Homework Assignment 4

STAT 544

Due on March 29, 2021

1. A headline in the February 17, 2014 edition of *The Gainesville Sun* proclaimed a worrisome spike in shark attacks during the previous two years. The reported total number of shark attacks in Florida per year from 2001–2013 were 33, 29, 29, 12, 17, 21, 31, 28, 19, 14, 11, 26, and 23. Are these counts consistent with a null Poisson model or a null negative binomial model? Test the Poisson model against the negative binomial model. Analyze the evidence of a positive linear trend over time.

2. Read Pregibon (1980) and perform a goodness-of-link test for the data in Problem 1. Use the model $\log \lambda_i = \beta_0 + \beta_1 \text{TIME}_i$ even if you did not select this model in Problem 1.

3. Aside from a formal goodness-of-fit test, one analysis that provides a sense of whether a particular GLM is plausible goes as follows. Suppose the ML fitted equation is the true equation. At the observed model matrix $\mathbf{X}$ for the n observations, pseudo-randomly generate n variates with distributions specified by the fitted GLM. Create scatterplots. Do they look like the scatterplots for the observed data?

   Do this for a Poisson loglinear model for the horseshoe crab data, with Y being the number of satellites and x being the carapace width. Does the variability about the fit resemble that in the observed data, including a similar number of zeros and large values? Repeat the procedure a few times to get a better sense of how the scatterplot for the observed data differs from what we would see if the Poisson GLM truly held.

4. Suppose that $\boldsymbol{Y}$ are iid Poisson outcomes with $\mathbb{E}Y_i = \mu$.

   (a) Find the likelihood equations for estimating μ.

   (b) Show that $\hat{\mu} = \bar{Y}$ regardless of the link function.

   (c) For testing $H_0 : \mu = \mu_0$, show that the likelihood-ratio statistic simplifies to

   $$-2(\ell_0 - \ell_1) = 2\{n(\mu_0 - \bar{Y}) + n\bar{Y}\log(\bar{Y}/\mu_0)\}.$$

5. Does the inflated-variance quasi-likelihood approach make sense as a way to generalize the OLM having $v(\mu_i) = \sigma^2$? Why or why not?

6. For the CMP-μ distribution, do the following.
   (a) Let $\nu = 1$, and show that the CMP-μ pmf reduces to the pmf of the Poisson distribution with mean μ.
   (b) Let $\nu = 0$, and show that the CMP-μ pmf reduces to the pmf of the geometric distribution with probability parameter $p = 1/(\mu + 1)$.
   (c) Show that as $\nu \to \infty$ the CMP-μ distribution approaches the Bernoulli distribution with probability parameter $p = \lambda/(1 + \lambda) = \mu$.

References

Pregibon, D. (1980). Goodness of link tests for generalized linear models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 29(1):15–24.