Stats\*3510 – Assignment 3

Graham Eckel – 0679576

Q3

library(fitdistrplus)

library(ggplot2)

library(reshape2)

dir = "E:\\Google Drive\\...\\Assignment 3\\"

file1 = "ONCHLOR.csv"

file2 = "Table3.csv"

dfChlor = read.table(file=paste(dir,file1, sep=""), header=TRUE, sep=',')

dfTable3 = read.table(file=paste(dir,file2, sep=""), header=TRUE, sep=',')

dfTable3 = dfTable3[order(dfTable3$Concentration..mg.Cl..L..), ]

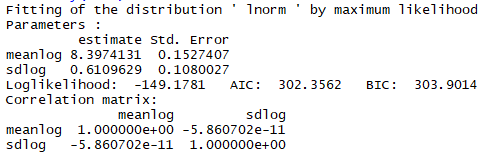
dfTable3$frac = ppoints(dfTable3$Concentration..mg.Cl..L.., 0.5)

**a)**

fit2 = fitdist(dfTable3$Concentration..mg.Cl..L.., 'lnorm')

summary(fit2)

**Summary of fitted model**



hc5\_2 = quantile(fit2, probs = 0.05)

hc5\_2

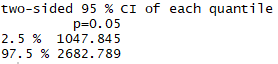
**HC5**



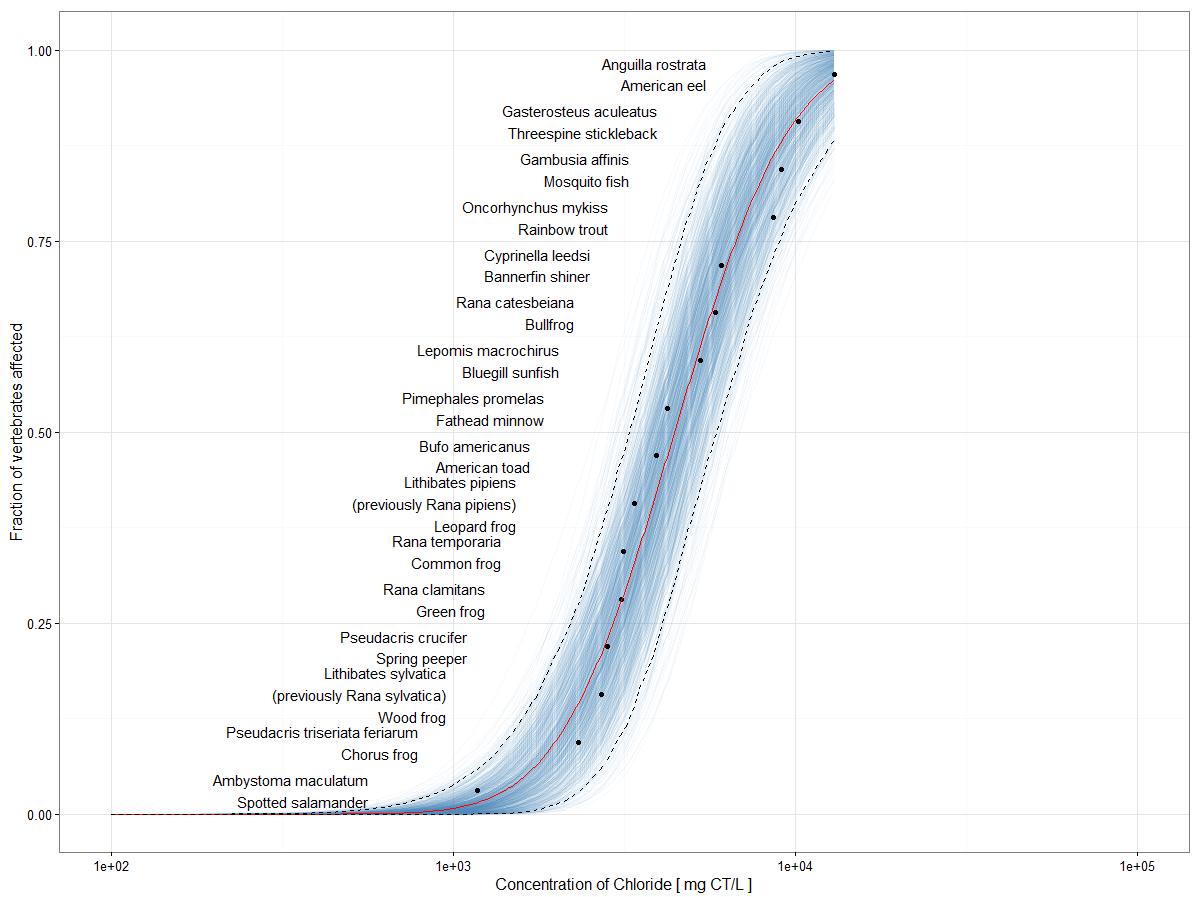
fit2\_bootp = bootdist(fit2, bootmethod = 'param', niter = 1000)

quantile(fit2\_bootp, probs = 0.05)

