lionfish\_GLM

GG

12/17/2018

## R Markdown

library(ggplot2)  
library(tidyverse)

## ── Attaching packages ───────────────────────────────────────────────────────────── tidyverse 1.2.1 ──

## ✔ tibble 1.4.2 ✔ purrr 0.2.4  
## ✔ tidyr 0.8.0 ✔ dplyr 0.7.4  
## ✔ readr 1.1.1 ✔ stringr 1.3.0  
## ✔ tibble 1.4.2 ✔ forcats 0.3.0

## ── Conflicts ──────────────────────────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

library(readxl)  
library(janitor)  
library(dbplyr)

##   
## Attaching package: 'dbplyr'

## The following objects are masked from 'package:dplyr':  
##   
## ident, sql

darwin <- read\_csv("~/Documents/darwin\_stats2.csv")

## Parsed with column specification:  
## cols(  
## lionfish\_den = col\_integer(),  
## fish\_div = col\_double(),  
## fish\_den = col\_double(),  
## fish\_bio = col\_double(),  
## temp = col\_double(),  
## c\_bda = col\_double(),  
## c\_enchry = col\_double(),  
## p\_furc = col\_double()  
## )

library(vegan)

## Warning: package 'vegan' was built under R version 3.4.4

## Loading required package: permute

## Loading required package: lattice

## This is vegan 2.5-1

Visualizing the GLM resutls with a PCA analysis

# remove unwanted variables   
darwin.pca <- darwin[, 2:8]  
darwin.pca[6] = NULL  
darwin.pca[5] = NULL  
darwin.den <- (darwin[, 1] +1)  
   
# apply PCA - scale. = TRUE is highly   
# advisable, but default is FALSE.   
pca <- prcomp(darwin.pca,  
 center = TRUE,  
 scale. = TRUE)

# print method  
print(pca)

## Standard deviations (1, .., p=5):  
## [1] 1.6527303 1.0391372 0.7639914 0.6501548 0.4269572  
##   
## Rotation (n x k) = (5 x 5):  
## PC1 PC2 PC3 PC4 PC5  
## fish\_div -0.007165718 -0.9169657 -0.3946599 0.04103037 -0.04102009  
## fish\_den 0.542395402 -0.1071666 0.2788520 -0.33027914 -0.71237616  
## fish\_bio 0.518714486 0.0871676 -0.3271884 -0.58406751 0.52454742  
## temp -0.457733358 -0.2765148 0.5690557 -0.59457821 0.19149965  
## p\_furc 0.476624942 -0.2522511 0.5793173 0.44110553 0.42310217

In order to create the biplot we need devtools and ggbiplot available from CRAN and github

library(devtools)

## Warning: package 'devtools' was built under R version 3.4.4

##   
## Attaching package: 'devtools'

## The following object is masked from 'package:permute':  
##   
## check

library(ggbiplot)

## Loading required package: plyr

## -------------------------------------------------------------------------

## You have loaded plyr after dplyr - this is likely to cause problems.  
## If you need functions from both plyr and dplyr, please load plyr first, then dplyr:  
## library(plyr); library(dplyr)

## -------------------------------------------------------------------------

##   
## Attaching package: 'plyr'

## The following objects are masked from 'package:dplyr':  
##   
## arrange, count, desc, failwith, id, mutate, rename, summarise,  
## summarize

## The following object is masked from 'package:purrr':  
##   
## compact

## Loading required package: scales

##   
## Attaching package: 'scales'

## The following object is masked from 'package:purrr':  
##   
## discard

## The following object is masked from 'package:readr':  
##   
## col\_factor

## Loading required package: grid

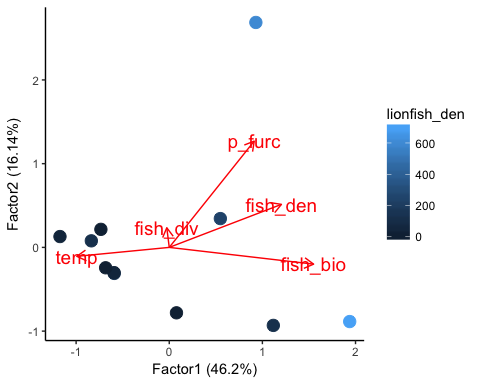
library(ggfortify)

