

Cod

2024-07-17

```
library(dplyr)

## Warning: package 'dplyr' was built under R version 4.3.3

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 4.3.3

library(bbmle)

## Warning: package 'bbmle' was built under R version 4.3.3

## Loading required package: stats4

##
## Attaching package: 'bbmle'

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

load("C:/Users/lucab/Downloads/stomach_dataset.Rdata")

sprat <- stom_df %>% filter(pred_taxa == "Gadus morhua")

ggplot(sprat, aes(x=pred_weight_g, y=log(ppmr))) +
  geom_point() +
  geom_smooth(method="gam", se=FALSE, aes(weight = nprey_perpred)) +
  labs(title="logPPMR vs Predator Weight", x="Predator Weight (g)", y="logPPMR")
```

```

## `geom_smooth()` using formula = 'y ~ s(x, bs = "cs")'

## Warning: Removed 177 rows containing non-finite outside the scale range
## ('stat_smooth()').

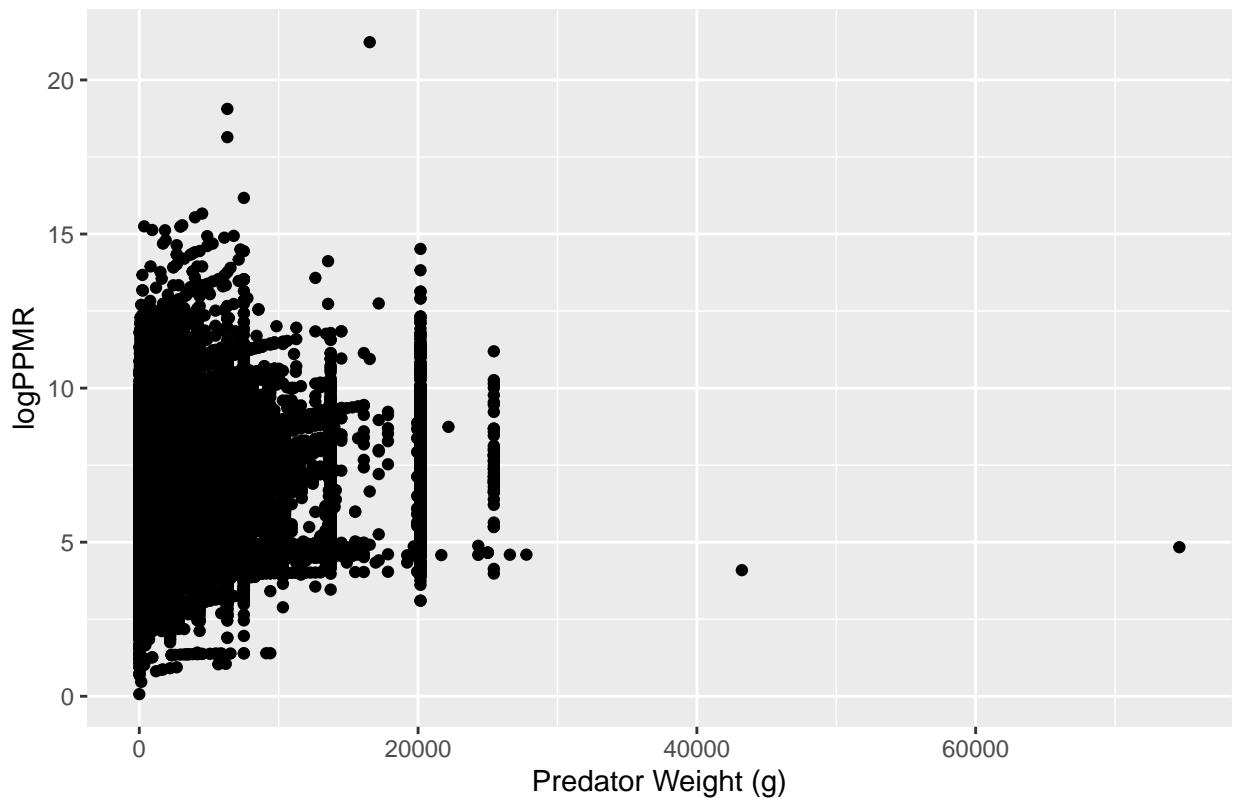
## Warning in sqrt(w): NaNs produced

## Warning: Failed to fit group -1.
## Caused by error in 'while (mean(ldxx / (ldxx + ldss)) > 0.4) {
##   def.sp <- def.sp * 10
##   ldss <- ldss * 10
## }':
## ! missing value where TRUE/FALSE needed

## Warning: Removed 177 rows containing missing values or values outside the scale range
## ('geom_point()').