**95% parametric bootstrapped CI**



* The estimated 95% parametric bootstrapped CI for is (1047.845, 2682.789)
* The estimated HC5 is 163.709
* The short term HC5 identified in the article was 640 mg CI’/L across all species and the long term was 120 mg CI/L. Using the short term number, the Northern riffleshell mussel and then Water flea fell below the HC5.



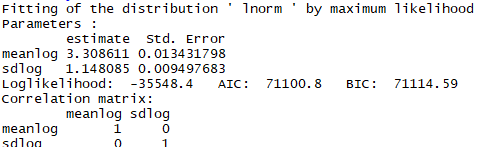
**b)**

dfChlor$frac = ppoints(dfChlor$Conc,0.5)

fit3 = fitdist(dfChlor$Conc,'lnorm')

summary(fit3)

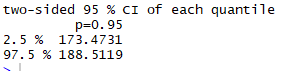
**Summary of fitted model**



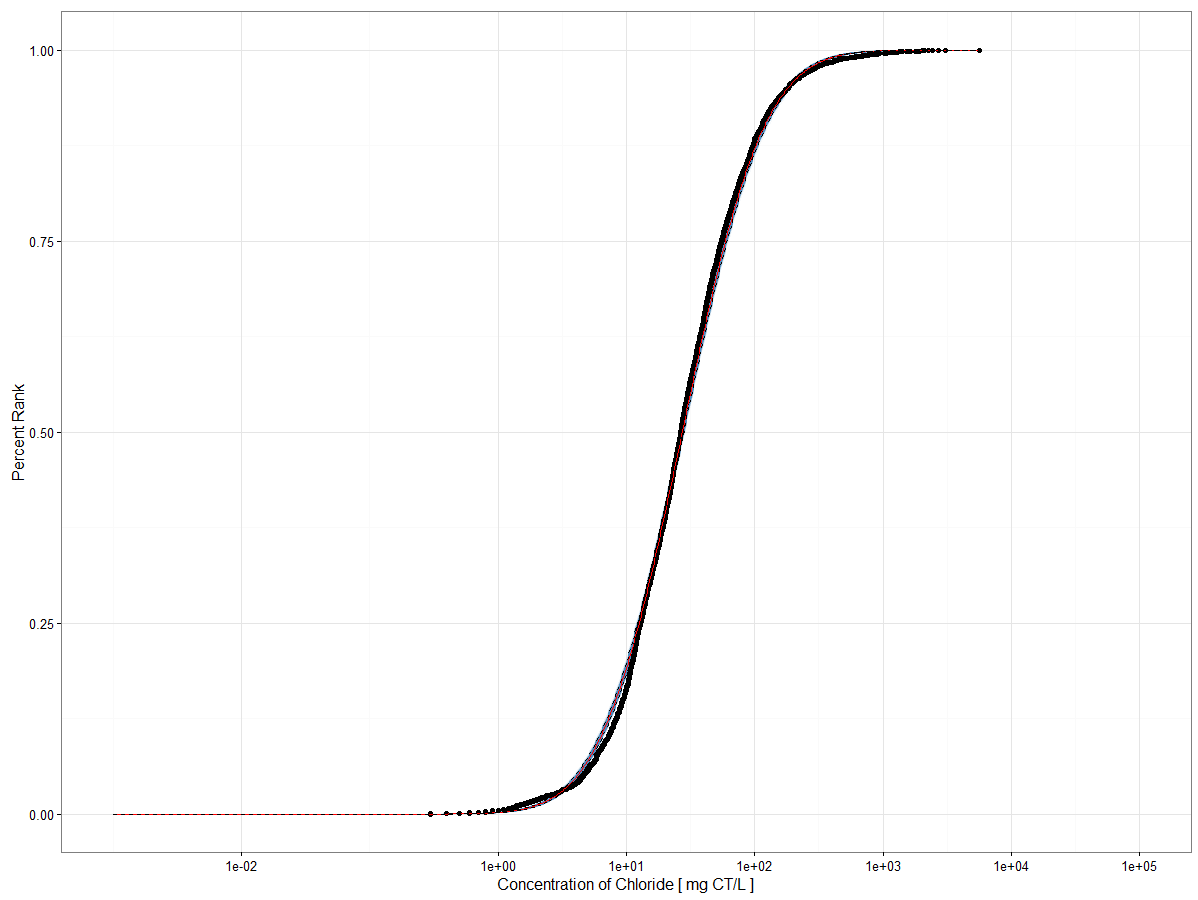
fit3\_bootp = bootdist(fit3, bootmethod = 'param', niter = 1000)

quantile(fit3\_bootp, probs = 0.95)