## Warning: package 'ggfortify' was built under R version 3.4.4

##   
## Attaching package: 'ggfortify'

## The following object is masked from 'package:ggbiplot':  
##   
## ggbiplot

d.factanal2 <- factanal(darwin.pca, factors = 2, scores = 'regression')  
pca.plot2 <- autoplot(d.factanal2, data = darwin, colour = 'lionfish\_den', size = 4,  
 loadings = TRUE, loadings.label = TRUE, loadings.label.size = 5) + theme\_classic()  
pca.plot2



ggsave(darwin, plot = pca.plot2, device = "eps", path = "~/Desktop/Lionfish")

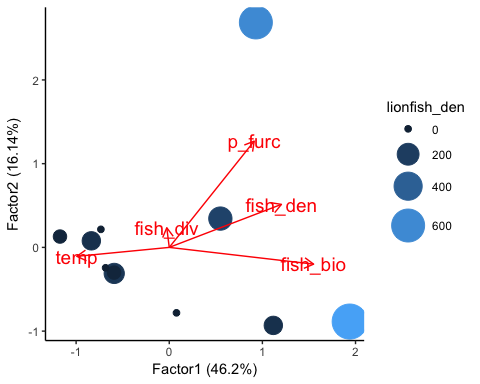
## Saving 5 x 4 in image

Changing size of bubbles to reflect lionfish density

pca.plot3 <- autoplot(d.factanal2, data = darwin, colour = 'lionfish\_den', size = 'lionfish\_den',  
 loadings = TRUE, loadings.label = TRUE, loadings.label.size = 5) + theme\_classic() + scale\_color\_continuous(limits=c(0, 700), breaks=seq(0, 700, by=200)) +  
 guides(color= guide\_legend(), size=guide\_legend()) + scale\_size\_continuous(limits=c(0, 700), breaks=seq(0, 700, by=200)) + scale\_size\_continuous(range = c(2,12))

## Scale for 'size' is already present. Adding another scale for 'size',  
## which will replace the existing scale.

pca.plot3



ggsave(filename = "lionfish\_pca\_rev3.eps", plot = pca.plot3, device = "eps", path = "~/Desktop/Lionfish")

## Saving 5 x 4 in image

NOW FOR THE GLM

GLM model with all variables included showing covariance among possible drivers.

scale((darwin[,2:8]), center = TRUE, scale = TRUE)

## fish\_div fish\_den fish\_bio temp c\_bda c\_enchry  
## [1,] 0.4809800 1.4365108 1.2281635 -0.1297499 0.09054847 2.60244718  
## [2,] -0.0497952 2.0336000 0.5682290 -0.8433746 0.23309789 -0.43401594  
## [3,] -0.1564501 0.8528393 2.0426555 -1.5569993 -0.37682665 0.99732680  
## [4,] 0.2879963 -0.6834913 0.5114883 -0.1297499 -0.73904237 0.08487951  
## [5,] -0.5657100 -0.7103267 -0.5480014 -0.1297499 -0.57546108 -0.64805727  
## [6,] 0.0724305 -0.7639977 -0.8468391 -0.8433703 -0.59882984 -0.48007135  
## [7,] 0.2267893 0.0410664 -0.5568844 -0.1297499 2.35211988 -0.66488296  
## [8,] -1.3837317 -0.1199464 0.1806320 0.5838747 1.36414569 -0.65962493  
## [9,] -0.7697882 -0.6029849 -0.6502334 -0.1297499 -0.73904237 0.10444226  
## [10,] 2.5176680 -0.6029849 -0.7515916 1.2974951 -0.27166725 -0.23756036  
## [11,] -0.6603889 -0.8802847 -1.1776184 2.0111241 -0.73904237 -0.66488296  
## p\_furc  
## [1,] -0.1154827  
## [2,] 2.7300946  
## [3,] 0.4089936  
## [4,] 0.6103919  
## [5,] -0.5969805  
## [6,] -0.4244987  
## [7,] -0.5969805  
## [8,] -0.5908870  
## [9,] -0.5969805  
## [10,] -0.2520169  
## [11,] -0.5756534  
## attr(,"scaled:center")  
## fish\_div fish\_den fish\_bio temp c\_bda   
## 1.043875e+00 3.456499e+03 2.160150e+05 2.232323e+01 7.658861e+03   
## c\_enchry p\_furc   
## 5.930354e+04 2.953466e+04   
## attr(,"scaled:scale")  
## fish\_div fish\_den fish\_bio temp c\_bda   
## 2.459291e-01 3.105964e+03 1.395158e+05 7.784982e-01 1.036322e+04   
## c\_enchry p\_furc   
## 8.919395e+04 4.947341e+04