```

logPPMR vs Predator Weight



```

dig <- 1

ggplot(sprat, aes(pred_weight_g, y=log(ppmr)))+
  geom_point()+
  facet_wrap(~data)+
  geom_smooth(method="gam", se=FALSE, aes(weight = nprey_perpred*sprat$prey_ind_weight_g^dig))+ 
  labs(title="logPPMR vs Predator Weight", x="Predator Weight (g)", y="logPPMR")

```

```

## Warning: Use of 'sprat$prey_ind_weight_g' is discouraged.
## i Use 'prey_ind_weight_g' instead.

## `geom_smooth()` using formula = 'y ~ s(x, bs = "cs")'

## Warning: Removed 177 rows containing non-finite outside the scale range
## ('stat_smooth()').

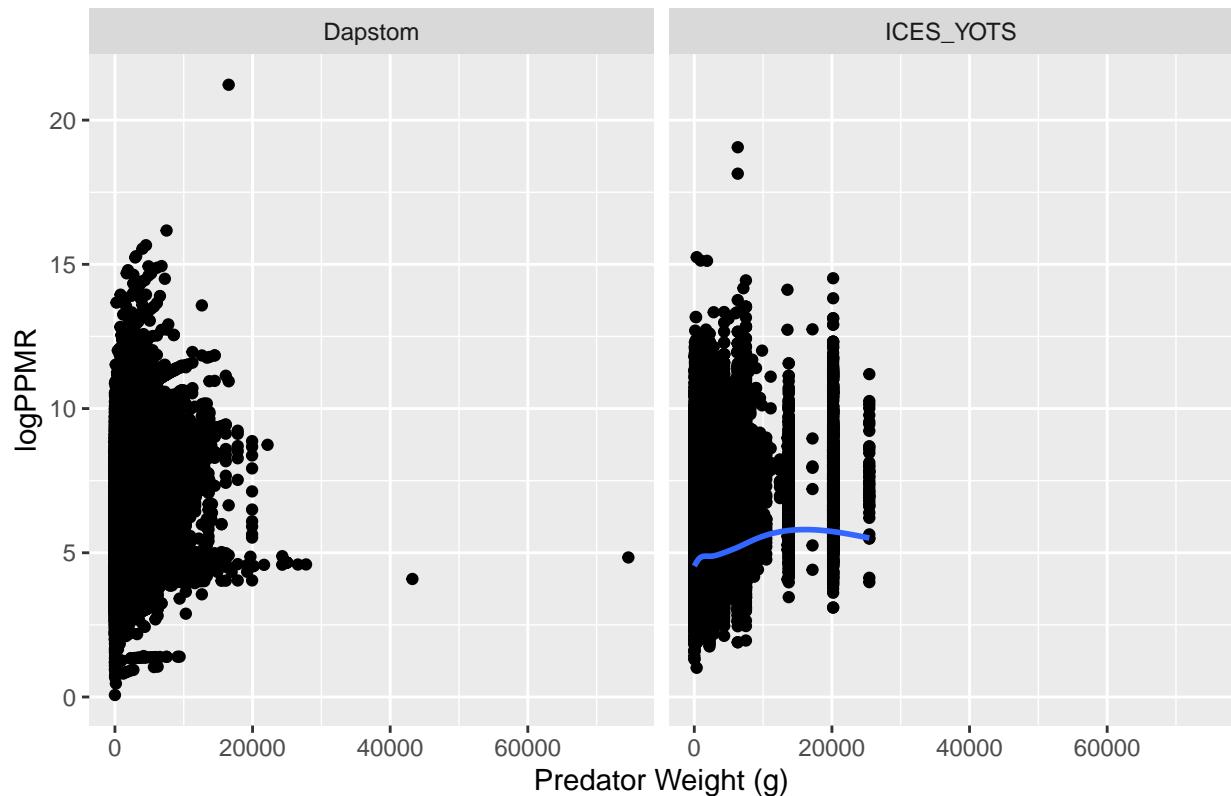
## Warning in sqrt(w): NaNs produced

## Warning: Failed to fit group -1.
## Caused by error in 'while (mean(ldxx / (ldxx + ldss)) > 0.4) {
##     def.sp <- def.sp * 10
##     ldss <- ldss * 10
## }':
## ! missing value where TRUE/FALSE needed

## Warning: Removed 177 rows containing missing values or values outside the scale range
## ('geom_point()').

