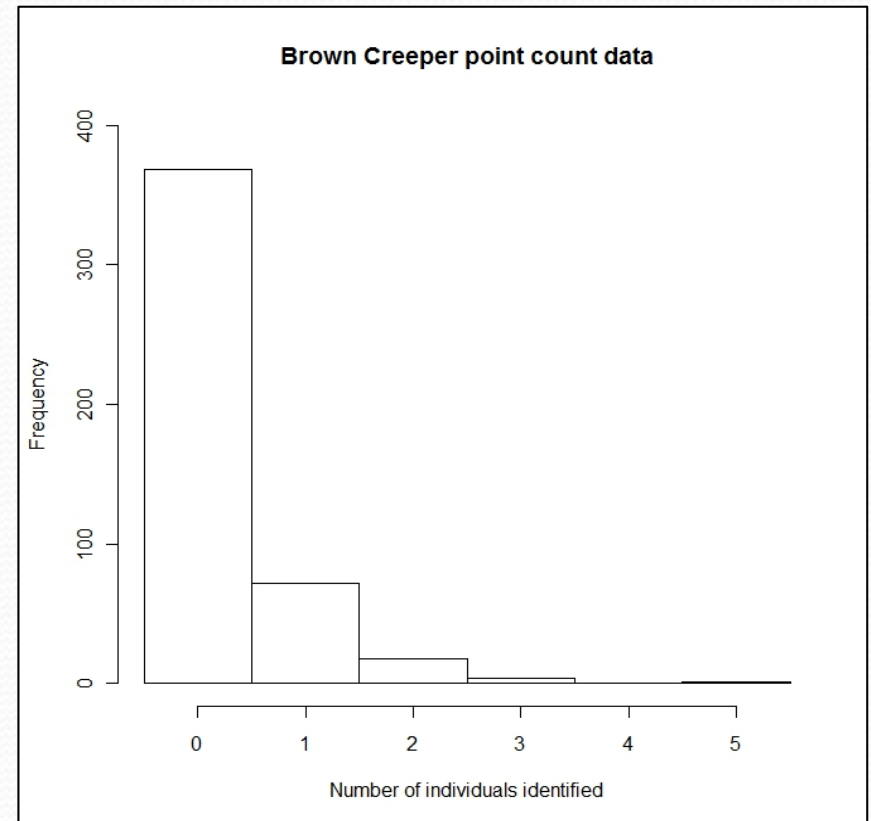


Topic 9: Over-dispersion + zero-inflation

Other GLM's and overdispersion

- Poisson
- Negative Binomial (NB)
- Zero Inflated Poisson
- Zero Inflated NB



Test which distribution fits the data best

- Histogram to visually identify the proportion of zero counts in the data set.
- log Likelihood, AIC (Akaike 1974) and correctly predicted zeros.
- Formal tests:
 - likelihood ratio tests
 - Poisson - negative binomial distribution.
 - zero-inflated Poisson and zero-inflated negative binomial.
 - Vuong test:

$$V = \frac{\text{sqrt}(N) * \text{mean}(m)}{s_m * m}$$

$$m = \ln \left(\frac{\mu_1}{\mu_2} \right)$$

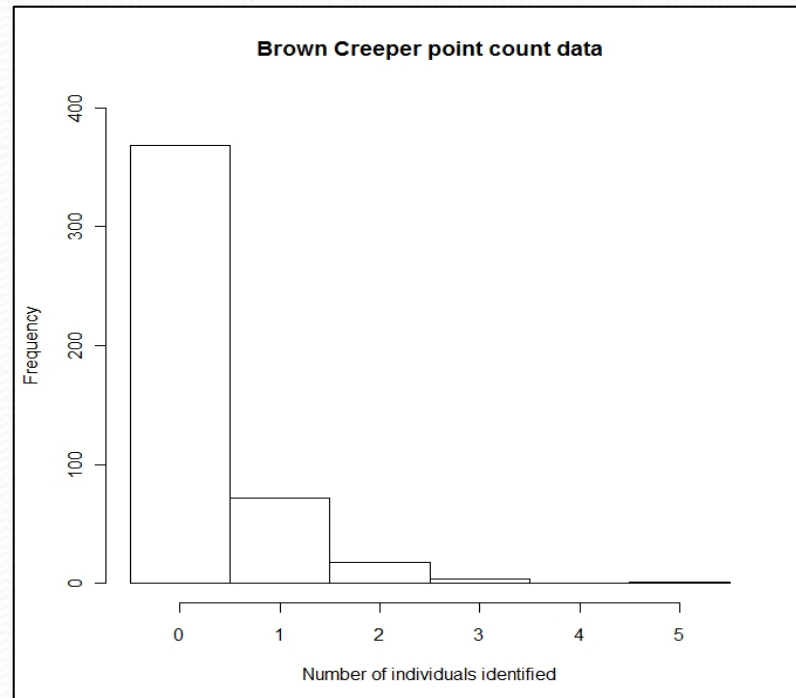
μ_1 = predicted probability of y for the zero-inflated model; μ_2 = predicted probability of y for the base model; s_m = standard deviation of m; N = number of observations in each model (both must use the same observations). test statistic V asymptotically normal.

$V > 1.96$ = zero-inflated model is preferred; $V < -1.96$ = the base model is preferred; between neither model is preferred).