fit3 <- glm((lionfish\_den + 1)~fish\_bio+fish\_den+p\_furc+fish\_div+temp, data=darwin, family = Gamma())  
summary(fit3)

##   
## Call:  
## glm(formula = (lionfish\_den + 1) ~ fish\_bio + fish\_den + p\_furc +   
## fish\_div + temp, family = Gamma(), data = darwin)  
##   
## Deviance Residuals:   
## 1 2 3 4 5 6 7   
## 0.99338 -0.01022 -0.07290 0.21486 1.00862 0.06599 -0.25767   
## 8 9 10 11   
## -2.39063 -2.52108 -2.36179 0.22330   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -1.874e-01 2.059e-01 -0.910 0.404  
## fish\_bio -1.507e-08 3.034e-08 -0.497 0.640  
## fish\_den 1.731e-06 1.923e-06 0.900 0.409  
## p\_furc -1.234e-07 9.613e-08 -1.284 0.255  
## fish\_div -8.316e-04 3.266e-02 -0.025 0.981  
## temp 9.128e-03 9.275e-03 0.984 0.370  
##   
## (Dispersion parameter for Gamma family taken to be 1.350902)  
##   
## Null deviance: 33.732 on 10 degrees of freedom  
## Residual deviance: 19.825 on 5 degrees of freedom  
## AIC: 135.85  
##   
## Number of Fisher Scoring iterations: 6

Testing covairance

library(car)

## Warning: package 'car' was built under R version 3.4.4

## Loading required package: carData

## Warning: package 'carData' was built under R version 3.4.4

##   
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':  
##   
## recode

## The following object is masked from 'package:purrr':  
##   
## some

sqrt(vif(fit3))

## fish\_bio fish\_den p\_furc fish\_div temp   
## 2.933760 3.842427 4.704210 1.189441 2.854775

*using Gamma distribution as this model results in the most “normal” data, see plots below of model fits*

scale((darwin[,2:8]), center = TRUE, scale = TRUE)

