HW1

HW1

Q1.

a. (4 points) Write an R function or chunk of code to simulate data from a negative binomial regression model using a single continuous covariate and a user specified value of ϕ.

b. (4 points) Use the following four values of $\phi : \{.01, .1, 1, 10\}$ and simulate a dataset. Create a paneled figure that shows y and x for each scenario.

c. (4 points) Use `stan_glm()` to fit Poisson models and Negative Binomial models for each of the four simulated datasets. Create a table or figure that contains the intercept and slope coefficient for each outcome. Then comment on the implications of your results.

d. (4 points) Use `posterior_predict()` and posterior predictive checks to further interrogate the model fit (using the Poisson models)

Q2.

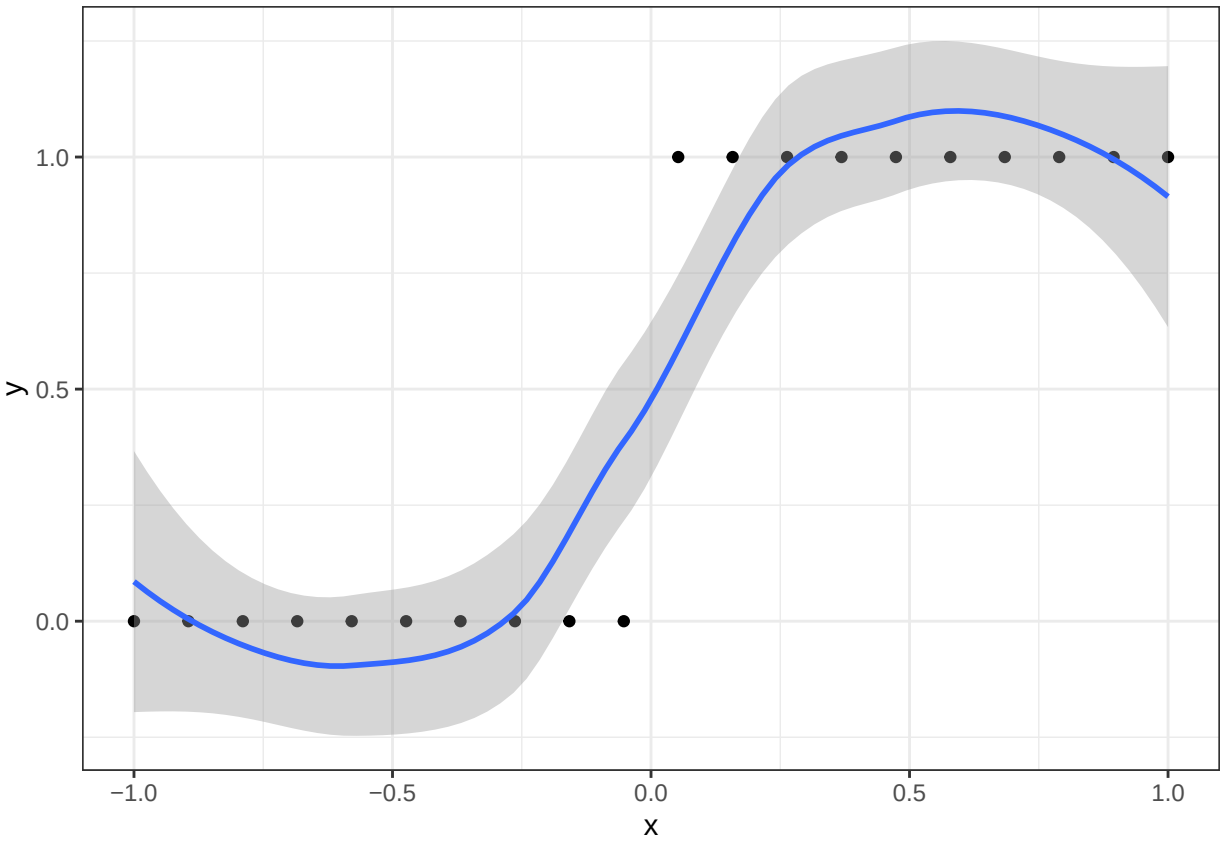
With binary regression, "separation" is a common problem. This occurs when a continuous predictor is perfectly separated with all zeros below a certain point and all zeros above a certain point. See the simulated data below for an example.

```
x <- seq(-1, 1, length.out = 20)
y <- rep(c(0,1), each = 10)

df_sep <- tibble(x=x, y=y)

df_sep %>% ggplot(aes(y=y, x=x)) + geom_point() + theme_bw() + geom_smooth()
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



a. (4 points) Using the figure above - and any other references - define separation and describe why it is problematic.

b. (4 points) Use both `glm` and `stan_glm` to fit the data. Identify the differences in the model output and discuss why they might differ.