

11.10

a. $\sum x_i = 3.25, \quad \sum y_i = 7.55, \quad \sum x_i y_i = 2.1825,$

$\sum x_i^2 = 1.2625, \quad \sum y_i^2 = 6.0725, \quad n = 10$

So, $S_{xy} = -.27125, \quad S_{xx} = .20625.$

Then $\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = -1.31515, \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 1.182424.$

OR: Use STATA

Put the data into Excel and save as .csv file. Import the data by clicking File->Import->ASCII data created by a spread sheet, or using command insheet directly.

. insheet using "C:\Documents and Settings\Sara\Desktop\sunfish.csv"
(2 vars, 10 obs)

. regress survivors time

Source	SS	df	MS	Number of obs = 10		
Model	.356734825	1	.356734825	F(1, 8) =	183.94	
Residual	.015515153	8	.001939394	Prob > F	=	0.0000
Total	.372249978	9	.041361109	R-squared	=	0.9583
				Adj R-squared	=	0.9531
				Root MSE	=	.04404

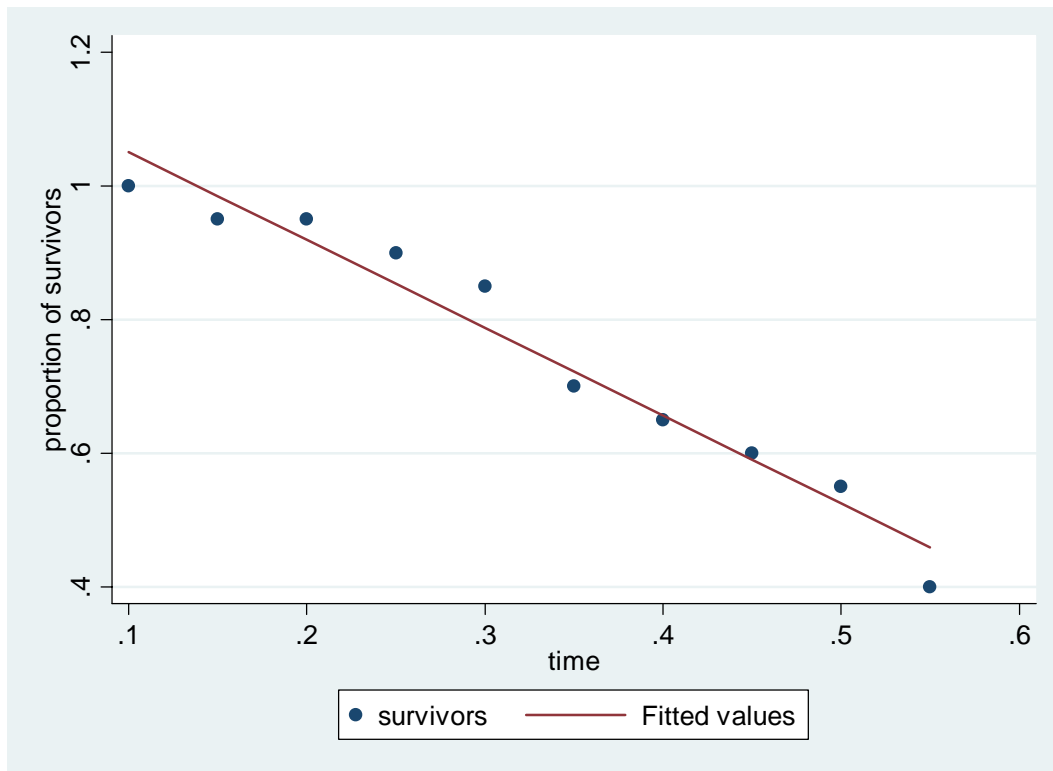
survivors	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
time	-1.315151	.0969697	-13.56	0.000	-1.538764	-1.091539
_cons	1.182424	.034455	34.32	0.000	1.102971	1.261878

So, $\hat{\beta}_1 = -1.31515, \hat{\beta}_0 = 1.182424.$

b.

. predict yhat
(option xb assumed; fitted values)

. scatter survivors time || line yhat time, xtitle(time) ytitle(proportion of survivors)



The regression line fits the data very well.

11.36

Refer to Exercise 11.10, calculate $S_{yy} = .37225$, $\bar{x} = .325$.

Then, $SSE = .0155$, and $S^2 = .0019394$.

Also, you can find these values in STATA.