## fish\_div fish\_den fish\_bio temp c\_bda c\_enchry  
## [1,] 0.4809800 1.4365108 1.2281635 -0.1297499 0.09054847 2.60244718  
## [2,] -0.0497952 2.0336000 0.5682290 -0.8433746 0.23309789 -0.43401594  
## [3,] -0.1564501 0.8528393 2.0426555 -1.5569993 -0.37682665 0.99732680  
## [4,] 0.2879963 -0.6834913 0.5114883 -0.1297499 -0.73904237 0.08487951  
## [5,] -0.5657100 -0.7103267 -0.5480014 -0.1297499 -0.57546108 -0.64805727  
## [6,] 0.0724305 -0.7639977 -0.8468391 -0.8433703 -0.59882984 -0.48007135  
## [7,] 0.2267893 0.0410664 -0.5568844 -0.1297499 2.35211988 -0.66488296  
## [8,] -1.3837317 -0.1199464 0.1806320 0.5838747 1.36414569 -0.65962493  
## [9,] -0.7697882 -0.6029849 -0.6502334 -0.1297499 -0.73904237 0.10444226  
## [10,] 2.5176680 -0.6029849 -0.7515916 1.2974951 -0.27166725 -0.23756036  
## [11,] -0.6603889 -0.8802847 -1.1776184 2.0111241 -0.73904237 -0.66488296  
## p\_furc  
## [1,] -0.1154827  
## [2,] 2.7300946  
## [3,] 0.4089936  
## [4,] 0.6103919  
## [5,] -0.5969805  
## [6,] -0.4244987  
## [7,] -0.5969805  
## [8,] -0.5908870  
## [9,] -0.5969805  
## [10,] -0.2520169  
## [11,] -0.5756534  
## attr(,"scaled:center")  
## fish\_div fish\_den fish\_bio temp c\_bda   
## 1.043875e+00 3.456499e+03 2.160150e+05 2.232323e+01 7.658861e+03   
## c\_enchry p\_furc   
## 5.930354e+04 2.953466e+04   
## attr(,"scaled:scale")  
## fish\_div fish\_den fish\_bio temp c\_bda   
## 2.459291e-01 3.105964e+03 1.395158e+05 7.784982e-01 1.036322e+04   
## c\_enchry p\_furc   
## 8.919395e+04 4.947341e+04

fit\_bio <- glm((lionfish\_den + 1)~fish\_bio, data=darwin, family = Gamma)  
fit\_den <- glm((lionfish\_den + 1)~fish\_den, data=darwin, family = Gamma)  
fit\_div <- glm((lionfish\_den + 1)~fish\_div, data=darwin,family = Gamma, start = c(0.5, 0.5))

## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced

## Warning: step size truncated due to divergence

## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced

## Warning: step size truncated due to divergence

## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced

## Warning: step size truncated due to divergence

## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced  
  
## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced

## Warning: step size truncated due to divergence

## Warning in log(ifelse(y == 0, 1, y/mu)): NaNs produced

fit\_pfur <- glm((lionfish\_den + 1)~p\_furc, data=darwin,family = Gamma)  
fit\_temp <- glm((lionfish\_den + 1)~temp, data=darwin, family = Gamma)  
AIC(fit\_bio, fit\_den, fit\_pfur, fit\_temp, fit\_div)

## df AIC  
## fit\_bio 3 132.1325  
## fit\_den 3 132.9021  
## fit\_pfur 3 133.0543  
## fit\_temp 3 130.5879  
## fit\_div 3 135.4520

coef(fit\_bio, fit\_den, fit\_pfur, fit\_temp, fit\_div)

## (Intercept) fish\_bio   
## 1.418382e-02 -2.590616e-08

summary(fit\_bio)

##   
## Call:  
## glm(formula = (lionfish\_den + 1) ~ fish\_bio, family = Gamma,   
## data = darwin)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.7729 -1.6877 -0.5795 0.4458 1.7658   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.418e-02 5.939e-03 2.388 0.0407 \*  
## fish\_bio -2.591e-08 1.270e-08 -2.039 0.0718 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Gamma family taken to be 1.485765)  
##   
## Null deviance: 33.732 on 10 degrees of freedom  
## Residual deviance: 26.879 on 9 degrees of freedom  
## AIC: 132.13  
##   
## Number of Fisher Scoring iterations: 7

summary(fit\_den)

##   
## Call:  
## glm(formula = (lionfish\_den + 1) ~ fish\_den, family = Gamma,   
## data = darwin)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.7576 -1.8288 -0.7249 0.3836 1.6406   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.144e-02 4.857e-03 2.355 0.0430 \*  
## fish\_den -1.024e-06 5.547e-07 -1.845 0.0981 .  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Gamma family taken to be 1.50623)  
##   
## Null deviance: 33.732 on 10 degrees of freedom  
## Residual deviance: 28.346 on 9 degrees of freedom  
## AIC: 132.9  
##   
## Number of Fisher Scoring iterations: 7

summary(fit\_div)

