If R is used, present the required output and the (relevant) syntax.

1. The dataset `melanoma` in the library `faraway` gives data on a sample of patients suffering from melanoma (skin cancer) cross-classified by the type of cancer and the location on the body.[1]

   a) Display the data in a contingency table and obtain a mosaic plot to visually check if `tumor` and `site` are independent. In addition, use a poisson model to determine if `tumor` and `site` are independent. Does your test agree with your visual conclusions?

   b) Make a two-way table of the deviance residuals from your model in a). Are there any larger residuals? Comment.

2. A student newspaper conducted a survey of student opinions about the Vietnam War in May 1967. Responses were classified by sex, year in the program and one of four opinions. The survey was voluntary. The data may be found in the dataset `uncviet` in the library `faraway` .

   a) Conduct an analysis of the patterns of dependence in the data assuming that all variables are nominal.

   b) Assign scores to the year and opinion and fit an appropriate model. Interpret the trends in opinion over the years. Check the sensitivity of your conclusions to the assignment of the scores (by trying other sensible alternatives).

3. For the multinomial logit model, the link function is given by $\eta_{ij} = log(p_{ij}/p_{i1})$ (Check ELM[2] Chapter 7 for more details). Show that

$$p_{ij} = \frac{e^{\eta_{ij}}}{1 + \sum_{k=2}^{J} e^{\eta_{ik}}}$$

   Hint: Since $\eta_{ij} = log(p_{ij}/p_{i1})$ then $p_{ij} = p_{i1}e^{\eta_{ij}}$ and $\sum_{i=1}^{J} p_{ij} = \sum_{i=1}^{J}(p_{i1}e^{\eta_{ij}})$.

4. The `hsb` data from the library `faraway` was collected as a subset of the High School and Beyond study conducted by the National Education Longitudinal Studies program of the National Center for Education Statistics. The variables are gender; race; socioeconomic status (SES); school type; chosen high school program type; scores on reading, writing, math, science, and social studies. We want to determine which factors are related to the choice of the type of program academic, vocational or general that the students pursue in high school. The response is multinomial with three levels.

   a) Make a table showing the proportion of males and females choosing the three different programs. Comment on the difference. Repeat this comparison but for SES rather than gender.

   b) Construct a plot like the right panel of Figure 7.1 in ELM that shows the relationship between program choice and reading score. Comment on the plot.

   c) Fit a multinomial response model for the program choice and examine the fitted coefficients. Interpret at least two coefficients. In addition, observe that of the five subjects, one gives unexpected coefficients. Why do you think this happens?

---

[1] Modified from ELM Ch7 Question 2

[2] Acronym for **E**xtending the **L**inear **M**odel in R by Faraway