**DRAFT - Sample Typology of Count Models**

Initial or "Standard" Regression Model for Counts

1. Poisson Regression:
    - Assumes the count data follow a Poisson distribution.
    - Assumes mean and variance are equal.
    - Does not account for overdispersion (i.e., when variance exceeds the mean).

Accommodating Underdispersion (or overdispersion)

2. Generalized Poisson Regression:
    - Generalizes the Poisson distribution to allow for extra dispersion.
    - Accounts for overdispersion by introducing an additional parameter.
    - Provides more flexibility in modeling the relationship between predictors and count outcomes.

Accommodating Overdispersion

3. Quasi-Poisson Regression:
    - Accounts for overdispersion without specifying the exact distribution.
    - Estimates the dispersion parameter without assuming a specific distribution.
    - Not preferred choice for modeling count data when the mean and variance differ.
4. Negative Binomial Regression (NB2, NB1):
    - Assumes the count data follow a negative binomial distribution.
    - Allows for overdispersion by introducing an additional dispersion parameter.
    - Can handle cases where variance exceeds the mean.
    - Has two forms: NB2 (variance is a quadratic function of the mean) and NB1 (variance is a linear function of the mean). NB2 is more commonly used.
5. Poisson Inverse Gaussian
    - Like NB, this is a mixture of two distributions: Poisson and Inverse Gaussian
    - Hilbe (2014) notes that the PIG model is better than NB when the data are highly overdispersed with data massed at counts for 1 and 2
    - Available in R (gamlss) and in Stata, not in other commercial software
6. Conway-Maxwell Poisson Regression:
    - Generalizes the Poisson distribution to account for both underdispersion and overdispersion.
    - Introduces a second dispersion parameter, the Conway-Maxwell parameter.
    - Provides flexibility in capturing various patterns of dispersion in count data.

Intentional Treatment of Zeros Through Zero-Inflation Models

7. Zero-Inflated Poisson Regression:
    - Accounts for excessive zeros in count data by modeling a separate process for generating zeros.
    - Consists of two components: a Poisson component and a zero-inflation component.
    - Assumes that excess zeros are generated by a separate process rather than being a result of the count process.

8. Zero-Inflated Negative Binomial Regression:
    - Combines the zero-inflation concept with the negative binomial distribution.
    - Models both the excess zeros and the count part of the data.
    - Suitable for count data with excessive zeros and overdispersion.

Models involving Truncation

9. Hurdle Regression:
    - Similar to zero-inflated models, but models the hurdle or the decision to have zero counts separately from the count part.
    - Consists of two components: a binary hurdle component and a truncated count component.
    - Zeros could be more or fewer than expected based on count component
    - Appropriate when there is a two-step process involved in generating zero counts and non-zero counts.
    - Typical Hurdle models are Hurdle-Poisson, and Hurdle-NB
10. Double Hurdle Model:
    - Dependent variables may take on the endpoints of an interval with positive probability and that are continuously distributed over the interior of the interval.
    - Involves a combination of truncated models, i.e., between two hurdle points, and then for the 2$^{nd}$ hurdle to the largest count.

Additional Complexities for Modeling Counts

11. Generalized Additive Models
    - Flexible extension of the generalized linear model (GLM) that allows for non-linear relationships between predictors and the response variable
    - Smooth functions are used to model the non-linear relationship between predictors and the response variable.
    - Splines are commonly used to represent the smooth functions.
    - The smooth functions are flexible and can capture complex patterns, such as curves, bends, and interactions.
12. Tobit Model:
    - The Tobit model assumes that there is an underlying continuous variable, but it is only partially observed due to censoring.
    - The censored observations are known to fall below or above certain thresholds, while the uncensored observations are fully observed.
    - Censoring can occur from below (left-censoring) or from above (right-censoring), or both.
    - The Tobit model assumes an unobserved continuous latent variable that follows a linear regression structure.
13. Random Effects (Mixed Effects) Models:
    - Incorporates random effects to account for clustering or nested structures in the count data.
    - Models the within-group correlations and allows for estimating group-level variation.
    - Can be combined with Poisson, negative binomial, or other count models to handle clustered count data.
    - Nearly all models described in this Typology have a Mixed-effects counterpart