

**Part I.** (50 points) Do all calculations in SAS. Use a word processor of your choice to write a report. Insert computer text output and graphics to support what you are saying, but you need to write something that looks like an academic paper — not a pile of computer output. Turn in a hard copy of your HW in class (i.e., don't email me your HW). Also:

1. Clearly specify parameters and hypotheses when appropriate.
2. Write a coherent conclusion.

(50<sup>pts</sup>) **1. Amazon forest clearing and butterflies:**

As an introduction to their study on the effect of Amazon forest clearing<sup>1</sup>, the researchers stated “Fragmentation of once continuous wild areas is a major way in which people are altering the landscape and biology of the planet.” Their study takes advantage of a Brazilian requirement that 50% of the land in any development project remain in forest and tree cover. As a consequence of this requirement, “islands” of forest of various sizes remain in otherwise cleared areas. The data below are the number of butterfly species found in 16 such islands. The area of the island is in hectares.

Summarize the role of area in the distribution of number of butterfly species. Write a brief report including a summary of statistical findings, graphical displays, and a section dealing with the methods used to answer the question of interest.

Reserve	Area (hectares)	Species (types)
1	1	14
2	1	50
3	1	55
4	1	34
5	1	40
6	1	57
7	10	43
8	10	103
9	10	33
10	10	53
11	10	50
12	100	110
13	100	70
14	100	119
15	100	60
16	1000	145

Create a SAS program, and provide the program as part of your writeup, to read in the data, and carry out the subsequent parts of the HW. It would be preferable for you create a text file with the data in it and use an INFILE statement to read the data in from that file. Include relevant comments as part of the SAS program; see the class notes for how to do this.

**Remark:** This problem is somewhat open-ended, but is easily formulated as a regression problem, where you are interested in whether the number of species changes with area, and if so, then how. You might consider the following steps in an analysis: (1) plot the data, (2) consider transforming the data to a scale where the relationship is roughly linear, and (3) if transformation is suggested, do the statistical analysis on the transformed scale, else do the analysis on the given data. Finally, (4) check model assumptions and (5) summarize your findings.

- (a) (10 pts) Plot the data, consider transforming the data and plot on transformed scale.
- (b) (20 pts) Do a statistical analysis on the scale (transformed or original) where the relationship is close to linear.
- (c) (10 pts) Check model assumptions.
- (d) (10 pts) Summarize your findings.

<sup>1</sup>Lovejoy, T.E., Rankin, J.M., Bierregaard, R.O., Brown, K.S., Emmons, L.H., and Van der Woot, M.E. (1984). “Ecosystem Decay of Amazon Forest Remnants” in *Extinctions*; edited by M. H. Nitecki. Chicago: University of Chicago Press. Example from *The Statistical Sleuth* by Ramsey and Schafer.

50 pts

50 pts