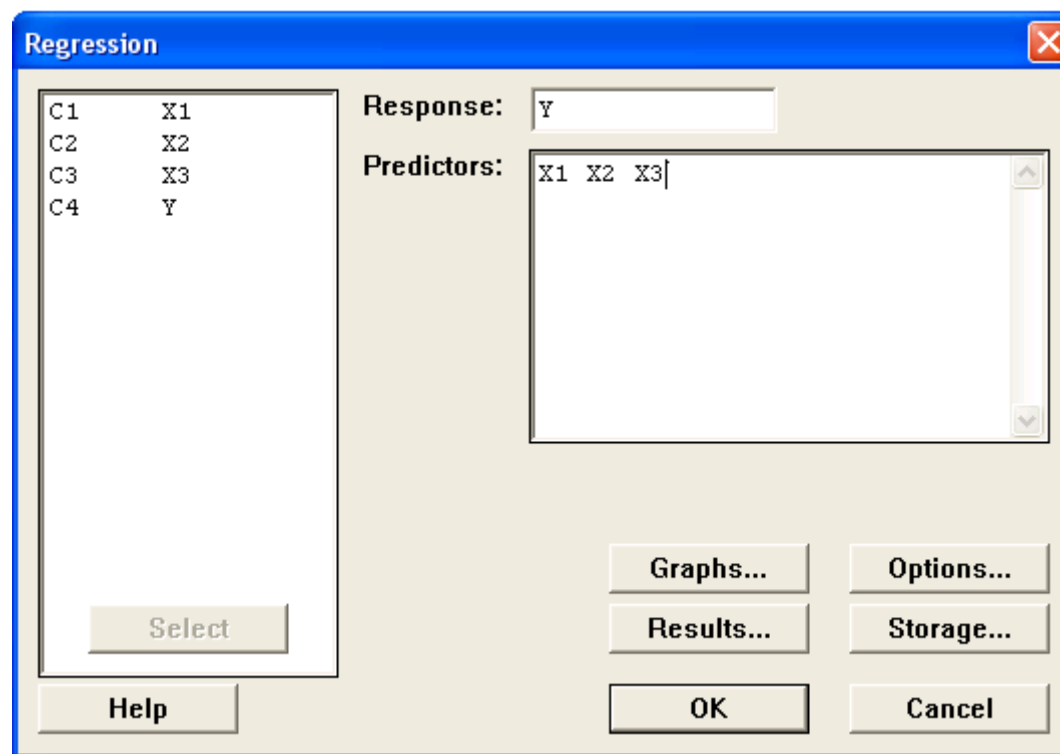
درصد گندم در کشتزار : X3

حل معادله ارزش هر هکتار را بر حسب متغیرهای X1,X2,X3 بدست آورید.

برای بدست آوردن معادله رگرسیون بصورت زیر عمل نمائید : (می توانید فایل Response.mtw را ابتدا باز نموده بصورت زیر عمل نمائید .)

Stat -----> Regression-----> Regression

حال در قسمت Response (شکل ۱-۸) ستون Y را وارد نمائید و در قسمت Predictors ستونهای X1,X2,X3 را وارد نموده و برای دیدن خروجی بر روی دکمه OK کلیک نمائید. خروجی نرم افزار بصورت شکل ۲-۸ خواهد بود.



شکل ۱-۸

شکل مربوط به رگرسیون

Regression Analysis: Y versus X1, X2, X3

The regression equation is

$$Y = -140 + 2.71 X1 + 6.33 X2 + 4.11 X3$$

Predictor	Coef	SE Coef	T	P
Constant	-140.44	61.45	-2.29	0.033
X1	2.713	1.884	1.44	0.165
X2	6.328	1.689	3.75	0.001
X3	4.111	1.526	2.69	0.014

S = 37.02 R-Sq = 79.8% R-Sq(adj) = 76.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	113676	37892	27.65	0.000
Residual Error	21	28780	1370		
Total	24	142456			

