Homework III (2022)

1. The following table summarizes eight studies in China about smoking and lung cancer.

   (a) Fit a logistic model with smoking and city as predictors. Interpret the smoking effect.

   (b) Conduct a Pearson test of goodness of fit. Interpret.

   (c) Check residuals to analyze further the quality of fit. Interpret.

| | | Lung Cancer | | | | Lung Cancer | |
|---|---|---|---|---|---|---|---|
| City | Smoking | Yes | No | City | Smoking | Yes | No |
| Beijing | Yes | 126 | 100 | Harbin | Yes | 402 | 308 |
| | No | 35 | 61 | | No | 121 | 215 |
| Shanghai | Yes | 908 | 688 | Zhengzhou | Yes | 182 | 156 |
| | No | 497 | 807 | | No | 72 | 98 |
| Shenyang | Yes | 913 | 747 | Taiyuan | Yes | 60 | 99 |
| | No | 336 | 598 | | No | 11 | 43 |
| Nanjing | Yes | 235 | 172 | Nanchang | Yes | 104 | 89 |
| | No | 58 | 121 | | No | 21 | 36 |

Source: Based on data in Z. Liu, *Int. J. Epidemiol.*, **21**: 197-201.

2. For the horseshoe crab data (available in icampus), fit a model using weight and width as predictors. Conduct (**a**) a likelihood-ratio test of $H_0 : \beta_1 = \beta_2 = 0$, and (**b**) separate tests for the partial effects. Why does neither test in part (b) show evidence of an effect when the test in part (a) shows strong evidence?