##   
## Call:  
## glm(formula = (lionfish\_den + 1) ~ fish\_div, family = Gamma,   
## data = darwin, start = c(0.5, 0.5))  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.93731 -2.04226 -0.38759 0.06849 1.69945   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 0.003937 0.010800 0.365 0.724  
## fish\_div 0.001435 0.010377 0.138 0.893  
##   
## (Dispersion parameter for Gamma family taken to be 1.923602)  
##   
## Null deviance: 33.732 on 10 degrees of freedom  
## Residual deviance: 33.694 on 9 degrees of freedom  
## AIC: 135.45  
##   
## Number of Fisher Scoring iterations: 8

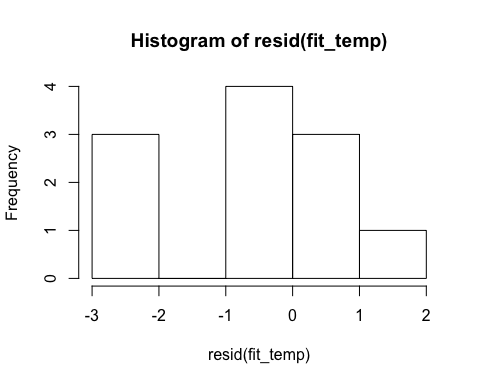
summary(fit\_pfur)

##   
## Call:  
## glm(formula = (lionfish\_den + 1) ~ p\_furc, family = Gamma, data = darwin)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.7662 -1.8014 -0.2507 0.2023 2.0405   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 8.872e-03 3.963e-03 2.239 0.0519 .  
## p\_furc -4.612e-08 2.690e-08 -1.714 0.1206   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Gamma family taken to be 1.917667)  
##   
## Null deviance: 33.732 on 10 degrees of freedom  
## Residual deviance: 28.644 on 9 degrees of freedom  
## AIC: 133.05  
##   
## Number of Fisher Scoring iterations: 7

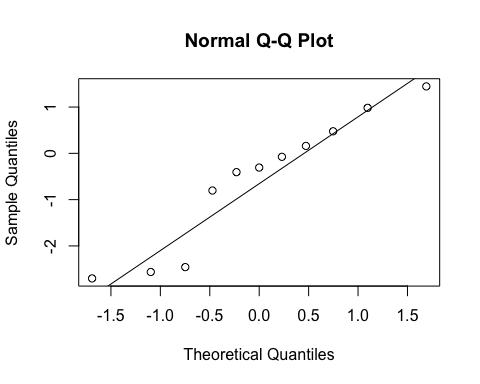
summary(fit\_temp)

##   
## Call:  
## glm(formula = (lionfish\_den + 1) ~ temp, family = Gamma, data = darwin)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.7016 -1.6302 -0.3070 0.3184 1.4462   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.162890 0.066513 -2.449 0.0368 \*  
## temp 0.007765 0.003134 2.478 0.0351 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Gamma family taken to be 1.165115)  
##   
## Null deviance: 33.732 on 10 degrees of freedom  
## Residual deviance: 24.127 on 9 degrees of freedom  
## AIC: 130.59  
##   
## Number of Fisher Scoring iterations: 6

hist(resid(fit\_temp))



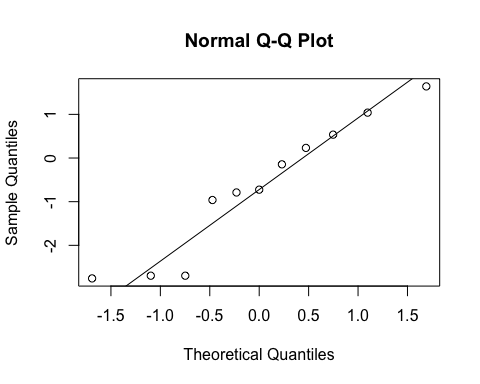
qqnorm(resid(fit\_temp))   
qqline(resid(fit\_temp))



hist(resid(fit\_den))



