**Name: Krista Miller**

**Question 1: Poisson/quasi-poisson/negative binomial model comprehension check**

1. How does quasi-poisson regression loosen the distributional assumption of the model residuals in the Poisson model? How does the negative binomial model do so?

The quasi-poisson regression takes the variance instead of being equal to the mean to be a constant multiple of the mean.

The negative binomial model allows the variance to grow as a multiple of the mean.

1. Comparing the Poisson regression model with the quasi-poisson regression model, are the coefficient estimates noticeably different? What about the standard errors?

The coefficients are the same for the Poisson and quasi-poisson regression models. However, for the quasi-poisson regression model, the standard errors are noticeably larger compared to the Poisson model.

1. Where in the output would you look to see if there is evidence of overdispersion in the Poisson regression model? How could you evaluate if the quasi-poisson model or the negative binomial model improve it?

You can see evidence of overdispersion in the Poisson regression model if there is a large residual deviance relative to the degrees of freedom.

**Question 2: Project check-in**

Now that you have your project topic, I want to check in with you about where you are in your work.

1. What is your project topic?

Latent Class Analysis:

Probability of high rating (4-5) for “How would you rate the quality of Starbucks compared to other brands(Coffee Bean, Old Town White Coffee..) to be” given membership in “Dine In” latent class.

1. What are the data requirements of your project topic? That is, what properties must your data have to use this method? You don’t need to have identified a data set to answer this question.

Static, categorical latent variable measured with categorical items. Datasets that are often used with this technique are surveys, where subgroups (AKA latent classes) can be formed from the set of response patterns.

Latent classes are exclusive and exhaustive- every person belongs to one (and only one) latent class, but because of measurement error, the true latent class is unknown.

1. What data set do you plan to use? If your data is publicly available, please provide a link to it. If you have not yet chosen a data set, what kind of data set do you want to obtain?

Currently, we plan to use a [dataset](https://www.kaggle.com/mahirahmzh/starbucks-customer-retention-malaysia-survey?select=Starbucks+satisfactory+survey+encode+cleaned.csv) available of travel reviews, found online from the Starbucks Customer Satisfaction survey on Kaggle.

(more questions on next page)

1. What R package/s do you anticipate using for your method? Please note that using R for your analysis and visualization is required for this project.

Packages: poLCA, glca (for multiple-group LCA models)

1. What resources are you using to learn more about your chosen method? Please list these resources here (e.g., links to websites, books, book chapters, links to videos, etc.). The reason I’m asking you this question is that I would like to build a list of resources for different methods to share; because you are a student, you are the best judge of what materials are most helpful!

Webinar on Latent Class Analysis from PennState Methodology Center:

<https://www.youtube.com/watch?v=BGdvoEzLCYA>

Latent Class Analysis Using R from Princeton Office of Population Research:

<https://pop.princeton.edu/events/2020/latent-class-analysis-using-r>