**2.5.1: SAS - Multiple Predictors**

Dr. Bean – Stat 5100

Example: (Exercises 6.15-6.17) A hospital administrator is studying the relation between patient satisfaction (Y, an index) and patient’s age (X1, in years), severity of illness (X2, an index), and anxiety level (X3, an index). Data are reported for 46 randomly selected patients. For all index variables, higher values indicate more (satisfaction, severity, anxiety).

**/\* Input data (see Exercises 6.15-6.17) \*/**

**data patient;**

**input satisfaction age severity anxiety @@; cards;**

**48 50 51 2.3 57 36 46 2.3**

**66 40 48 2.2 70 41 44 1.8**

**89 28 43 1.8 36 49 54 2.9**

**46 42 50 2.2 54 45 48 2.4**

**26 52 62 2.9 77 29 50 2.1**

**89 29 48 2.4 67 43 53 2.4**

**47 38 55 2.2 51 34 51 2.3**

**57 53 54 2.2 66 36 49 2.0**

**79 33 56 2.5 88 29 46 1.9**

**60 33 49 2.1 49 55 51 2.4**

**77 29 52 2.3 52 44 58 2.9**

**60 43 50 2.3 86 23 41 1.8**

**43 47 53 2.5 34 55 54 2.5**

**63 25 49 2.0 72 32 46 2.6**

**57 32 52 2.4 55 42 51 2.7**

**59 33 42 2.0 83 36 49 1.8**

**76 31 47 2.0 47 40 48 2.2**

**36 53 57 2.8 80 34 49 2.2**

**82 29 48 2.5 64 30 51 2.4**

**37 47 60 2.4 42 47 50 2.6**

**66 43 53 2.3 83 22 51 2.0**

**37 44 51 2.6 68 45 51 2.2**

**59 37 53 2.1 92 28 46 1.8**

**;**

**run;**

**/\* Look at scatterplot matrix \*/**

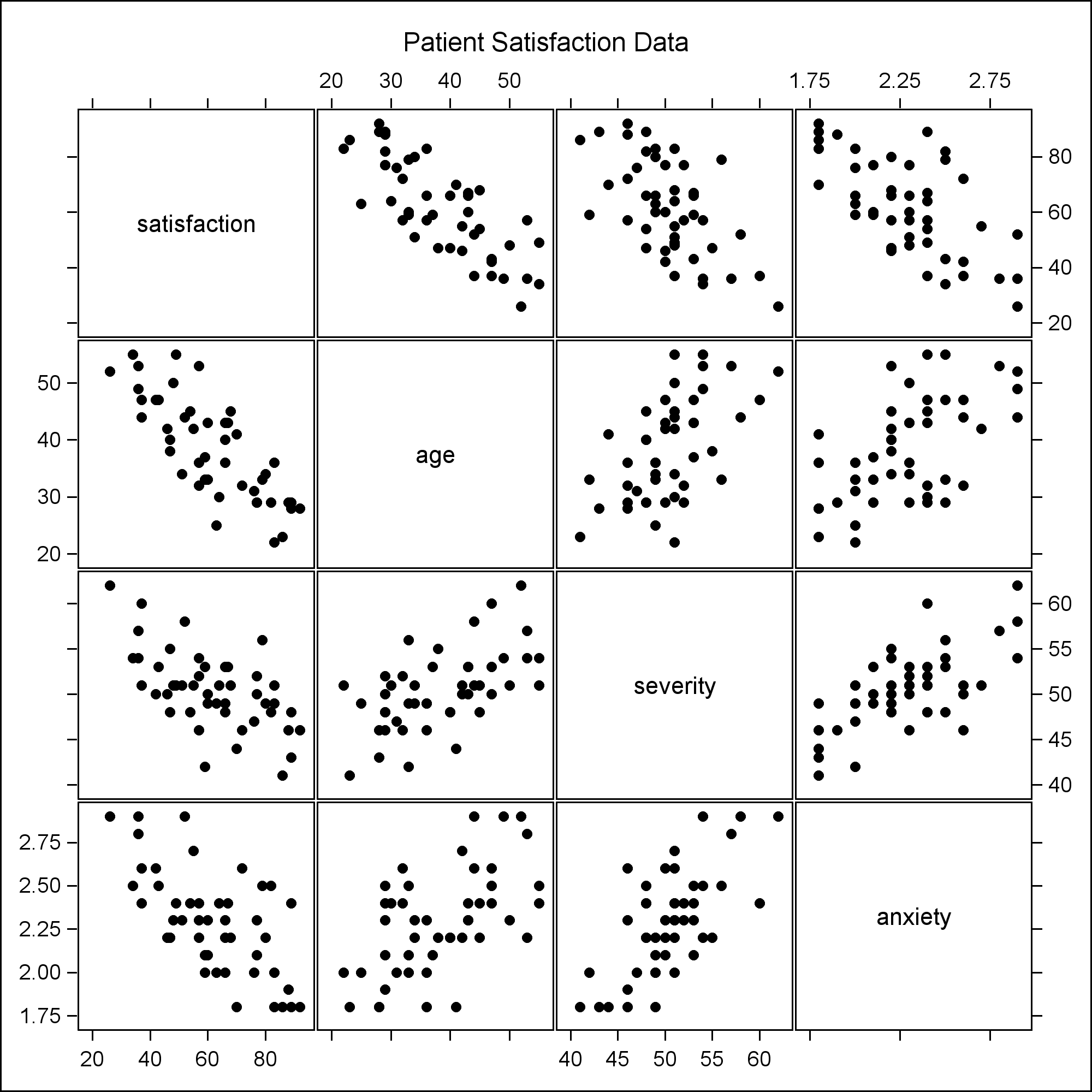
