Topics to be covered this week

Statistics 138
Fall Quarter, 2018

Monday, Dec 3 Loglinear models (Handout 14, and chap 7 from the text).

Wednesday, Dec 5 Loglinear models (handouts 14 and 15, chap 7 from the text).

Friday, Dec 5 Comments on loglinear models.

Homework 7 (Due: Friday, Dec 5)
You may form a group of three registered in this course and submit one completed homework for the group. The front page should be blank. Please write down the names of the students in the group on the submitted work.

1. For a log-linear model with $I = J = K = 3$, answer the following questions. Assume the "set to zero" constraints instead of the "summation" constraints.
(a) How many non-zero terms are there for $\{\lambda_i^X\}$, $\{\lambda_j^Y\}$ and $\{\lambda_k^Z\}$?
(b) How many non-zero terms are there for $\{\lambda_{ij}^{XY}\}, \{\lambda_{ik}^{XZ}\}$ and $\{\lambda_{jk}^{YZ}\}$?
(c) How many non-zero terms are there for $\{\lambda_{ijk}^{XYZ}\}$?
(d) How many non-zero terms are there for the saturated model, total?

2. The dataset HD.csv has three columns: weight (Over=overweight, NotOver=not overweight), heart (Yes=heart disease present, No=heart disease absent), and Freq (the frequency in each category). Consider heart to be the Y variable and weight to be the X variable.
(a) Fit the loglinear model of independence, and obtain the estimates of the λ-parameters.
(b) Obtain the estimated expected frequencies and check that they correspond to the estimates $n_{i+}n_{+j}/n$.
(c) Obtain the goodness of fit statistic G^2, its p-value and the AIC value.
(d) Write down the null and the alternative hypotheses corresponding to the p-value found in part (c). State your conclusion if level $\alpha = 0.05$.

3. A medical study wanted to see if an additictive narcotic drug being "cured" (no relapse in 2 years), differed by treatment and by gender. The information given per subject is:
(i) Y : If the the patient was cured (Yes) or not (No).
(ii) X : If the patient was male (M) or female (F).
(iii) Z: What treatment group they were in - One, Two, Three.
The data can be found in the file bepatient.csv.
(a) Fit the model for conditional independence of cured and gender given treatment group. Obtain the goodness of fit statistic G^2 and its p-value. Can you conclude at a level $\alpha = 0.05$, if this model for conditional independence is reasonable for this data?
(b) Starting with the saturated model, run a backward stepwise regression to obtain the appropriate model. For this final model, write down the parameter estimates, their standard errors, the z-values and the p-values. [Fit the saturated model in R, and then use the R function "step"].

(c) For the final model selected in part (b), obtain the the conditional (conditioning over treatment) odds ratio of being cured for females vs males for each of the treatment groups.