Source	DF	Seq SS
X1	1	41587
X2	1	62141
X3	1	9948

Unusual Observations

Obs	X1	Y	Fit	SE Fit	Residual	St Resid
5	39.0	363.00	259.24	10.23	103.76	2.92R

شکل ۸-۲ خروجی نرم افزار برای رگرسیون چند متغیره

ابتدای صفحه

رگرسیون

روش Stepwise

در این روش ابتدا برای تمام مقادیر F آمار F حساب شده و کوچکترین آنها که از مقدار

$F_{to\ remove}$ کوچکتر است حذف می شود ، Minitab کوچکترین مقادیر F را حذف می نماید ، سپس آمار F جدیدی برای آنها پی که در جدول فعلی وجود ندارند حساب می شود و اگر یکی از مقادیر دارای F بزرگتری از عدد $F_{to\ enter}$ قسمت Option باشد در نظر گرفته می شود. در آخر بزرگترین این مقادیر در شکل گیری مدل در نظر گرفته می شود.

روش FORWARD

در این روش متغیرها به مدل اضافه می شوند ولی هرگز از مدل حذف نمی گردد . این روش زمانی به پایان می رسد که هیچ متغیری از مقدار $F_{to\ enter}$ بزرگتر نباشد ، برای اینکه مسئله از این روش حل شود . باید مقدار $F_{to\ remove}$ را صفر وارد نمایید.

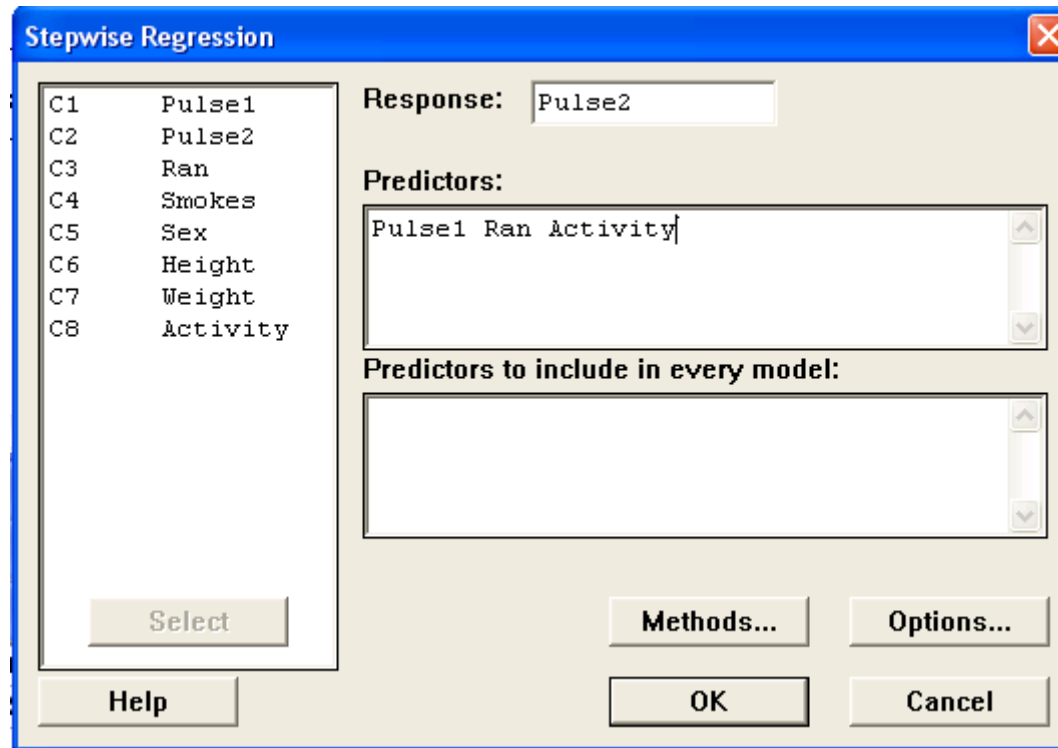
روش BACKWARD

در این روش مسئله با تمام متغیرهای ممکن شروع به حل می کند و شروع به حذف متغیرهای ممکن می کند ، به طوری که در انتها هیچ متغیری باقی نمانده باشد که مقدار F آن از مقدار $F_{to\ remove}$ کمتر باشد . برای آشنایی بیشتر مسئله ای را بصورت زیر حل نمایید: ابتدا فایل PULSE.MTW را باز

نموده و سپس بصورت زیر عمل نمائید :

Stat-----> Regression----->Stepwise Regression

متغیر PULSE 2 را به عنوان متغیر پاسخ در قسمت Response (شکل ۸-۳) وارد نمائید و در قسمت Predictors متغیرهای Activity , Ran, Pulse1 را وارد کنید.