**proc sgscatter data=patient;**

**matrix satisfaction age severity anxiety /**

**markerattrs=(symbol=CIRCLEFILLED size=2pt);**

**title1 'Patient Satisfaction Data';**

**run;**



**/\* Fit regression model \*/**

**proc reg data=patient;**

**model satisfaction = age severity anxiety;**

**output out=out1 r=resid p=pred;**

**title1 'Patient Satisfaction Regression';**

**run;**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | ***Patient Satisfaction Regression*** |      |  |  | | --- | --- | | **Number of Observations Used** | 46 |      | **Analysis of Variance** | | | | | | | --- | --- | --- | --- | --- | --- | | **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** | | **Model** | 3 | 9120.46367 | 3040.15456 | 30.05 | <.0001 | | **Error** | 42 | 4248.84068 | 101.16287 |  |  | | **Corrected Total** | 45 | 13369 |  |  |  |      |  |  |  |  | | --- | --- | --- | --- | | **Root MSE** | 10.05798 | **R-Square** | 0.6822 | | **Dependent Mean** | 61.56522 | **Adj R-Sq** | 0.6595 | | **Coeff Var** | 16.33711 |  |  |      | **Parameter Estimates** | | | | | | | --- | --- | --- | --- | --- | --- | | **Variable** | **DF** | **Parameter Estimate** | **Standard Error** | **t Value** | **Pr > |t|** | | **Intercept** | **1** | 158.49125 | 18.12589 | 8.74 | <.0001 | | **age** | **1** | -1.14161 | 0.21480 | -5.31 | <.0001 | | **severity** | **1** | -0.44200 | 0.49197 | -0.90 | 0.3741 | | **anxiety** | **1** | -13.47016 | 7.09966 | -1.90 | 0.0647 | |