From the result of the simple regression in 11.10, we see that $S^2 = .0019394$.

. summarize time

Variable	Obs	Mean	Std. Dev.	Min	Max
-----+-----					
time	10	.325	.1513825	.1	.55

So, $S_{xx} = 0.1513825^2 * (n-1) = 0.1513825^2 * 9 = 0.20625$.

For $x = .3$, $\hat{y} = 1.182424 - 1.31515 * (.3) = .7878$

Use the formula on P560, we have the 90% confidence interval for $E(Y)$ is

$$.7878 \pm 1.86 \sqrt{.0019394} \sqrt{\frac{1}{10} + \frac{(.3 - .325)^2}{.20625}} = .7878 \pm .0263.$$

11.43

Refer to Exercise 11.10. When $x = .6$, $\hat{y} = 1.182424 - 1.31515 * (.6) = .3933$.

Use the formula on P564, we have the 95% confidence interval for Y at $x = .6$ is

$$.3933 \pm 2.306 \sqrt{.0019394} \sqrt{1 + \frac{1}{10} + \frac{(.6 - .325)^2}{.20625}} = .3933 \pm .12, \text{ equivalently, } [.27, .51].$$

11.56

Fit the linear model: $y = \beta_0 + \beta_1 x + \beta_2 x^2$ using the method of least squares.

For this exercise:

$$X = \begin{bmatrix} 1 & -3 & 9 \\ 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix}, Y = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \\ -1 \\ 0 \\ 0 \end{bmatrix}.$$

In MATLAB, calculate

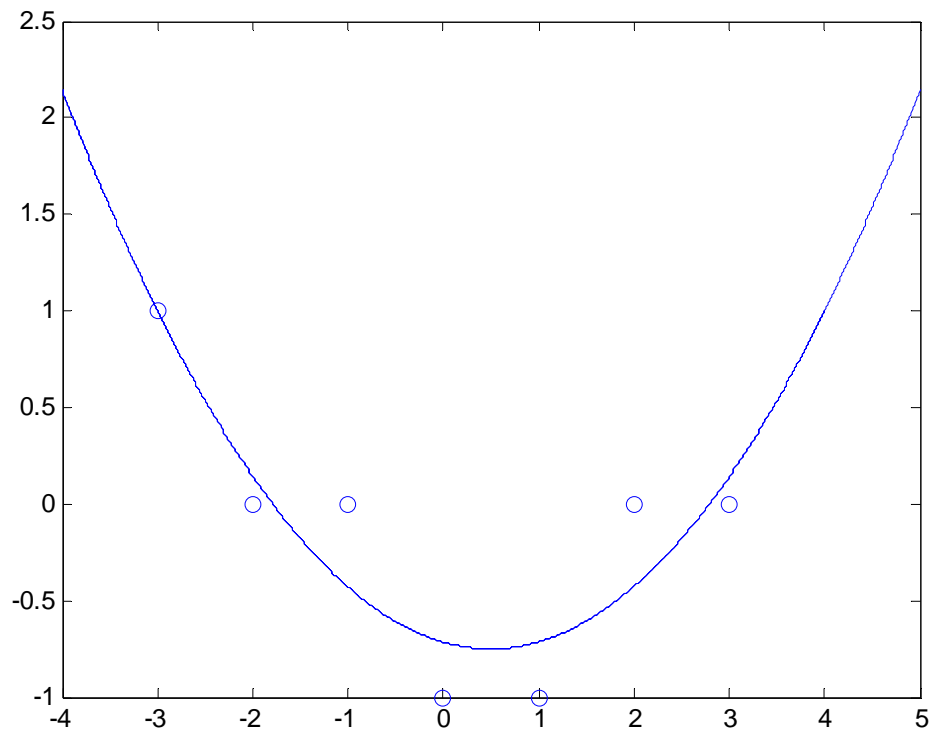
$$X'Y = \begin{bmatrix} -1 \\ -4 \\ 8 \end{bmatrix}, X'X = \begin{bmatrix} 7 & 0 & 28 \\ 0 & 28 & 0 \\ 28 & 0 & 196 \end{bmatrix}, (X'X)^{-1} = \begin{bmatrix} .3333 & 0 & -.0476 \\ 0 & .0357 & 0 \\ -.0476 & 0 & .0119 \end{bmatrix}.$$

$$\text{So, } \hat{\beta} = (X'X)^{-1} X'Y = \begin{bmatrix} -.7143 \\ -.1429 \\ .1429 \end{bmatrix}.$$

The fitted parabola is $\hat{y} = -.71 - .14x + .14x^2$.

$$SSE = Y'Y - \hat{\beta}' X'Y = .5714 \Rightarrow \hat{\sigma}^2 = SSE / (n-3) = .14286.$$

The fitted parabola and raw data are shown as the following:



11.51 a. Let $W = \ln Y$ so that $E(W) = \ln \alpha_0 - \alpha_1 X = \beta_0 + \beta_1 X$, where $\ln \alpha_0 = \beta_0$ & $-\alpha_1 = \beta_1$.