شکل ۸-۳

شکل مربوط به رگرسیون به روش Stepwise Regression

حال بر روی دکمه Ok کلیک نمائید خروجی نرم افزار بصورت شکل ۴-۸ خواهد بود.

Stepwise Regression: Pulse2 versus Pulse1, Ran, Activity

Alpha-to-Enter: 0.15 Alpha-to-Remove: 0.15

Response is Pulse2 on 3 predictors, with N = 92

Step	1	2	3
Constant	10.28	44.48	51.84
Pulse1	0.957	0.912	0.902
T-Value	7.42	9.74	9.73
P-Value	0.000	0.000	0.000
Ran		-19.1	-19.2
T-Value		-9.05	-9.19
P-Value		0.000	0.000
Activity			-3.1
T-Value			-1.83
P-Value			0.071
S	13.5	9.82	9.70
R-Sq	37.97	67.71	68.89
R-Sq(adj)	37.28	66.98	67.83
C-p	87.5	5.3	4.0

دوقسمت دیگر در بخش Stepwise وجود دارد که توضیح هر کدام بصورت زیر می باشد :

ENTER : اگر متغیری را در این قسمت وارد کنید. Minitab آنها را بصورت متغیر آغازین در نظر می گیرد و اگر مقدار F آنها از مقدار F to remove قسمت Option کمتر شود ، آنها را حذف می نماید.

FORCE : در این قسمت متغیرهایی را وارد نمائید ، که نمی خواهید در طول حل مسئله از مدل حذف گردند.

روش BEST SUBSET

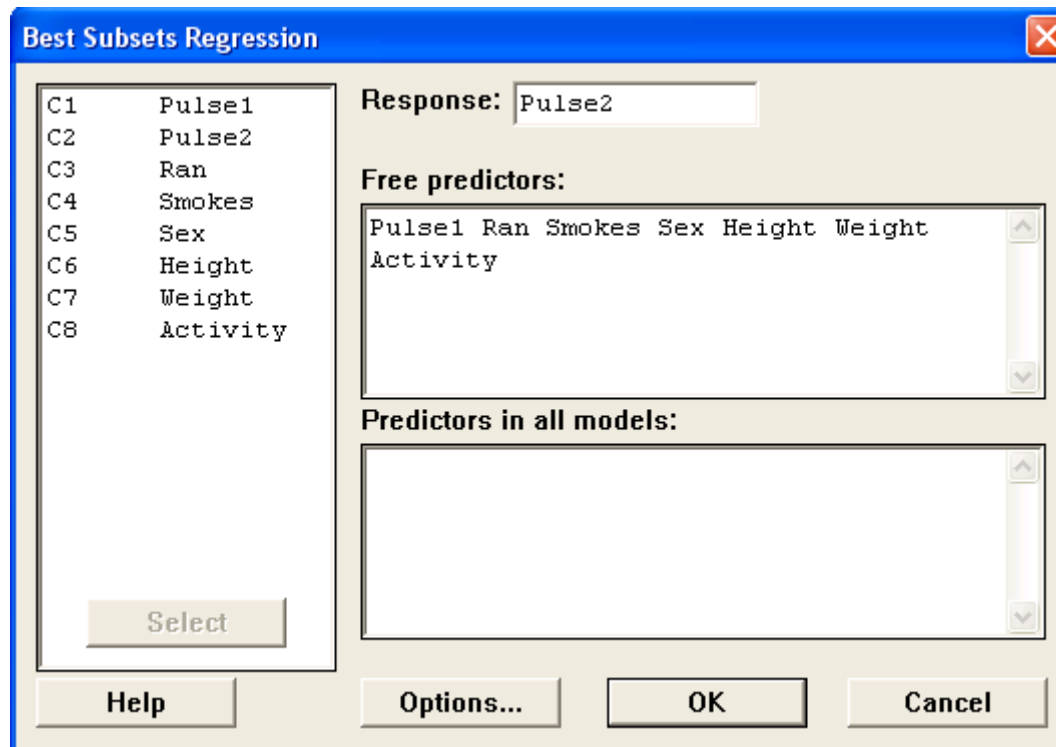
این روش برای بدست آوردن جواب خود از معیار R-Squared استفاده می نماید . لازم به ذکر است که این روش در واقع مقدار R-Square را برای تمام حالت‌های مسئله حساب می کند و ما می توانیم با دیدن مقادیر مختلف R-Square بزرگترین آنها را انتخاب نمائیم و مدل رگرسیون خود را بر اساس متغیرهایی که در بزرگترین مقدار R-Square وجود دارند بسازیم . برای آشنایی بیشتر مثال زیر را مورد بررسی قرار می دهیم. ابتدا فایل PULSE.MTW را باز نموده و بصورت زیر عمل نمائید :