```

logPPMR vs Predator Weight



The PPMR does seem to increase at the largest sizes, should I remove or just ignore?

```

sprat$weight_numbers <- sprat$nprey_perpred
sprat$weight_biomass <- sprat$nprey_perpred*sprat$prey_ind_weight_g^dig

ggplot() +
  geom_density(data = sprat, aes(x = log(ppmr), weight = weight_biomass), color = "red") +
  geom_density(data = sprat, aes(x = log(ppmr), weight = weight_numbers), color = "blue") +
  labs(title = "Density Plot of log(ppmr) Weighted by Biomass and Numbers",
       x = "log(ppmr)",
       y = "Density") +
  theme_minimal()

## Warning: Removed 177 rows containing non-finite outside the scale range
## ('stat_density()').

## Warning: Computation failed in 'stat_density()' .
## Caused by error in 'density.default()' :
## ! 'weights' must not be negative

## Warning: Removed 177 rows containing non-finite outside the scale range
## ('stat_density()').

## Warning: Computation failed in 'stat_density()' .
## Caused by error in 'density.default()' :
## ! 'weights' must not be negative

```

Density Plot of log(ppmr) Weighted by Biomass and Numbers

Density

log(ppmr)

```

sprat <- sprat%>%filter(nprey_perpred>0)

library(mclust)

## Warning: package 'mclust' was built under R version 4.3.3

## Package 'mclust' version 6.1.1
## Type 'citation("mclust")' for citing this R package in publications.

repeat_elements <- function(data, weights) {

  valid_indices <- !is.na(data) & !is.na(weights)
  data <- data[valid_indices]
  weights <- weights[valid_indices]

  final_vector <- c()

  for (i in seq_along(data)) {

    rounded_weight <- round(weights[i])

    repeated_values <- rep(data[i], times = rounded_weight)

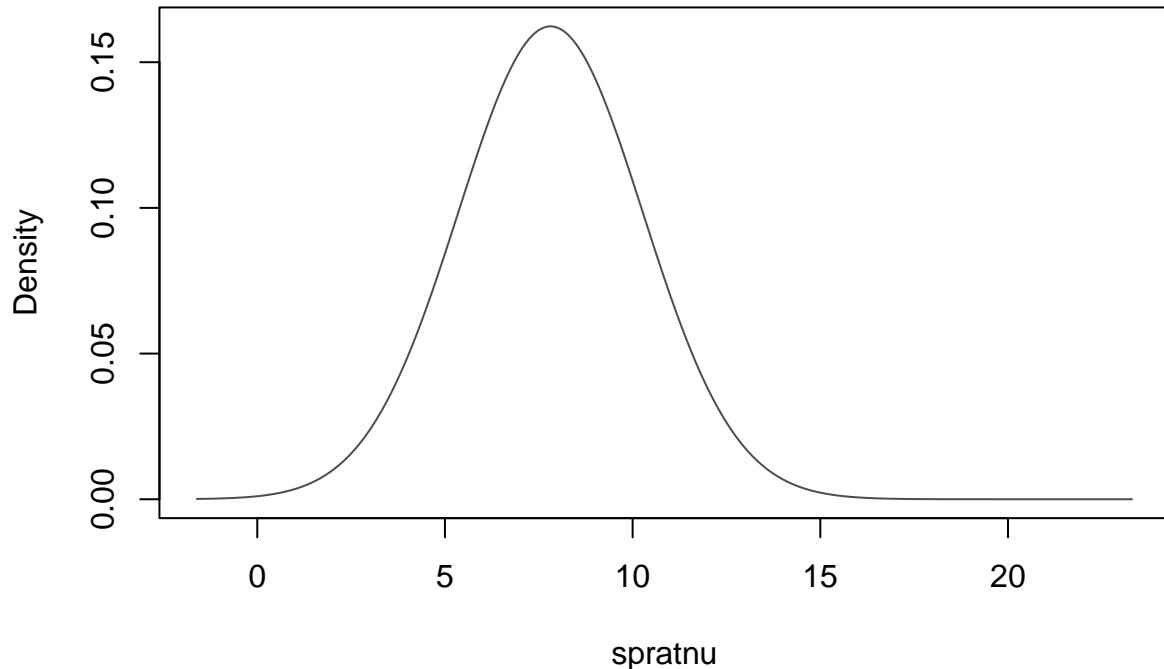
    final_vector <- c(final_vector, repeated_values)
  }

  return(final_vector)
}

spratnu <- repeat_elements(log(sprat$ppmr), sprat$nprey_perpred)

gmm <- densityMclust(spratnu, G=1)

