Extensive Assignment 1

STAT 601

Andee Kaplan & Maggie Johnson

February 28, 2014

Model

Let $\{Y_{ij} : i = 1, \ldots, n_j, j = 1, \ldots, m\}$ represent the quantity of green beans sold on day i at store j, let $\{Z_{ij} : i = 1, \ldots, n_j, j = 1, \ldots, m\}$ represent the unobservable construct of consumer interest on day i at store j, and let $\{x_{ij} : i = 1, \ldots, n_j, j = 1, \ldots, m\}$ be the price of green beans on day i at store j. There are $m = 191$ stores in the midwest region. Then, we impose the following model.

$$
\begin{aligned}
Z_{ij} &\stackrel{\text{iid}}{\sim} \text{Bern}(p_{ij}) \\
Y_{ij}|Z_{ij} = 1 &\stackrel{\text{indep}}{\sim} \text{Pois}(\lambda_{ij}) \\
Pr(Y_{ij} = 0|Z_{ij} = 0) &= 1
\end{aligned}
$$

This yields the following marginal distribution of Y_{ij} for $\lambda_{ij} > 0$ and $0 < p_{ij} < 1$:

$$
f(y_{ij}|p_{ij}, \lambda_{ij}) = [(1 - p_{ij}) + p_{ij} \exp(-\lambda_{ij})] \, \mathbb{I}\{y_{ij} = 0\} + \left[\frac{p_{ij}}{y_{ij}!} \lambda_{ij}^{y_{ij}} \exp(-\lambda_{ij})\right] \mathbb{I}\{y_{ij} > 0\}
$$

Systematic Components

We would like to have both the parameters p_{ij} from the Bernoulli distribution and the parameters λ_{ij} from the Poisson portion of the model to be further modeled as a function of price (x_{ij}). So, we will use the constructs of GLM to model the expected values of Y_{ij} and Z_{ij} as inverse link functions of the simple regression equations.

$$
\begin{aligned}
\mathrm{E}(Z_{ij}) &= p_{ij} \\
\mathrm{E}(Y_{ij}) &= \sum_{y_{ij}=1}^{\infty} p_{ij} y_{ij} \lambda_{ij}^{y_{ij}} \exp(-\lambda_{ij}) \\
&= \sum_{y_{ij}=0}^{\infty} p_{ij} y_{ij} \lambda_{ij}^{y_{ij}} \exp(-\lambda_{ij}) \\
&= p_{ij}\lambda_{ij}
\end{aligned}
$$

We will use a logit-link for $\mathrm{E}(Z_{ij})$ and a log-link for $\mathrm{E}(Y_{ij})$, which yields the following values.

$$
\begin{aligned}
\log\left(\frac{p_{ij}}{1-p_{ij}}\right) &= \beta_0 + \beta_1 x_{ij} \\
\Rightarrow p_{ij} &= \frac{\exp(\beta_0 + \beta_1 x_{ij})}{1 + \exp(\beta_0 + \beta_1 x_{ij})}
\end{aligned}
$$

$$
\begin{aligned}
\log(p_{ij}\lambda_{ij}) &= \beta_2 + \beta_3 x_{ij} \\
p_{ij}\lambda_{ij} &= \exp(\beta_2 + \beta_3 x_{ij}) \\
\Rightarrow \lambda_{ij} &= \frac{\exp(\beta_2 + \beta_3 x_{ij})}{\exp(\beta_0 + \beta_1 x_{ij})}(1 + \exp(\beta_0 + \beta_1 x_{ij}))
\end{aligned}
$$

Likelihood

To obtain the store likelihood $L_j(\beta_0, \beta_1, \beta_2, \beta_3, \beta_4)$, first we will write the joint density for store j.

$$
\begin{aligned}
L_j(\beta_0, \beta_1, \beta_2, \beta_3, \beta_4) &= f(y_{1j}, \ldots, y_{n_j,j} | \beta_0, \beta_1, \beta_2, \beta_3, \beta_4) \\
&= \prod_{i=1}^{n_j} f(y_{ij} | \beta_0, \beta_1, \beta_2, \beta_3, \beta_4) && \text{(independence)} \\
&= \prod_{i=1}^{n_j} \Big([(1 - p_{ij}(\boldsymbol{\beta})) + p_{ij}(\boldsymbol{\beta}) \exp(-\lambda_{ij}(\boldsymbol{\beta}))] \, \mathbb{I}\{y_{ij} = 0\} + \\
&\qquad \left[\frac{p_{ij}(\boldsymbol{\beta})}{y_{ij}!} \lambda_{ij}(\boldsymbol{\beta})^{y_{ij}} \exp(-\lambda_{ij}(\boldsymbol{\beta}))\right] \mathbb{I}\{y_{ij} > 0\} \Big)
\end{aligned}
$$