Stat ----->Regression-----> Best Subsets

سپس در قسمت Response (شکل ۵-۸) متغیر پاسخ را وارد نمائید . متغیر پاسخ در اینجا PULSE2 می باشد در قسمت Free Predictors متغیرهای پیش بینی کننده را وارد نمائید .

لازم به ذکر است متغیرهایی که در این قسمت وارد می کنید ، می توانند از مدل حذف شوند و R-Square های مختلف را حاصل کنند ، در این قسمت متغیرهای PULSE1 , Ran ,

Smokes, Activity, Weight , Sex , Height را وارد نمائید . اگر می خواهید متغیری در تمامی حالات وجود داشته باشد ، آنرا در قسمت Predictors In All Models وارد نمائید . (در اینجا ما همه متغیرها را قابل حذف انتخاب کرده ایم) بدیهی است که اگر متغیری را در این قسمت وارد کردید دیگر نباید آنرا در قسمت Free Predictors وارد کنید.



شکل ۵-۸

شکل مربوط به مثال Subset Regression

سپس بر روی دکمه Ok کلیک نمائید ، خروجی نرم افزار بصورت شکل ۶-۸ خواهد بود .

Best Subsets Regression: Pulse2 versus Pulse1, Ran, ...

Response is Pulse2

Vars	R-Sq	R-Sq(adj)	C-p	S	A c							
					P u l s e	S R k e a	H m o S g e e	W e i i v h h t	T t y			
1	38.0	37.3	108.1	13.538	X							
1	33.3	32.5	122.9	14.041	X							
2	67.7	67.0	16.1	9.8219	X	X						
2	47.2	46.0	80.9	12.560	X	X						
3	72.1	71.2	4.1	9.1751	X	X	X					
3	70.8	69.9	8.1	9.3853	X	X		X				
4	72.9	71.7	3.5	9.0929	X	X	X	X				
4	72.3	71.1	5.4	9.1948	X	X	X	X				
5	73.2	71.7	4.6	9.0951	X	X	X	X	X			
5	73.1	71.6	4.9	9.1136	X	X	X	X	X	X		
6	73.4	71.5	6.2	9.1260	X	X	X	X	X	X	X	
6	73.3	71.4	6.3	9.1349	X	X	X	X	X	X	X	X
7	73.4	71.2	8.0	9.1716	X	X	X	X	X	X	X	X

شکل ۶-۸

خروجی مربوط به مثال Subset Regression

تذکر: اگر بخواهید برای متغیرهای قابل حذف در مدل محدودیتهایی را فائل شوید، می توانید با Minimum , Maximum در قسمت Free Predictors In Each Model که در بخش Option قرار دارد، حداقل و حداکثر آنها را تعیین نمایید.