**95% parametric bootstrapped CI**



* The estimated 95% parametric bootstrapped CI for is (173.4731, 188.5119)



**c)**

ggplot(data = dfTable3) +

geom\_point(aes(x = Concentration..mg.Cl..L.., y = frac), size = 3) +

geom\_point(data = dfChlor, aes(x = Conc, y = frac), size = 2) +

geom\_point(aes(x=hc5\_2$quantiles[1], y = 0.05), size = 5, col = "red") +

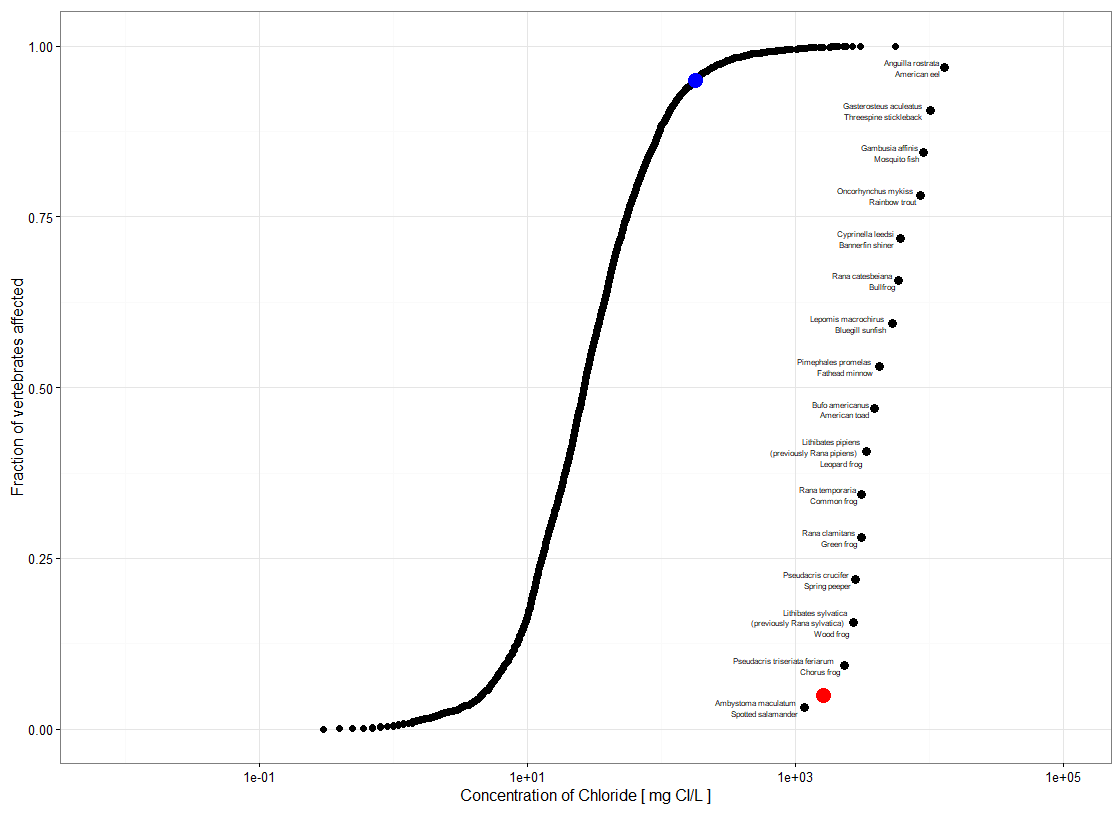
geom\_point(aes(x=hc95\_3$quantiles[1,1], y = 0.95), size = 5, col = "blue") +

geom\_text(aes(x = Concentration..mg.Cl..L.., y = frac, label = Species), hjust = 1.1, size = 2) +

theme\_bw() +

scale\_x\_log10(limits = c(0.0075, 100000)) +

labs(x = expression(paste('Concentration of Chloride [ mg CI/L ]')), y = 'Fraction of vertebrates affected')



This plot contains both the species sensitivity data for the vertebrates and environmental exposure data where the x-axis is on the log10 scale and the y-axis the species affected. The red dot is the HC5 and the blue dot is the 95th percentile. We can see that the number of species affected by the 95th percentile to be 0%. That is, approximately 95% of the time no species will be affected.