

STAT 8311/PUBH 8401 HW 12

Due Dec. 14, 2022

Exercise 1. Show that the gamma distribution can be written in the form

$$f(y_i) = \exp\left\{ [y_i\theta_i - b(\theta_i)]/a(\phi) - c(y_i, \phi) \right\}.$$

Hint: parameterize the density in terms of the mean and coefficient of variation.

1. Derive the mean and variance using the properties of the exponential family.
2. Derive the canonical link.

Exercise 2. Provide the detailed steps of the IWLS algorithm for gamma regression with the canonical link (using parameterization from previous exercise).

Exercise 3. Derive the deviance and Pearson's χ^2 statistic for the gamma distribution (using parameterization from previous exercises).

Exercise 4. Consider a logistic regression model for disease status (a binary variable coded so that 1 means diseased and 0 means no disease) with a single explanatory variable that is an indicator variable for an exposure (e.g. the disease is lung cancer and the indicator is for smoking). Define

n_{00} = the number of times the predictor and the response are both 0

n_{01} = the number of times the predictor is 0 and the response is 1

n_{10} = the number of times the predictor is 1 and the response is 0

n_{11} = the number of times the predictor and the response are both 1

1. Find expressions for the MLEs of the regression coefficients in terms of n_{ij} for $i, j = 0, 1$.
2. Show that one can approximate the standard error of the regression coefficient associated with the indicator variable with $1/n_{00} + 1/n_{10} + 1/n_{01} + 1/n_{11}$.