بدست آوردن رگرسیون خطی ، درجه دو ودرجه سه

برای بدست آوردن معادلات رگرسیونی فوق می توانید بصورت زیر عمل نمایید:

Stat----->Regression -----> Fitted Line Plot

در قسمت Type Of Regression Model می توانید نوع مدل خود را انتخاب نمایید . این مدل می تواند بصورت زیر باشد :

مدل ریاضی

لازم به ذکر است که اگر قسمت Stat----> Fit Intercept علامت خورده باشد آنگاه برای معادلات رگرسیون ضریب ثابت B در نظر گرفته می شود ، ودر غیر این صورت معادلات فوق بدون ضریب ثابت در نظر گرفته می شوند . برای آشنایی بهتر با نحوه کار این قسمت مثال زیر را مورد بررسی قرار داده ایم .

مثال : داده های زیر را در نظر بگیرید ، این داده ها به ترتیب نشان دهنده مقدار ماده افزودنی (X) و زمان خشک شدن رنگ هستند :

X : میزان ماده اضافه شده (گرم)

X	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
Y	۲۸۲	۲۶۰	۲۷۰	۲۲۵	۲۷۰	۲۴۳	۲۹۳	۳۳۸	۳۱۵	۲۹۳

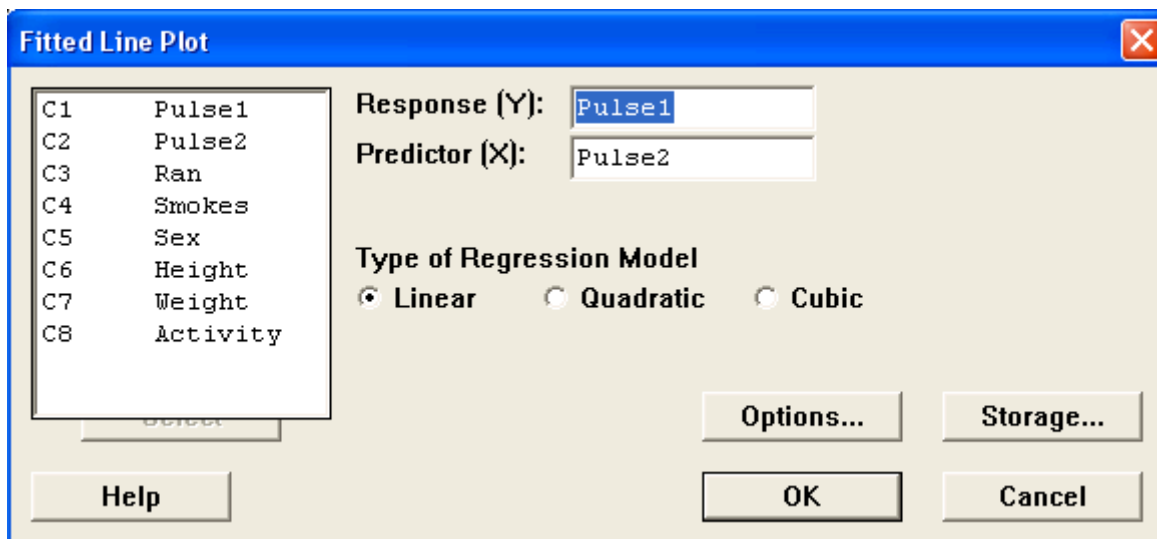
حل معادله رگرسیوني را بصورت خطي ، درجه دو ودرجه سه بدست آورید؟

براي حل این مسئله بصورت زیر عمل نمائید:

