**Evaluating Fence Marker Effectiveness Using Zero-Inflated Poisson Regression**

**Bryan Stevens**

**Background:**

Greater sage-grouse (*Centrocercus urophasianus*) are a widespread and relatively common bird species in high-desert sagebrush-steppe habitats of western North America. Sage-grouse have a polygamous mating strategy called lekking, and congregate at traditional display grounds called leks from March-May every year to attract mates and breed. Sage-grouse commonly use the same lekking areas for breeding in repeated years in the absence of disturbance, and leks are known that have been active in the same location since the 1950’s. Fences are common on many western rangelands inhabited by sage-grouse, and observational studies have documented sage-grouse mortality as a function of the birds flying into fences. Research from Europe and to a lesser extent here in North America has suggested grouse as a group are highly susceptible to colliding with man-made infrastructure (e.g., fences, power lines, etc.). Researchers working with lesser prairie-chickens (*Tympanuchus pallidicinctus*) in Oklahoma developed a fence marking mitigation method designed to reduce grouse fence collision frequency by increasing the visibility of fences via vinyl siding trim markers. Anecdotal evidence suggests this method likely reduces grouse fence collision risk, however, no studies systematically studied effectiveness of the fence markers. As part of my wildlife MS thesis project I conducted a field experiment to quantify the effectiveness of fence markers at reducing sage-grouse fence collision frequency in high-risk areas.

**Study Design:**

I used field sampling during spring of 2009 as a pilot study to locate potentially high risk areas for sage-grouse fence collision. I used the identified high-risk areas to conduct the field experiment, and replicated the same design on each site. I started with a 3 km segment of fence on each study area, which I subsequently partitioned into 6, 500 m segments, each separated by 50 m. I then randomly selected 3, 500 m segments to apply the fence marking treatment, and used the remaining 3, 500 m segments as unmarked controls. This simple design was replicated on 8 different study areas across south central and southeast Idaho, resulting in 48 total 500 m fence segments, which served as the experimental units in the analysis. I then monitored each 500 m fence segment for sage-grouse collisions across the breeding season (March-May) during spring 2010. Also, since each of the 500 m fence segments were not homogenous with respect to biological factors that may influence sage-grouse fence collision, I included 2 quantitative variables as potential statistical controls the regression analysis: distance (m) from the center of the fence segment to the nearest sage-grouse lek, and size of the nearest sage-grouse lek (number of birds) for each of the 500 m fence segments. This resulted in a zero-inflated count dataset of sage-grouse fence collisions for each of the 48 fence segments over the entire breeding season.

**Zero-Inflated Poisson Regression Model**

Since the data were counts and dominated by zero observations, I used zero-inflated Poisson regression to analyze the data. Zero-inflated Poisson models (hereafter ZIP models) are statistical mixture models, where observed counts are treated as a binomial mixture of a point mass at zero with a Poisson random variable. As such, ZIP models take the following form:

$$P(Y=0)=1-p+ p×e^{-λ},$$

$P(Y=r)=p\frac{λ^{r}e^{-λ}}{r!}$, $ r=1,2,…,$

where

$$logit\left(p\right)= β\_{0}+ β\_{1}\left(X\_{1}\right)+ … + β\_{k}\left(X\_{k}\right)$$

$log(λ)= γ\_{0}+ γ\_{1}\left(X\_{1}\right)+ … + γ\_{k}\left(X\_{k}\right)$.

In this analysis $r$ represents integer valued count data, $λ$ represents the mean number of collisions at a given fence segment, and *p* represents the binomial mixture probability (i.e., probability of the data coming from the Poisson distribution, where 1 - $p$ is probability of data coming from inflated zero count; Lambert 1992, Martin et al. 2005). As such, ZIP models are a type of generalized linear model that facilitate modeling both expected counts and probability of an group membership (i.e., zero, Poisson) simultaneously as a function of covariates. For this analysis I modeled the binomial mixture probability using an intercept only logistic regression, and thus assumed constant probability of data point inclusion in the Poisson distribution across varying environmental conditions at each fence segment. This assumption may have oversimplified reality, however, my sample sizes were not large enough to fit a more parameterized model with the mixture probability as a function of covariates. This is performed by combining logistic and Poisson regression into 1 joint likelihood model. This analysis was completed using the pscl package and zeroinfl function (Zeileis et al. 2008) in the R statistical computing language. Analysis results and model predictions will be presented and discussed in the accompanying presentation. More information on zero-inflated Poisson regression, including the full likelihood, can be found in the accompanying citations, as well as pages 392-394 of the Regression 550 textbook.

References:

Lambert, D. 1992. Zero-inflated Poisson regression, with an application to defects manufacturing. Technometrics 34:1-14.

Martin, T. G., B. A. Wintle, J. R. Rhodes, P. M. Kuhnert, S. J. Low-Choy, A. J. Tyre, and H. P. Possingham. 2005. Zero tolerance ecology: improving ecological inference by modeling the source of zero observations. Ecology Letters 8:1235-1246.

Zeileis, A., C. Kleiber, and S. Jackman. 2008. Regression models for count data in R. Journal of Statistical Software 27:1-25.