```

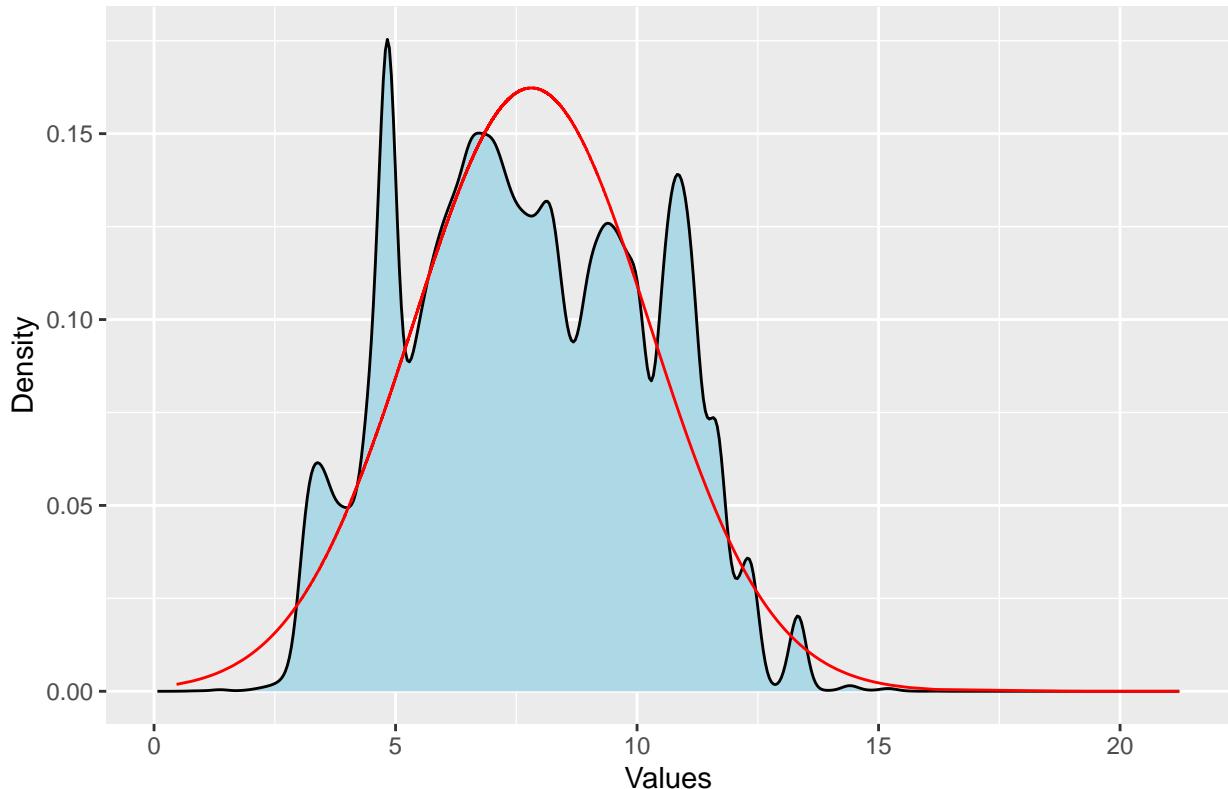


```
dplot <- data.frame(x=gmm[["data"]], density=gmm[["density"]])

(numbfit <- ggplot() +
  geom_density(data=sprat, aes(log(ppmr), weight=weight_numbers), fill="lightblue") +
  geom_line(data=dplot, aes(x = x, y = density), color="red") +
  labs(x = "Values", y = "Density") +
  ggtitle("Number Density Plot from Number Distribution"))

## Warning: Removed 173 rows containing non-finite outside the scale range
## ('stat_density()'').
```

