BIST P8130: Biostatistics Methods I

Recitation 08 Simple Linear Regression in SAS

Zilan Chai

Fall 2017

This recitation's big ideas:

- Use PROC CORR to calculate the correlation coefficient in SAS
- Practice problem (2.4, 2.13) in SAS

Example: Refer to Grade point average dataset (Kutner Problem 1.19)

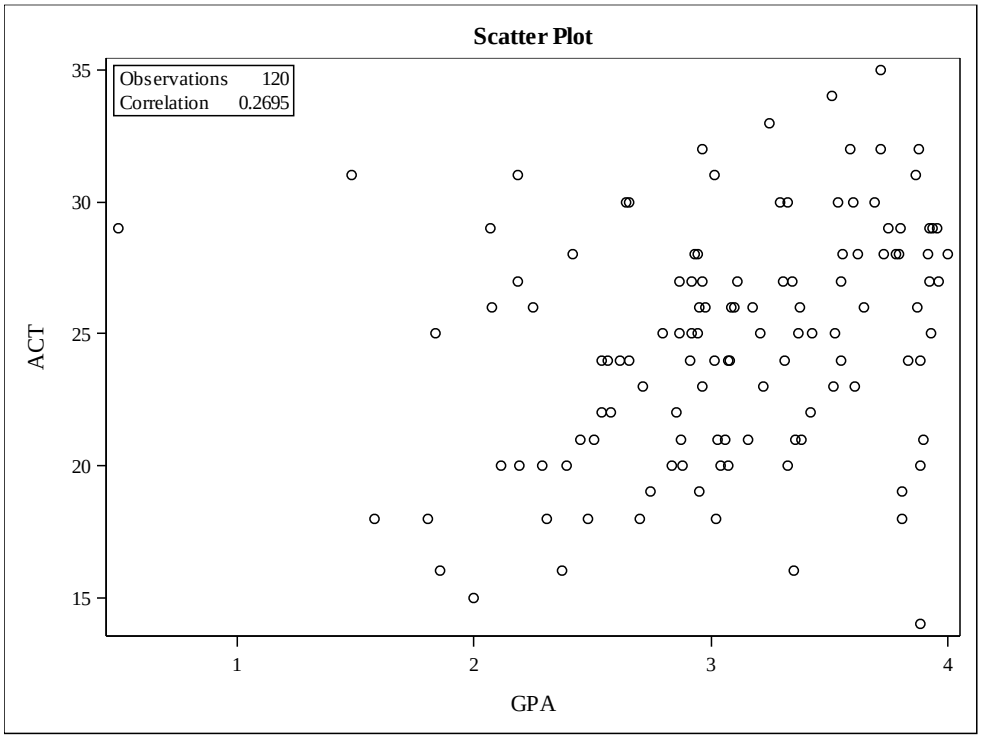
Obs	GPA	ACT
1	3.897	21
2	3.885	14
…	…	…
118	3.914	28
119	1.86	16
120	2.948	28

SAS code for calculating correlation coefficient:

```
PROC CORR data = GPA plots = scatter (ellipse= none);
  var GPA ACT;
RUN;
```

Simple Statistics						
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
GPA	120	3.07405	0.64434	368.88600	0.50000	4.00000
ACT	120	24.72500	4.47207	2967	14.00000	35.00000

Pearson Correlation Coefficients, N = 120 Prob > |r| under H0: Rho=0		
	GPA	ACT
GPA	1.00000	0.26948 0.0029
ACT	0.26948 0.0029	1.00000



Practice problem 2.4 in SAS:

Note: This is only a brief introduction of the calculation in SAS. For all the hand calculation and interpretation, please refer to the Solution of Practice Problem.

Kutner 2.4 a)

```
proc reg data=GPA;
model GPA = ACT /CLB ALPHA = 0.01;
run;
```

Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	99% Confidence Limits	
Intercept	1	2.11405	0.32089	6.59	<.0001	1.27390	2.95420
ACT	1	0.03883	0.01277	3.04	0.0029	0.00539	0.07227

Kutner 2.4 b&c)

```
PROC CORR data = GPA  Fisher(alpha = 0.01);
  var GPA ACT;
RUN;
```

Pearson Correlation Coefficients, N = 120 Prob > |r| under H0: Rho=0		
	GPA	ACT
GPA	1.00000	0.26948 0.0029
ACT	0.26948 0.0029	1.00000

Pearson Correlation Statistics (Fisher's z Transformation)									
Variable	With Variable	N	Sample Correlation	Fisher's z	Bias Adjustment	Correlation Estimate	99% Confidence Limits		p Value for H0:Rho=0
GPA	ACT	120	0.26948	0.27630	0.00113	0.26843	0.037020	0.472519	0.0028

Kutner 2.13 (a. b)

```
*Kutner 2.13 a&b;
Data Xvalue;
INPUT ACT GPA;
CARDS;
28 .
;
run;

Data GPA;
SET GPA Xvalue;
run;

proc reg data=GPA;
model GPA = ACT /CLM CLI ALPHA = 0.05;
run;

/* The option CLM will give us CIs for the expected value of Y */
/* The option CLI will give us CIs for the individual predicted value of Y */
```

Output Statistics								
Obs	Dependent Variable	Predicted Value	Std Error Mean Predict	95% CL Mean		95% CL Predict		Residual
1	3.8970	2.9294	0.0742	2.7826	3.0763	1.6868	4.1721	0.9676
2	3.8850	2.6576	0.1483	2.3639	2.9514	1.3892	3.9261	1.2274
...								
121	.	3.2012	0.0706	3.0614	3.3410	1.9594	4.4431	.