**/\* Check model assumptions \*/**

**%macro resid\_num\_diag(dataset,datavar, ...**

**%*resid\_num\_diag*(dataset=out1, datavar=resid,**

**label='Residual', predvar=pred,**

**predlabel='Predicted Value');**

**run;**

**/\* Ouput not included here;**

**This give BF\_pvalue = .81453 and correlation .98851**

**(N=46; check text Table B.6 for threshold)**

**\*/**

**/\* Joint 90% intervals for beta1, beta2, and beta3 \*/**

**proc reg data=patient;**

**model satisfaction = age severity anxiety /**

**clb alpha=.0333;**

**title1 'Simultaneous 90% intervals for three predictors effects';**

**run;**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Simultaneous 90% intervals for three predictors effects***     | **Parameter Estimates** | | | | | | | | | --- | --- | --- | --- | --- | --- | --- | --- | | **Variable** | **DF** | **Parameter Estimate** | **Standard Error** | **t Value** | **Pr > |t|** | **96.67% Confidence Limits** | | | **Intercept** | **1** | 158.49125 | 18.12589 | 8.74 | <.0001 | 118.59967 | 198.38283 | | **age** | **1** | -1.14161 | 0.21480 | -5.31 | <.0001 | -1.61434 | -0.66888 | | **severity** | **1** | -0.44200 | 0.49197 | -0.90 | 0.3741 | -1.52473 | 0.64072 | | **anxiety** | **1** | -13.47016 | 7.09966 | -1.90 | 0.0647 | -29.09514 | 2.15482 | |

**/\* Simultaneous 90% prediction limits on two new patients**

**(using Scheffe and Bonferroni), with profiles**

**age=35, severity=45, anxiety=2.2**

**and**

**age=42, severity=61, anxiety=1.8**

**\*/**

**data dummy; input age severity anxiety check; cards;**

**35 45 2.2 1**

**42 61 1.8 1**

**;**

**data temp; set patient dummy;**

**proc reg data=temp noprint;**

**model satisfaction = age severity anxiety;**

**output out=out1 p=Yhat stdi=seYhatnew;**

**/\* KEY: stdi is SE of individual prediction \*/**

**data out1; set out1;**

**alpha = 0.10; /\* 1-alpha is simult. pred. level \*/**

**p = 4; /\* # of beta's (including intercept) \*/**

**n = 46; /\* sample size \*/**

**g = 2; /\* number of simultaneous intervals \*/**

**S = sqrt(g\*finv(1-alpha,g,n-p)); /\* Scheffe crit val \*/**

**t = tinv(1-alpha/(2\*g),n-p); /\* Bonf. crit. val. \*/**

**S\_upper = Yhat + S\*seYhatnew;**

**S\_lower = Yhat - S\*seYhatnew;**

**B\_upper = Yhat + t\*seYhatnew;**

**B\_lower = Yhat - t\*seYhatnew;**

**proc print data=out1;**

**where check = 1;**

**var age severity anxiety Yhat S\_lower S\_upper**

**B\_lower B\_upper;**

**title1 'Simultaneous 90% intervals of individual prediction';**

**title2 'at two X-profiles, using Scheffe and Bonferroni';**

**run;**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | ***Simultaneous 90% intervals of individual prediction*** | | ***at two X-profiles, using Scheffe and Bonferroni*** |  | **Obs** | **age** | **severity** | **anxiety** | **Yhat** | **S\_lower** | **S\_upper** | **B\_lower** | **B\_upper** | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | **47** | 35 | 45 | 2.2 | 69.0103 | 46.0553 | 91.9652 | 48.0122 | 90.0083 | | **48** | 42 | 61 | 1.8 | 59.3350 | 31.3797 | 87.2903 | 33.7629 | 84.9071 | |

**/\* What if a transformation were needed? \*/**

**proc transreg data=patient;**

**model boxcox(satisfaction / lambda=-.6 to 1.8 by 0.2)**

**= identity(age severity anxiety);**

**title1 'Box-Cox Transformation with Multiple Predictors';**

**run;**

|  |  |
| --- | --- |
| |  | | --- | | ***Box-Cox Transformation with Multiple Predictors*** | |