Number Density Plot from Number Distribution



Ok this gaussian fits nicely, does it shift?

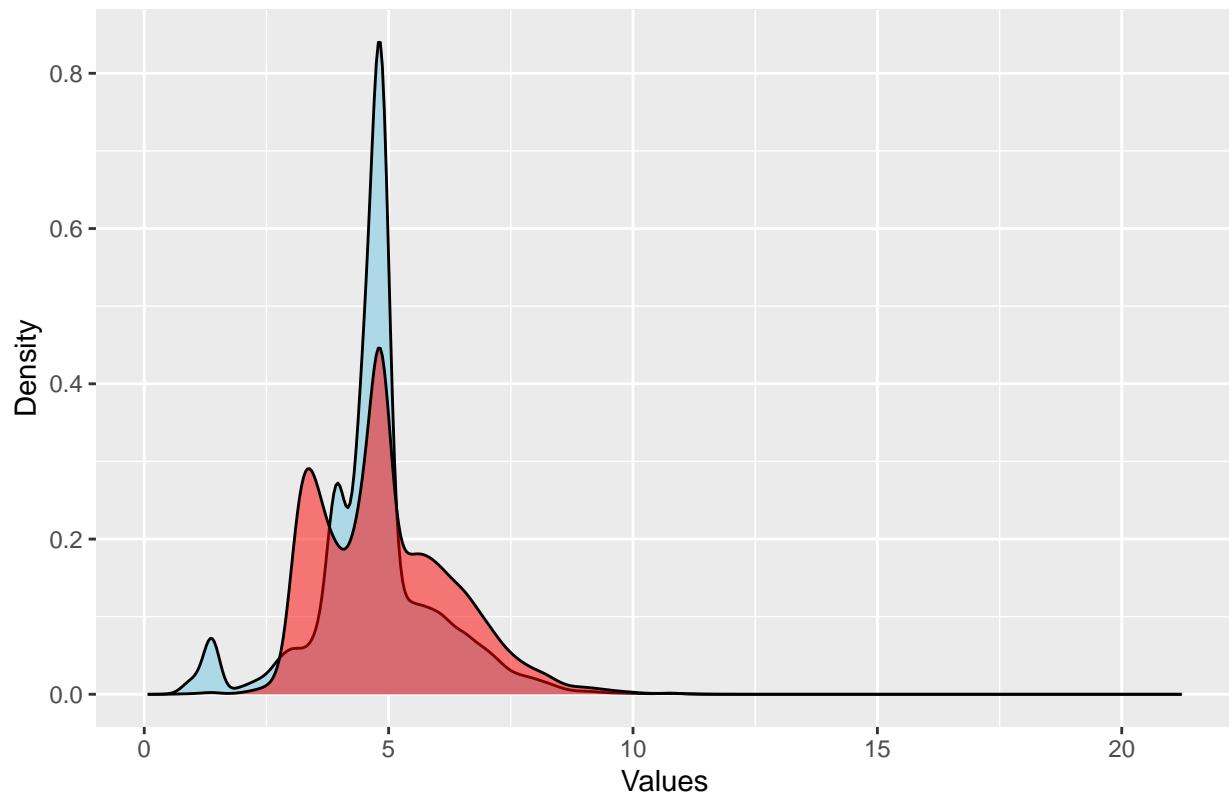
```
newspratdens <- gmm[["density"]]*exp(-dig*gmm[["data"]])
shifted_pdf_normalized <- newspratdens / sum(newspratdens)

#making new dataframe
dplot <- data.frame(x=gmm[["data"]], density=shifted_pdf_normalized)

(numbfitbio <- ggplot() +
  geom_density(data=sprat, aes(log(ppmr), weight=weight_biomass), fill="lightblue")+
  geom_density(data=dplot, aes(x, weight=shifted_pdf_normalized), fill="red", alpha=0.5) +
  #geom_line(data=dplot, aes(x = x, y = density), color="red") +
