

## Homework # 13

due Friday, 2 December

The index of biotic integrity (IBI) is a measure of the water quality in streams. IBI and land-use measures for a collection of streams in the Ozark Highland ecoregion of Arkansas were collected as part of a study. The IBI data set (next page) gives data for the 49 streams in the original sample with area less than or equal to  $70 \text{ km}^2$ . The variables are the area of the watershed (in square kilometers), the IBI, and the percent of the area that was forested.

For the following, treat IBI as the response variable.

- (1) Use numerical and graphical methods to describe the variable IBI. Do the same for area. Summarize your results.
- (2) Plot the data and describe the relationship between IBI and area. Are there any outliers or unusual patterns?
- (3)
  - (a) Give the statistical model for simple linear regression for this problem.
  - (b) State the null and alternative hypotheses for examining the relationship between IBI and area.
  - (c) Run the simple linear regression and summarize the results.
- (4)
  - (a) Obtain the residuals and plot them versus area. Is there anything unusual in the plot?
  - (b) Do the residuals appear to be approximately normal? Give reasons for your answer.
- (5) Do the assumptions for the analysis of these data using the model you gave in (3) appear to be reasonable? Explain your answer.
- (6)
  - (a) Find a 95% confidence interval for mean response corresponding to an area of  $30 \text{ km}^2$ .
  - (b) Find a 95% prediction interval for a future response in an area of  $30 \text{ km}^2$ .
  - (c) Write a short paragraph interpreting the meaning of these intervals in terms of Ozark Highland streams.
  - (d) Do you think that these results can be applied to other streams in Arkansas or in other states? Explain why or why not.
- (7)
  - (a) Find the correlations among all three variables, perform the test of statistical significance for each correlation, and summarize the results.
  - (b) Which of these test results could have been obtained from the analysis that you performed in question (3)?

Area	IBI	Forest
21	47	0
29	61	0
31	39	0
32	59	0
34	72	0
34	76	0
49	85	3
52	89	3
2	74	7
70	89	8
6	33	9
28	46	10
21	32	10
59	80	11
69	80	14
47	78	17
8	53	17
8	43	18
58	88	21
54	84	22
10	62	25
57	55	31
18	29	32
19	29	33
39	54	33

Area	IBI	Forest
49	78	33
9	71	39
5	55	41
14	58	43
9	71	43
23	33	47
31	59	49
18	81	49
16	71	52
21	75	52
32	64	59
10	41	63
26	82	68
9	60	75
54	84	79
12	83	79
21	82	80
27	82	86
23	86	89
26	79	90
16	67	95
26	56	95
26	85	100
28	91	100