STA 312f12 Assignment Ten[1]

Please bring your R printouts to the quiz. The non-computer questions are practice for the quiz on Friday Nov. 23d, and are not to be handed in. **Bring a calculator to the quiz**.

1. In your text, the second full paragraph on Page 206 gives data for examining the relationship of race to belief in life after death. In the text, the result of fitting a log-linear model of independence is shown in Table 7.1.

   (a) First, make a 3×2 table, and write μ_{ij} in each cell, using the λ notation.

   (b) The software (possibly SAS `proc genmod`) uses a different parameterization of the model than R does. In lecture, we have been R's parameterization. Please fit the same model with R, obtaining both the $\widehat{\mu}$ values and the $\widehat{\lambda}$ values.

      i. In the book's computer output of Table 7.1, The "Deviance" line has a value of 0.3565. Is this a p-value? Explain.

      ii. Again for that Deviance statistic, why should the degrees of freedom equal two? Explain in terms of β values from a dummy variable regression model for $\log \mu_{ij}$, with effect coding.

      iii. My R output says $\widehat{\mu}_{11} = 1339.62839$ (Whites who believe in an afterlife). Obtain this value from your R output using only the $\widehat{\lambda}$ values.

      iv. Now obtain the same number from the textbook's computer output in Table 7.1

      v. Reproduce $\widehat{\mu}_{32}$ (Other who do not believe in an afterlife) from the book's output.

      vi. Under the independence model, the book finds that the estimated odds of belief in an afterlife are roughly 4.5 to 1 for each race. Obtain this number more precisely for Blacks from your R output using first the $\widehat{\lambda}$ values, and then from the $\widehat{\mu}$ values.

      vii. From R's output, you can easily calculate the estimated grand mean $\widehat{\lambda}$ from the $\widehat{\mu}_{ij}$ values. You know how to do this because $\lambda = \beta_0$. Please obtain R's $\widehat{\lambda}$ from R's $\widehat{\mu}_{ij}$ values.

      viii. A good guess is that the book's software is doing something like a regression model with indicator dummy variables. If so, you should be able to reproduce their estimated intercept $\widehat{\lambda}$ from R's $\widehat{\mu}_{ij}$ values. Try it. Did you already do this question??

      ix. Finally, Table 7.2 in the text shows the result of fitting a saturated model to the Race and Belief in afterlife table. Assuming the book's interaction terms correspond to the regression coefficients of *products* of dummy variables, reproduce the Intercept term in Table 7.2 from the table of *observed* frequencies. What is your reasoning?

[1] Copyright information is at the end of the last page.

2. For the data of Table 7.3,

   (a) How many unique main effects are there in a saturated model? The answer is a positive integer.

   (b) How many unique two-factor interaction effects are there in a saturated model? The answer is a positive integer.

   (c) How may unique 3-factor interaction effects are there in a saturated model? The answer is a positive integer.

   (d) Suppose you were testing goodness of fit for the model (AC, AM, CM), which could also be written $(AC)(AM)(CM)$. What are the degrees of freedom for the test? Briefly explain why in terms of β values.

3. Consider the model (AC, AM, CM) for the data of Table 7.3. The model could also be written $(AC)(AM)(CM)$.

   (a) Get the data of Table 7.3 into R, and fit the model.

   (b) Compare your estimated expected frequencies with Table 7.4. Do they match?

   (c) Compare your tests of fit with Table 7.7. Do they match?

   (d) Table 7.6 is for this same model. How can you tell just from the degrees of freedom?

   (e) Reproduce the Likelihood Ratio test of A*M in Table 7.6.

4. Recall the Florida Death Penalty data. Prisoners were either Black or White, the murder victim was either Black or White, and they either got the death penalty or not. For each question below, you would need to do a hypothesis test in order to give an answer. Using bracket notation, give the full and the reduced model for each question.

   (a) Controlling for Victim's race, is the Prisoner's race related to whether he got the death penalty?

   (b) Controlling for Prisoner's's race, is the Victim's race related to whether the prisoner got the death penalty?

   (c) Are victim's race and prisoner's race conditionally independent given Death penalty?

   (d) Are victim's race and the death penalty conditionally independent given prisoner's race?

   (e) Is $(PR, VR)(DP)$ an improvement over the model of complete independence?

   (f) Does $(PR, VR)(DP)$ fit adequately?

5. Why would we say $(PR)(VR)(DP)$ is "nested" within $(PR, VR)(DP)$?

6. Give the model $(PR, VR)(DP)$ in λ notation.

7. How many tests of conditional independence are possible for the Death Penalty data? What are the degrees of freedom for each one?