  labs(x = "Values", y = "Density") +
  ggtitle("Diet Density Plot from Number Distribution"))

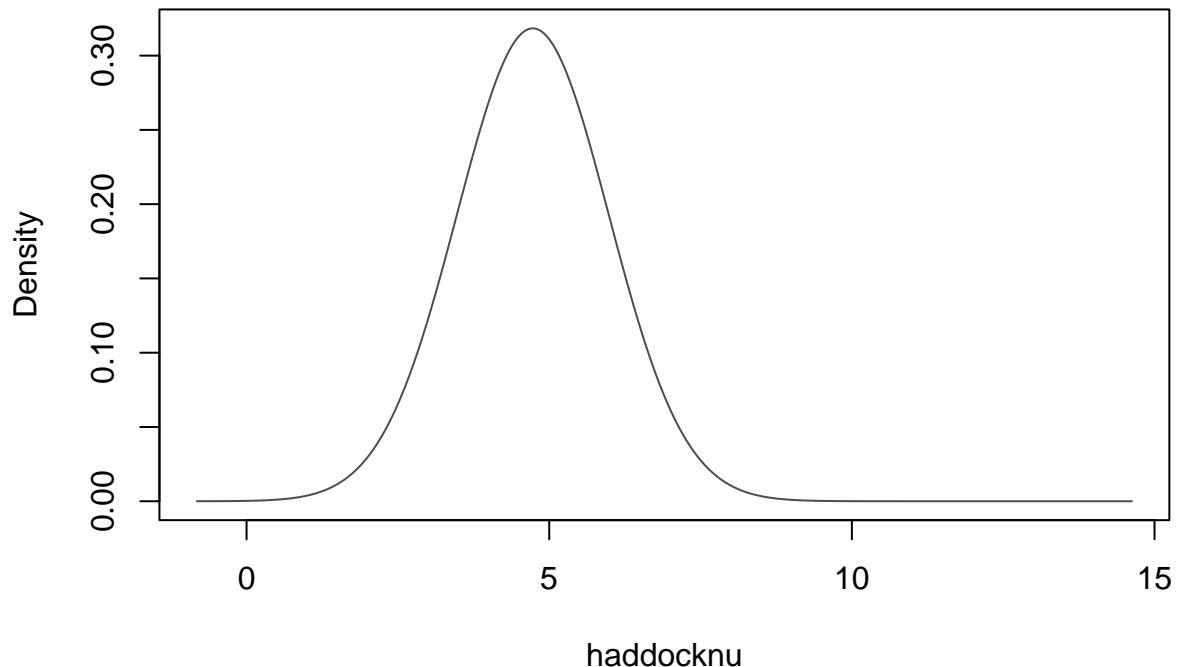
## Warning: Removed 173 rows containing non-finite outside the scale range
## ('stat_density()'').
```

Diet Density Plot from Number Distribution



This isn't the best shift, let's try to shift it the other way.

```
stomach <- sprat  
  
haddocknu <- repeat_elements(log(stomach$ppmr), stomach$nprey_perpred *  
                           stomach$prey_ind_weight_g^dig)  
gmm <- densityMclust(haddocknu, G=1)
```

