STA 312f22 Assignment Seven[1]

These questions are practice for the quiz on Friday Nov. 11th, and are not to be handed in. The R part is deferred until next week.

1. Prove that the greater the log offs, the greater the probability.
2. If two events have equal probability, the odds ratio equals ____.
3. For a multiple logistic regression model, if the value of the jth explanatory variable is increased by c units and everything else remains the same, the odds of Y=1 are ____ times as great. Show the calculation.
4. For a multiple logistic regression model, let $P(Y_i = 1|x_{i,1}, \ldots, x_k) = \pi(\mathbf{x}_i)$. Show that a linear model for the log odds is equivalent to

$$\pi(\mathbf{x}_i) = \frac{e^{\beta_0+\beta_1 x_1+\ldots+\beta_{p-1}x_k}}{1+e^{\beta_0+\beta_1 x_1+\ldots+\beta_{p-1}x_k}} = \frac{e^{\mathbf{x}_i'\boldsymbol{\beta}}}{1+e^{\mathbf{x}_i'\boldsymbol{\beta}}}$$

5. Write the log likelihood for the last question, and simplify it as much as possible.
6. In logistic regression, the *null model* is a model with no explanatory variables. That is, $\beta_1 = \beta_2 = \cdots = \beta_k = 0$. There is just one unknown parameter, β_0. For this special case, a closed-form expression for the MLE is available. Derive it.
7. That last question was an example of the *invariance principle* of maximum likelihood estimation, which says the MLE of a function of the parameter is that function of the MLE. It is very handy. Now, still considering a logistic regression model with no explanatory variables,
   (a) Suppose p (the sample proportion of $Y = 1$ cases) is 0.57. What is $\widehat{\beta}_0$? Your answer is a number.
   (b) Suppose $\widehat{\beta}_0 = -0.79$. What is p? Your answer is a number.
8. It is natural to estimate $\mathbf{x}_i'\boldsymbol{\beta}$ with $\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n$. What is the asymptotic (approximate large-sample) distribution of $\mathbf{x}_i'\widehat{\boldsymbol{\beta}}_n$? Use the formula sheet and show your work. Give the asymptotic expected value and variance.
9. We are working toward the Wald test of $H_0 : \mathbf{L}\boldsymbol{\beta} = \mathbf{h}$.
   (a) What is the asymptotic distribution of $\mathbf{L}\widehat{\boldsymbol{\beta}}_n$? What facts from the formula sheet are you using? Give the expected value and covariance matrix.
   (b) The asymptotic distribution of W_n on the formula sheet uses another fact from the formula sheet. Which one?
   (c) Why does the Wald statistic have r degrees of freedom?
10. In logistic regression, the z-test of $H_0 : \beta_j = 0$ uses the test statistic $z = \frac{\widehat{\beta}_j}{se_{\widehat{\beta}_j}}$. Show that for the Wald test of this null hypothesis, $W_n = z^2$.

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

11. Consider a logistic regression in which the cases are newly married couples with both people from the same religion, the explanatory variable is religion (A, B, C and None – let's call "None" a religion), and the response variable is whether the marriage lasted 5 years (1=Yes, 0=No).

    (a) Make a table with four rows, showing how you would set up indicator dummy variables for Religion, with None as the reference category.

    (b) Add a column showing the odds of the marriage lasting 5 years. The *symbols* for your dummy variables should not appear in your answer, because they are zeros and ones, and different for each row. But of course your answer contains β values.

    (c) What is the ratio of the odds of a marriage lasting 5 years or more for Religion C to the odds of lasting 5 years or more for No Religion? Answer in terms of the β symbols of your model.

    (d) What is the ratio of the odds of lasting 5 years or more for religion A to the odds of lasting 5 years or more for Religion B? Answer in terms of the β symbols of your model.

    (e) You want to test whether Religion is related to whether the marriage lasts 5 years. State the null hypothesis in terms of one or more β values.

    (f) You want to know whether marriages from Religion A are more likely to last 5 years than marriages from Religion C. State the null hypothesis in terms of one or more β values.

    (g) You want to test whether marriages between people of No Religion have a 50-50 chance of lasting 5 years. State the null hypothesis in terms of one or more β values.
