Categorical Data Analysis: Midterm 1, Take Home

Instructions: This part of Midterm 1 is due next week Friday (March 29th) in class. Please show all relevant R commands and type up your answers neatly. You may use all your class material for this midterm (lecture notes, all material that is available on our course website) or any books that you want to consult, but you *may not show to or discuss this exam with anyone*! If I should get aware of any cheating, the student(s) involved will fail the class. This exam needs to be your own work exclusively.

1. In the general population, the proportion of males and females that are left-handed is about the same. A biologist suspects that the student population in Graz is not like the general population in terms of handedness. He randomly selects 100 male students and finds that 22 are left-handed. He also randomly selects 100 female students and finds that 15 are left-handed. Do these data support the biologist's theory? Answer by computing a confidence interval for an appropriate parameter of interest. (10 points)

2. At the R prompt, type

    ```
    xtabs(Freq ~ Class+Survived, Titanic)
    ```

   to obtain a table that shows how many passengers in first, second and third class and the crew survived the fateful journey of the Titanic. (Titanic is a built in dataset in R.) Test if there is an association between class membership and surviving. If there is a relationship, try to explain where it is mostly coming from using (standardized) residuals. (15 points)

3. The following table shows the results of a retrospective study comparing radiation therapy with surgery in treating cancer of the larynx. The response indicates whether the cancer was controlled for at least two years following treatment. (10 points)

	Cancer	
	controlled	not controlled
Surgery	21	2
Radiation Therapy	15	3

   Report and interpret the P-value for Fisher's exact test with (a) $H_A : \theta > 1$ and (b) $H_A : \theta \neq 1$.