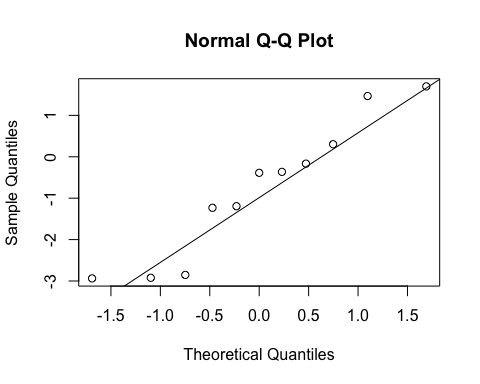
qqnorm(resid(fit\_den))  
qqline(resid(fit\_den))



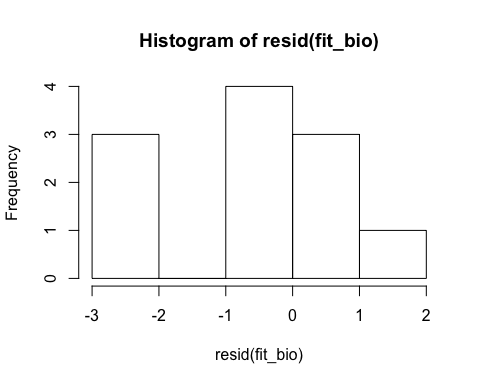
hist(resid(fit\_div))



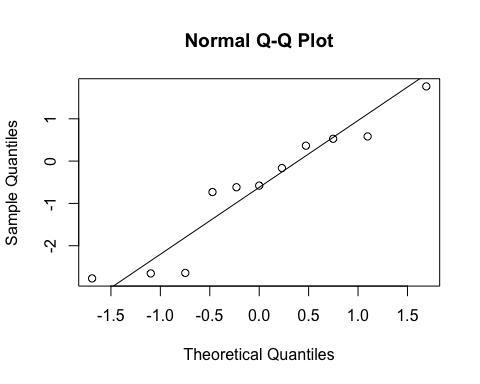
qqnorm(resid(fit\_div))  
qqline(resid(fit\_div))



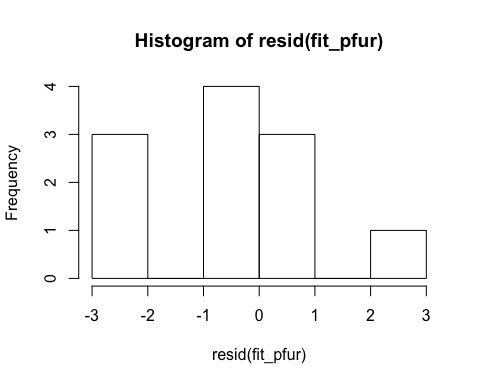
hist(resid(fit\_bio))



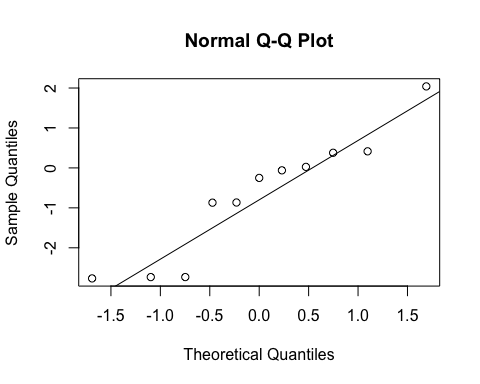
qqnorm(resid(fit\_bio))  
qqline(resid(fit\_bio))



hist(resid(fit\_pfur))



qqnorm(resid(fit\_pfur))  
qqline(resid(fit\_pfur))



coef(fit\_bio)

## (Intercept) fish\_bio   
## 1.418382e-02 -2.590616e-08

coef(fit\_pfur)

## (Intercept) p\_furc   
## 8.871575e-03 -4.611608e-08

coef(fit\_temp)

## (Intercept) temp   
## -0.162889556 0.007764716

coef(fit\_div)

## (Intercept) fish\_div   
## 0.003937398 0.001435102

coef(fit\_den)

## (Intercept) fish\_den   
## 1.143692e-02 -1.023463e-06