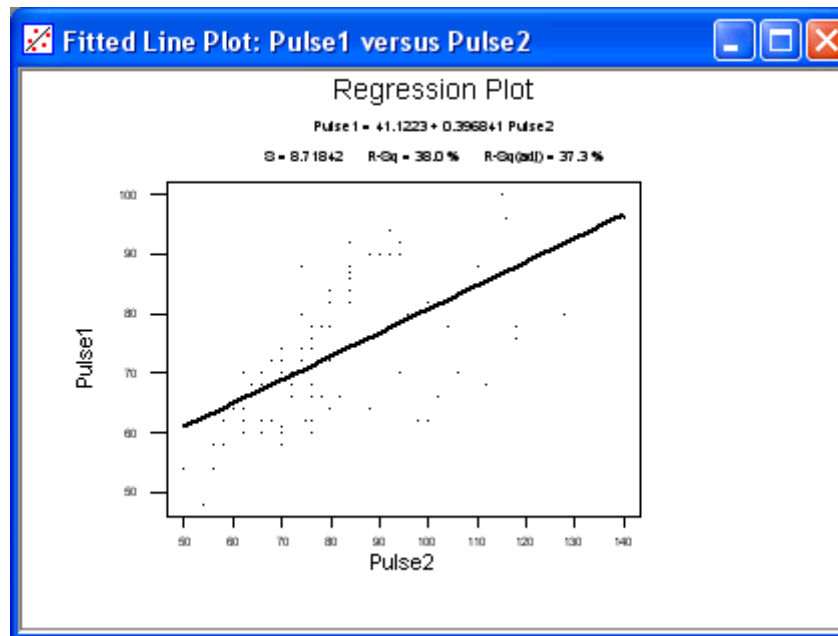
ابتدا مقادیر X را در ستون C1 وارد کنید ومقادیر Y را در ستون C2 وارد نموده ، سپس بصورت زیر عمل نمائید :

Stat-----> Regression-----> Fitted Line Plot

درقسمت Response (Y) (شکل ۸-۷) ستون () Y را وارد نموده و در قسمت Predictor(X) ستون C را وارد نمائید . اکنون یکبار قسمت Linear را علامت زده و بر روی دکمه Ok کلیک نمائید ، خروجي نرم افزار بصورت شکل ۸-۸ و ۸-۹ خواهد بود.



شکل ۷-۸ شکل مربوط به بر آورد رگرسیون خطی ، درجه ۲ ودرجه ۳



شکل ۸-۸ خروجی گرافیکی مربوط به بر آورد رگرسیون خطی

Regression Analysis: Pulse1 versus Pulse2

The regression equation is
Pulse1 = 41.1223 + 0.396841 Pulse2

S = 8.71842 R-Sq = 38.0 % R-Sq(adj) = 37.3 %

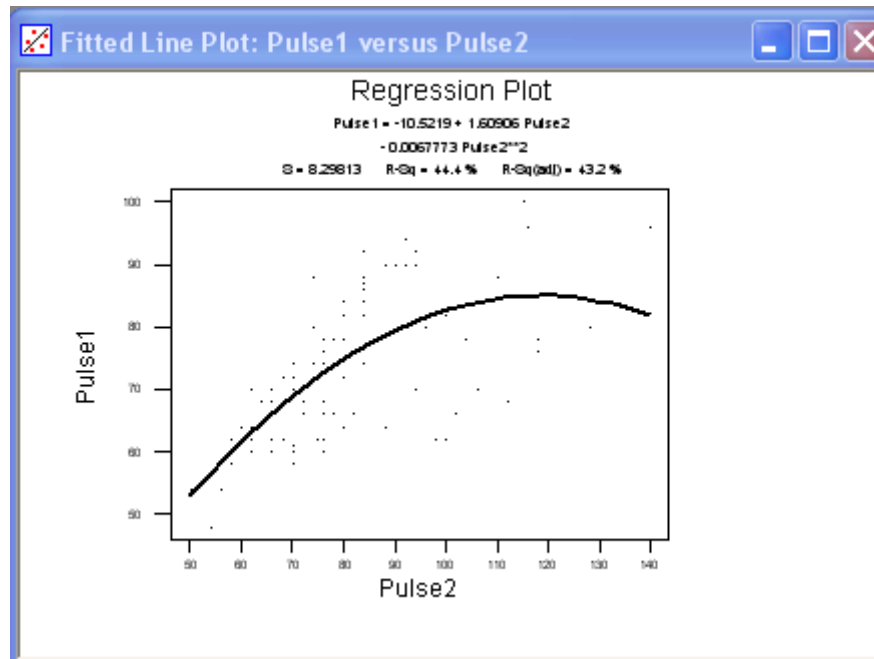
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	4187.5	4187.47	55.0904	0.000
Error	90	6841.0	76.01		
Total	91	11028.4			

Fitted Line Plot: Pulse1 versus Pulse2

شکل ۹-۸ خروجی مربوط به برآورد رگرسیون خطی

وبه همین ترتیب برای مشاهده معادله درجه دو و درجه سه به ترتیب Quadratic و Cubic را علامت بزنید و بر روی دکمه Ok کلیک نمایید. تا نتیجه را ببینید خروجی نرم افزار بصورت شکل‌های ۱۰-۸ و ۱۱-۸ و ۱۲-۸ و ۱۳-۸ خواهد بود.