After calculating individual values of W for each of the $n=10$ observations, we have

$$\sum X = 55, \quad \sum W = 35.505412, \quad \sum XW = 194.49729,$$

$$\sum X^2 = 385, \quad \sum W^2 = 126.07188, \quad n = 10,$$

$$S_{XW} = -.782481, \quad S_{XX} = 82.5, \quad S_{WW} = .008448$$

$$\text{Hence, } \hat{\beta}_1 = \frac{S_{XW}}{S_{XX}} = -\frac{.782481}{82.5} = -.00948$$

$$\& \hat{\beta}_0 = \bar{W} - \hat{\beta}_1 \bar{X} = 3.6027$$

Transforming to the original variables, we have

$$\hat{\alpha}_1 = -\hat{\beta}_1 = .00948$$

$$\& \hat{\alpha}_0 = e^{\hat{\beta}_0} = 36.70$$

and the prediction equation is $\hat{Y} = 36.70 e^{-.00948X}$

b. In order to find a CI for α_0 , we first find a CI for β_0 & then transform the endpoints of the interval.

The LSE of β_0 is $\hat{\beta}_0$, which has variance

$$V(\hat{\beta}_0) = \frac{\sigma^2 \sum X_i^2}{n \sum (X_i - \bar{X})^2} = \frac{\sigma^2 \sum X_i^2}{n S_{XX}} \quad (\text{given in Section 11.5})$$

It's necessary to calculate

$$\begin{aligned} SSE &= S_{WW} - \hat{\beta}_1 S_{XW} = .008448 - (-.00948)(-.782481) \\ &= .0010265 \end{aligned}$$

$$\& s^2 = SSE / (n-2) = .0001283$$

Then 90% CI for β_0 is

$$\hat{\beta}_0 \pm t_{.05, 8} \sqrt{V(\hat{\beta}_0)} = 3.6027 \pm 1.86 \sqrt{.0001283 \left[\frac{385}{10(82.5)} \right]}$$

yielding the interval $[3.5883, 3.6171]$

Transforming, we have the interval for α_0 as
 $[e^{3.5883}, e^{3.6171}] = [36.17, 37.23]$

11.52 Similar to Exercise 11.51

Notice $\ln(e^{-\alpha_0 X^{\alpha_1}}) = -\alpha_0 X^{\alpha_1}$,

& $\ln[-\ln(e^{-\alpha_0 X^{\alpha_1}})] = \ln \alpha_0 + \alpha_1 \ln X$

Therefore, we would expect $\ln(-\ln Y)$ to be linear in $\ln X$

Define $V = \ln(X)$, $Z = \ln(-\ln Y)$, $\beta_0 = \ln \alpha_0$, $\beta_1 = \alpha_1$.

We fit the model $E(Z) = \beta_0 + \beta_1 V$

(We are assuming the error remains additive when we take two natural logs, this may not be a valid assumption).

To fit the model, we have

$$\bar{\Sigma} V = -10.15243, \quad \bar{\Sigma} Z = -13.1544, \quad \Sigma VZ = 18.521577$$

$$\bar{\Sigma} V^2 = 12.977909, \quad \bar{\Sigma} Z^2 = 28.50516, \quad n = 9,$$

$$S_{VZ} = 3.6828016, \quad S_{VV} = 1.51548$$

$$\Rightarrow \hat{\beta}_1 = \frac{S_{VZ}}{S_{VV}} = 2.4142$$

$$\hat{\beta}_0 = \bar{Z} - \hat{\beta}_1 \bar{V} = 1.2617$$

Transforming, we have $\hat{\alpha}_1 = \hat{\beta}_1 = 2.4142$

$$\hat{\alpha}_0 = e^{\hat{\beta}_0} = 3.5315$$

The predictor is $\hat{Y} = \exp(-3.5315 X^{2.4142})$