MSDS 6372: UNIT 11 HW

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **non-smokers** | | |  | **smokers** | | |
| **mortality** | **exposure** | |  | **mortality** | **exposure** | |
| **yes** | **no** |  | **yes** | **no** |
| **deaths** | 17 | 6 |  | **deaths** | 131 | 69 |
| **referents** | 98 | 103 |  | **referents** | 274 | 240 |

The table above references a case-referent (case-control) study of lung cancer deaths (Deaths) and non-deaths (Referents) and their relationship to asbestos exposure (yes,no) and the worker’s smoking status. The data for analysis with SAS is in hw11Data.xlsx. The study was conducted by Liddell in 2001 from previous study Liddell had performed in 1984. A link to the 2001 paper is:

<https://pdfs.semanticscholar.org/db24/28cdad85c64e3ec9643aa658805875af679e.pdf>

Let’s assume the data were collected by sampling 223 workers who died from lung cancer and 715 that did not die from lung cancer. These workers were then classified as either exposed to asbestos dust or not as well as classified as a smoker or non-smoker. Note that the response is listed in the rows in the above table and the explanatory variable is listed in the columns. When you run proc freq in SAS make sure the explanatory variable is listed in the row and the response in the columns. Also, sort the input data with the code below so the analysis by all students is consistent:

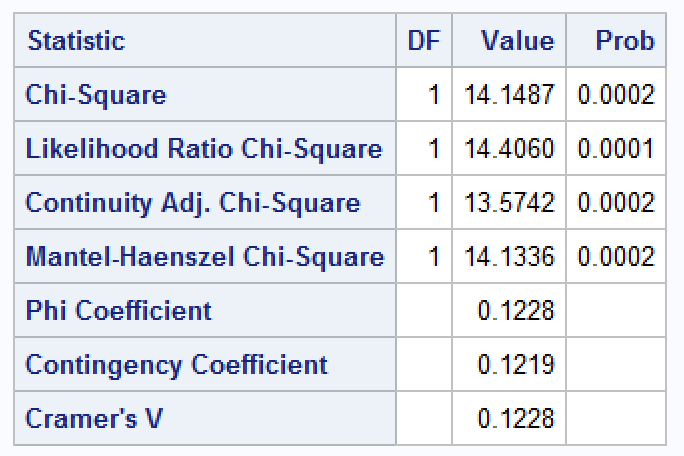
**proc** **sort** data=datIn;

by descending exposure descending response;

**run**;

1. Combine the smoking and non-smoking data to create one table of data with the responses (Death / Referents (non-deaths) on the columns and the explanatory variable (Exposure) on the rows. Be sure and clearly label your rows and columns. Simply cut and paste your table below.

|  |  |  |
| --- | --- | --- |
| **Exposure** | **Mortality** | |
| **referents** | **deaths** |
| **no exposure** | 343 | 75 |
| **exposure** | 372 | 148 |

1. Calculate a chi-square statistic to evaluate if exposure affects deaths. Do your calculations manually and confirm them with SAS. What do you conclude from the results?

|  |  |  |
| --- | --- | --- |
| **Exposure** | **Mortality** | |
| **referents** | **deaths** |
| **no exposure** | 343 | 75 |
| **exposure** | 372 | 148 |
|  | 715 | 223 |
| chi |  | 14.14870628 |

The results suggest there is sufficient evidence that there is frequencies across the exposure are not equal.

1. Use the table from part A to test if the proportion (odds, if proportions are not appropriate) of dying from lung cancer is greater for those that are exposed to asbestos than those that are not. Provide the null and alternative hypothesis of your test as well as any p-values and conclusion. Please include a 95% confidence interval in your conclusion. Do you calculations manually and check them with SAS.

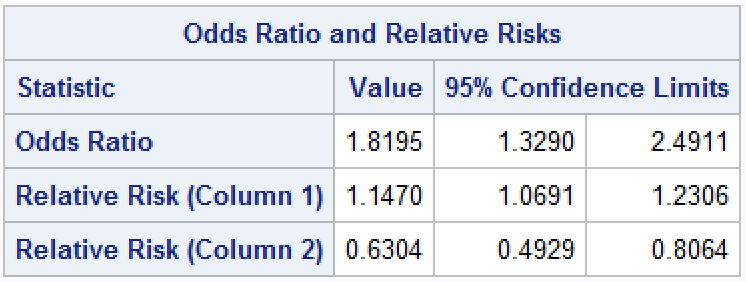
1.863

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Exposure** | **Mortality** | |  |  |  |  |
| **referents** | **deaths** |  |  |  | 8.1469E+221 |
| **no exposure** | 346 | 72 | 418 |  | 1.30674E+82 |  |
| **exposure** | 369 | 151 | 520 |  | 4.4579E+134 |  |
|  | 715 | 223 | 938 |  |  |  |
|  |  |  |  |  | p = | 7.15042E-06 |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  | 343 | 75 |  |
|  |  |  |  | 372 | 148 | 4.78967E-05 |
|  |  |  |  | 344 | 74 |  |
|  |  |  |  | 371 | 149 | 2.60714E-05 |
|  |  |  |  | 345 | 73 |  |
|  |  |  |  | 370 | 150 | 1.38312E-05 |
|  |  |  |  | 346 | 72 |  |
|  |  |  |  | 369 | 151 | 7.15042E-06 |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  | 9.49498E-05 |

