Quiz11- MA478

Clark

Consider data that shows survival and death for 539 males diagnosed with lung cancer. Consider a model for the number of deaths as a factor of stage of disease, histology, and follow-up time interval. The data can be downloaded from:

```
cancer_dat <- read.table("https://users.stat.ufl.edu/~aa/glm/data/Cancer.dat",header=T)
```

Note that for this model we want to ensure that `histology`, `stage`, and `time` are treated as factors NOT numeric values.

Use this model to test whether `stage` is significant given `histology` and `time`.

Conduct a goodness of fit test for your model. Are there any concerns?

Repeat your analysis using log of risktime as an offset. Are there still concerns using your goodness of fit test? Which model do you prefer and why?