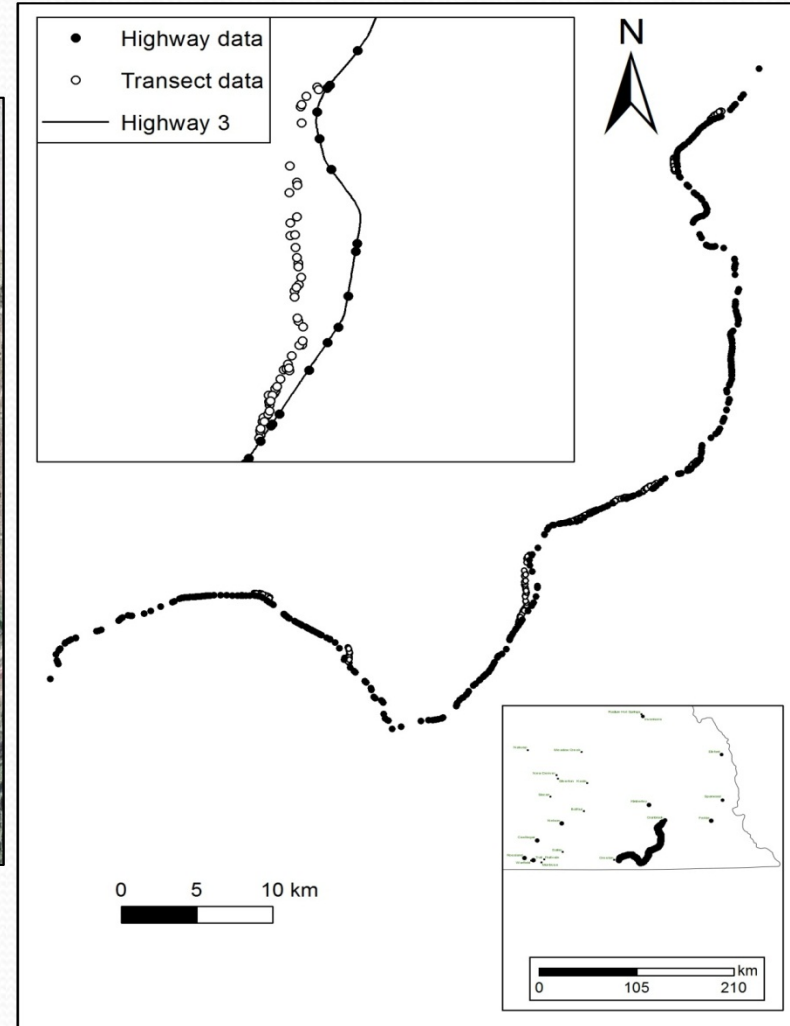
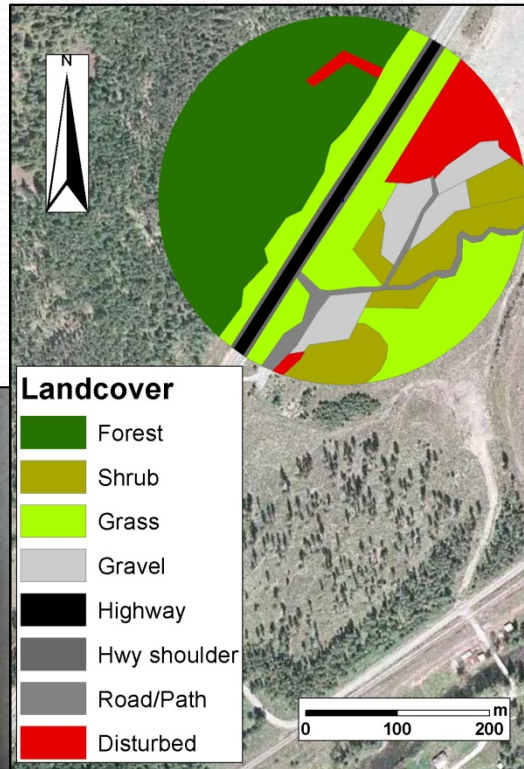
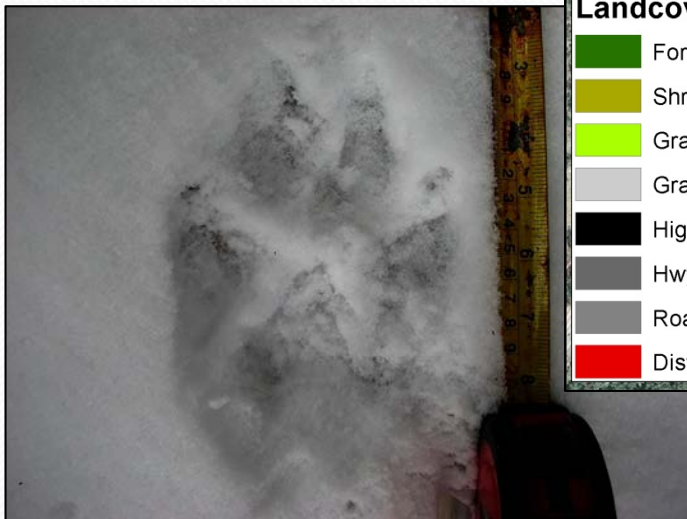
Results

True o's	GLM		Zero inflated	
368		negative		negative
	Poisson	binomial	Poisson	binomial
no. parameters	4	5	8	9
-2* log Lik	-274.924	-272.284	-266.696	-266.518
AIC	557.847	554.569	549.392	551.036
predicted o's	361	368	368	369

- Likelihood ratio tests
 - Poisson – **negative binomial**:
 - $\chi^2 = 5.2785$, $df=1$, $p=0.010795$
 - ZI Poisson – ZI NB:
 - $\chi^2 = 0.3552$, $df=1$, $p=0.2756$
- Vuong tests:
 - **ZI NB – NB**:
 - 1.739 ($p = 0.041$)
 - ZI Poisson – Poisson:
 - 1.505 ($p = 0.066$)



Multiscale highway effects on a large mammal community



Setup



- Highway 3 in SE BC between Creston and Cranbrook
- Identify potential crossing sites
- 2 seasons of snow track data
 - Highway 96km
 - 9 Transects off highway total 13km
- Species data: deer, moose, elk, coyote, bobcat, cougar, wolf, ...



Analysis

- Distribution choices similar to previous example
- Models:
 - 3 scales individually and in combination (200, 500, 1000m) using Landsat derived variables
 - hand digitized variables (200m)
 - Combine all
- Predictive maps using Landsat data

