Biostat 200C Homework 2

Due Apr 26 @ 11:59PM

Q1. Beta-Binomial

Let Y_i be the number of successes in n_i trials with

$$Y_i \sim \text{Bin}(n_i, \pi_i),$$

where the probabilities π_i have a Beta distribution

$$\pi \sim \text{Be}(\alpha, \beta)$$

with density function

$$f(x; \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad x \in [0,1], \alpha > 0, \beta > 0.$$

1.1

Find the mean and variance of π.

1.2

Find the mean and variance of Y_i and show that the variance of Y_i is always larger than or equal to that of a Binomial random variable with the same batch size and mean.

Q2. Poisson regression log-likelihood

Let $Y_1, \ldots, Y_n$ be independent random variables with $Y_i \sim \text{Poisson}(\mu_i)$ and $\log \mu_i = \mathbf{x}_i^T \boldsymbol{\beta}$, $i = 1, \ldots, n$.

2.1

Write down the log-likelihood function.

2.2

Derive the gradient vector of the log-likelihood function with respect to the regression coefficients $\boldsymbol{\beta}$, i.e. taking derivative with respect to each β_j.

2.3

Show that for the fitted values $\widehat{\mu}_i$ from maximum likelihood estimates

$$\sum_i \widehat{\mu}_i = \sum_i y_i.$$

Therefore the deviance reduces to

$$D = 2 \sum_i y_i \log \frac{y_i}{\widehat{\mu}_i}.$$

Q3. Simpson's paradox

The dataset `death` contains data on murder cases in Florida in 1977. The data is cross-classified by the race (black or white) of the victim, of the defendant and whether the death penalty was given.

3.1

Consider the frequency with which the death penalty is applied to black and white defendants, both marginally and conditionally, with respect to the race of the victim. Is this an example of Simpson's paradox? Are the observed differences in the frequency of application of the death penalty statistically significant?

3.2

Determine the most appropriate dependence model between the variables.

3.3

Fit a binomial regression with death penalty as the response and show the relationship to your model in the previous question.