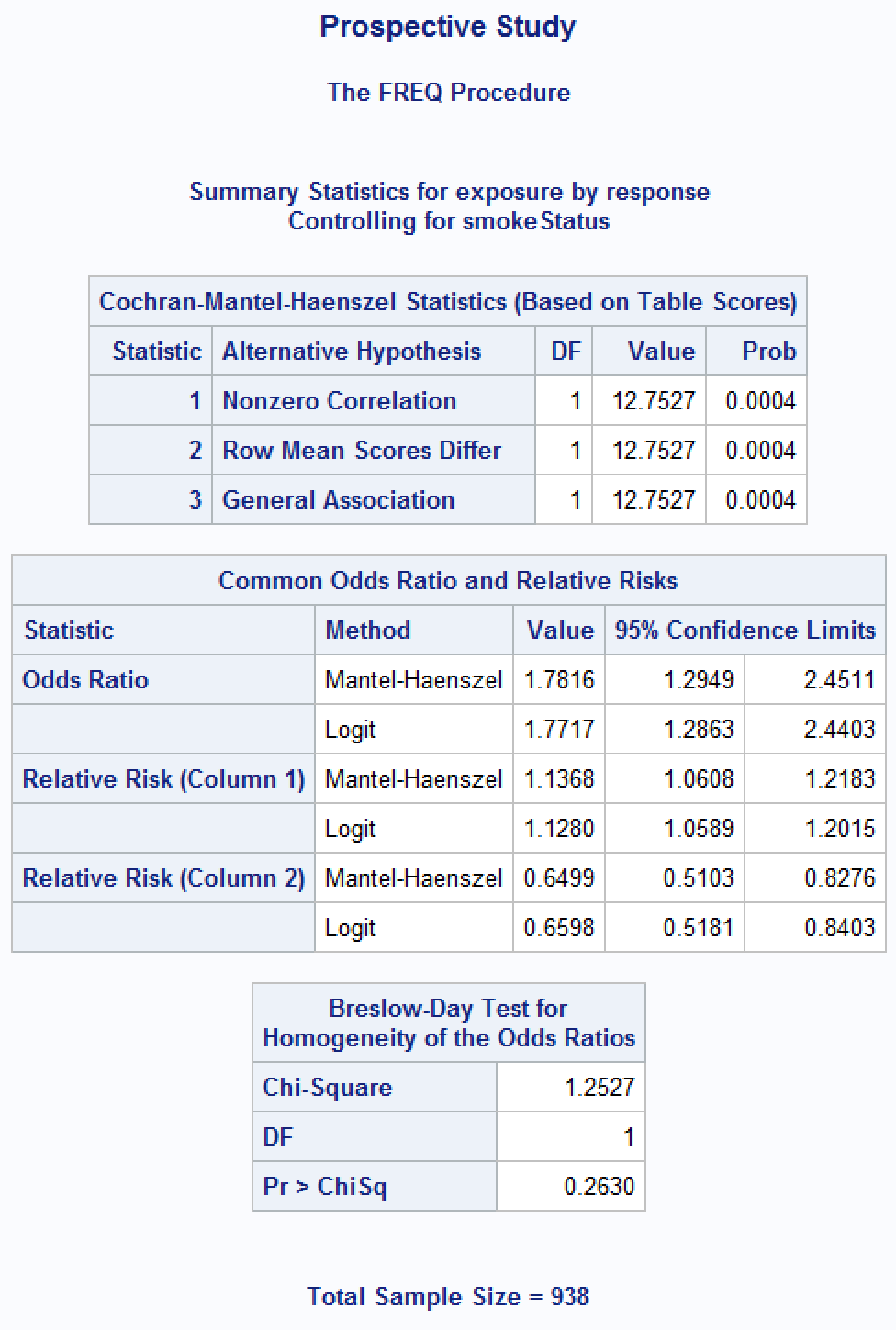
Null hypothesis: odds ratio =1

Alternative hypothesis: odds ratio not equal to 1

The odds of dying from lung caster are 1.85 times greater (approximately 95% CI: ) when exposed to asbestos than when not exposed (p value < 0.0001).



1. Using SAS conduct a Mantel-Haenszel test to test if the proportion of dying of lung cancer is greater for those exposed to asbestos than those that are not after accounting for smoking. Be sure and first check the assumption of the test with the Breslow-Day test. Provide an analysis of your results.



Based on the results from the Breslow-Day test, we would fail to reject the null hypothesis. In other words, there is insufficient evidence to suggest the odds ratio between mortality and exposure is equivalent for smokers/non-smokers.

Having met the assumptions, we proceed to the Mantel-Haenszel test. The results from the Mantel-Haenszel test reject the null hypothesis that the odds ratio is equal to 1. We would need to look further to see if smoke status is affecting the results.