شکل ۱۰-۸

خروجی گرافیکی مربوط به برآورد رگرسیون درجه ۲

Polynomial Regression Analysis: Pulse1 versus Pulse2

The regression equation is
Pulse1 = -10.5219 + 1.60906 Pulse2
- 0.0067773 Pulse2**2

S = 8.29813 R-Sq = 44.4 % R-Sq(adj) = 43.2 %

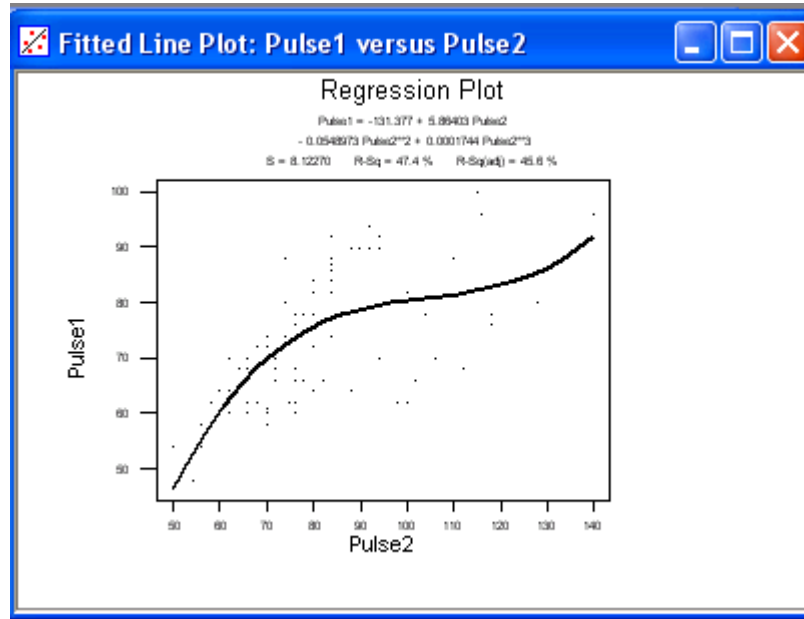
Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	4900.0	2449.99	35.5799	0.000
Error	89	6128.4	68.86		
Total	91	11028.4			

Source	DF	Seq SS	F	P
Linear	1	4187.47	55.0904	0.000
Quadratic	1	712.52	10.3476	0.002

شکل ۱۱-۸

خروجی مربوط به برآورد رگرسیون درجه ۲



شکل ۸-۱۲

خروجی گرافیکی مربوط به برآورد رگرسیون درجه ۳

Polynomial Regression Analysis: Pulse1 versus Pulse2

The regression equation is

$$\text{Pulse1} = -131.377 + 5.86403 \text{ Pulse2} - 0.0548973 \text{ Pulse2}^{**2} + 0.0001744 \text{ Pulse2}^{**3}$$

S = 8.12270 R-Sq = 47.4 % R-Sq(adj) = 45.6 %

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	5222.4	1740.78	26.3842	0.000
Error	88	5806.1	65.98		
Total	91	11028.4			

Source	DF	Seq SS	F	P
Linear	1	4187.47	55.0904	0.000
Quadratic	1	712.52	10.3476	0.002
Cubic	1	322.36	4.8859	0.030

شکل ۱۳-۸

خروجی مربوط به برآورد رگرسیون درجه ۳

ابتدای صفحه

نحوه رسم نمودار باقیمانده ها

برای رسم این نمودار بصورت زیر عمل نمائید :

Stat-----> Regression ----->Residual plots

که در قسمت Residuals ستون باقیمانده ها وارد و در قسمت Fits ستونی که در آن مقدار پیش بینی شده توسط مدل رگرسیونی آمده است را باید وارد نمود .

ابتدای صفحه

ابتدای فصل <=====> محاسبات در Minitab