NAME:

UNI:

Homework 8

P8120 Spring 2017

Due: **Monday, May 1st, 2017 at 5:30pm**.

* All homework MUST BE submitted electronically via CourseWorks in PDF format!
* Use file name: LastName\_FirstName\_UNI\_HW8
* HW has to be TYPED, with your name and **UNI** written on top.
* Problems must be ordered, starting with Problem #1.
* Any question must be answered with a full and meaningful sentence using the words of the problem!
* Don’t forget to **SHOW KEY STEPS OF YOUR WORK**! You get partial credit for showing meaningful steps. A final answer is not enough for full credit.
* Any hypothesis test must follow the steps outlined in class.
* You can work in groups or discuss the problems with your classmates. However, your **written answers and solutions must be strictly your own and cannot be an exact copy of your classmates answers!**
* **USING PAST YEARS’ SOLUTIONS IS NOT ACCEPTABLE!**

When computing and reporting results using SAS:

* Even when using SAS, specify and check all assumptions (if any) for both confidence intervals and hypothesis tests, write down the steps of all hypothesis tests.
* For each question, only provide relevant parts of SAS output properly annotated. The relevant parts (tables) of SAS output have to be incorporated into your answers (cut and paste) in the correct order.
* We don’t need to see your SAS code, however if you prefer to have it with your homework, staple it at the end.

1. A retrospective study was conducted to investigate the effects of aluminum on the development of Alzheimer’s disease. Researchers compared a group of 177 Alzheimer’s patients with a control group of 158 subjects who did not have the disease. Each subject was classified according to the use of antacids that contain aluminum. The data are provided in the table below. Use SAS to answer the following questions.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Antacid Use | | | |
| None | Low | Medium | High |
| Diseased | 109 | 13 | 23 | 32 |
| Not Diseased | 123 | 12 | 13 | 10 |

1. Perform an overall test for association between antacid use and disease status ignoring the ordinal nature of the exposure variable.
2. Now using logistic regression techniques, assess the research question in part a).
3. Assign the scores 0, 1, 2, and 3 to the exposure levels and perform a test to determine if antacid use is linearly associated with Alzheimer’s disease.
4. Now using logistic regression techniques, assess the research question in part c).
5. Complete the table below with information found in parts (a) - (d).

|  |  |  |  |
| --- | --- | --- | --- |
| Test | Test Statistic  Value | Null  Distribution (df) | p-value |
| Part A |  |  |  |
| Part B |  |  |  |
| Part C |  |  |  |
| Part D |  |  |  |

What observations do you make based on this table? Which analysis is most appropriate here?

1. In a short paragraph, explain your findings from the appropriate analysis. Be sure to indicate the direction of association if there is an association.
2. Data in the file “breast” consist of a subset of data from a larger study on benign breast disease. Cases consist of women with a biopsy-confirmed diagnosis of fibrocystic breast disease from two hospitals in New Haven, CT. Controls were selected from among female patients admitted for other services at the same two hospitals. Cases and controls were matched on age at interview by the study investigators. The variables in the data set are:

STR (stratum 1 - 50)

AGMT (age at interview, years)

FNDX (final diagnosis 1 = Case, 0 = Control)

CHK (regular medical checkups 1 = Yes, 0 = No)

AGMN (age at menarche, years)

WT (weight of subject at interview, pounds)

MST (marital status 1 = married, 2 = divorced, 3 = separated, 4 = widowed, 5 = never married)

1. What type of study design is this?
2. What type of analysis is appropriate for this study?
3. First create a new binary variable called NVMR that takes on the value of 1 if a subject was never married and 0 if the subject was ever married. Then fit a model that includes CHK, AGMN, WT, and NVMR. Write down the model and interpret the estimated model parameters (slopes).
4. In a brief summary, explain your findings for each of the explanatory variables from the fitted model in (c). Be sure to assess the significance of any associations, interpreting those associations that are significant (*no need to go through all seven steps*).
5. A podiatrist has designed a new shoe insert that he hopes will correct for overpronation. However, he is concerned that the material that he is using for the insert will induce heel tenderness as a result of some loss of cushioning on the strike of each step. A study was conducted on 87 subjects who participate in high impact exercise and who used the insert for a month. Researchers asked the participants whether or not they experienced occasional heel tenderness before and after they used the new insert in their shoes.

|  |  |  |  |
| --- | --- | --- | --- |
|  | | After | |
|  | Heel Tenderness | Yes | No |
| Before | Yes | 19 | 5 |
| No | 15 | 48 |

1. What are the matched pairs in this study?
2. What proportion of subjects had heel tenderness before using inserts? What proportion of subjects had heel tenderness after one month of using inserts?
3. Perform an appropriate analysis to determine if there is an association between using the insert and heel tenderness. Confirm your answer using SAS.
4. Compute by hand the estimated OR for heel tenderness comparing after insert use to before insert use. Also provide the 95% confidence interval for the true OR. Interpret both the point and interval estimate.
5. Use SAS to confirm your answers in (d). The individual level data are given below:

**data** insert;

input subject time $ response $ @@;

datalines;

1 before no 1 after no 2 before no 2 after no

3 before no 3 after no 4 before no 4 after no

5 before no 5 after no 6 before no 6 after no

7 before no 7 after no 8 before no 8 after no

9 before no 9 after no 10 before no 10 after no

11 before no 11 after no 12 before no 12 after no

13 before no 13 after no 14 before no 14 after no

15 before no 15 after no 16 before no 16 after no

17 before no 17 after no 18 before no 18 after no

19 before no 19 after no 20 before no 20 after no

21 before no 21 after no 22 before no 22 after no

23 before no 23 after no 24 before no 24 after no

25 before no 25 after no 26 before no 26 after no

27 before no 27 after no 28 before no 28 after no

29 before no 29 after no 30 before no 30 after no

31 before no 31 after no 32 before no 32 after no

33 before no 33 after no 34 before no 34 after no

35 before no 35 after no 36 before no 36 after no

37 before no 37 after no 38 before no 38 after no

39 before no 39 after no 40 before no 40 after no

41 before no 41 after no 42 before no 42 after no

43 before no 43 after no 44 before no 44 after no

45 before no 45 after no 46 before no 46 after no

47 before no 47 after no 48 before no 48 after no

49 before no 49 after yes 50 before no 50 after yes

51 before no 51 after yes 52 before no 52 after yes

53 before no 53 after yes 54 before no 54 after yes

55 before no 55 after yes 56 before no 56 after yes

57 before no 57 after yes 58 before no 58 after yes

59 before no 59 after yes 60 before no 60 after yes

61 before no 61 after yes 62 before no 62 after yes

63 before no 63 after yes 64 before yes 64 after no

65 before yes 65 after no 66 before yes 66 after no

67 before yes 67 after no 68 before yes 68 after no

69 before yes 69 after yes 70 before yes 70 after yes

71 before yes 71 after yes 72 before yes 72 after yes

73 before yes 73 after yes 74 before yes 74 after yes

75 before yes 75 after yes 76 before yes 76 after yes

77 before yes 77 after yes 78 before yes 78 after yes

79 before yes 79 after yes 80 before yes 80 after yes

81 before yes 81 after yes 82 before yes 82 after yes

83 before yes 83 after yes 84 before yes 84 after yes

85 before yes 85 after yes 86 before yes 86 after yes

87 before yes 87 after yes

;

**run**;