RA module 22: Poisson and Gamma GLMs practice problems

(The attached PDF file has better formatting.)

Fox Regression analysis Chapter 15 Structure of Generalized linear models

** Exercise 22.1: Conditional distribution

The range of the dependent variable depends on the conditional distribution in the generalized linear model.

What are the ranges of the following conditional distributions?'

A. Gaussian (normal) distribution
B. Binomial distribution
C. Poisson distribution
D. Gamma distribution
E. Lognormal distribution

Part A: The Gaussian (normal) distribution has a range of $(-\infty, +\infty)$. This distribution is not appropriate for variables which don't have negative values or distributions which are skewed.

Part B: The binomial distribution has a range of $((0, 1, \ldots, n_j) / n_j)$, where n_j is the number of exposures.

Illustration: A study of retention rates asks how many policyholders renew. If there are n_j policyholders, 0 to n_j may renew. (Property-casualty insurance uses *renewal rate* instead of *retention rate*.) Studies of new drugs (medications) ask whether the patient recovers from illness or does not recover (or dies vs does not die).

Part C: The Poisson distribution has a range of $(0, 1, 2, \ldots)$. Claim counts may be 0, 1, 2, ... (any integer).

Illustration: Studies of insurance claim counts (or claim frequencies) use Poisson distributions.

Parts D and E: The Gamma distribution and the lognormal distribution have ranges of $(0, +\infty)$, and they are positively skewed.

Illustration: The Gamma and lognormal distributions are used for stock prices and insurance claim severities.

Jacob: Fox doesn't discuss the lognormal distribution. This seems strange, as lognormal distributions are used for many actuarial and financial distributions. For example, the Black-Scholes formula assumes stock prices have a lognormal distribution.

Rachel: The lognormal distribution is not a member of the exponential family of distributions. The Gamma distribution is similar to the lognormal distribution and is a member of the exponential family of distributions. For generalized linear models, we use Gamma distributions instead of lognormal distributions.

Jacob: The Gamma distribution differs from the lognormal distribution. With modern computers, we can do maximum likelihood estimates using either distribution. Why not use the distribution that fits the data better?

Rachel: The important item is the relative variance as a function of the expected value. For both the Gamma and lognormal distributions, the variance is proportional to the square of the mean. Maximum likelihood estimation gives (nearly) the same result for these two distributions.

** Exercise 22.2: Residual deviance

The likelihood for the model being tested is 8% and the likelihood for the saturated model is 10%.

A. What is the loglikelihood for the model being tested?
B. What is the loglikelihood for the saturated model?
C. What the residual deviance for the model being tested?

Part A: $ln(0.08) = -2.52573$

Part B: $ln(0.10) = -2.30259$

Part C: $2 \times [\, ln(0.10) - ln(0.08) \,] = 0.44629$