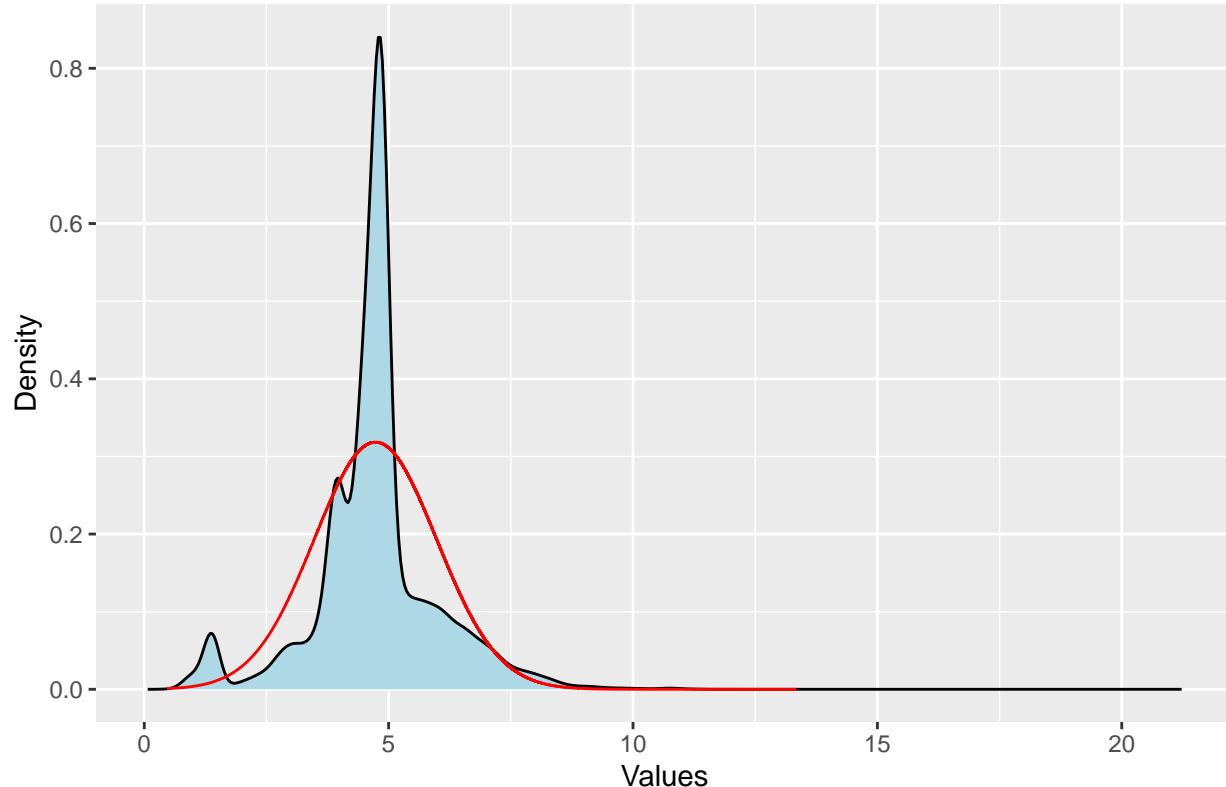


```
dplot <- data.frame(x=gmm[["data"]], density=gmm[["density"]])

(biofit <- ggplot() +
  geom_density(data=stomach, aes(log(ppmr), weight=weight_biomass), fill="lightblue")+
  geom_line(data=dplot, aes(x = x, y = density), color="red") +
  labs(x = "Values", y = "Density") +
  ggtitle("Diet Density Plot from Diet Distribution"))
```

```
## Warning: Removed 173 rows containing non-finite outside the scale range
## ('stat_density()'').
```

Diet Density Plot from Diet Distribution



```

shifted_fit <- gmm
shifted_fit[["parameters"]][["mean"]] <- shifted_fit[["parameters"]][["mean"]]+
  (1)*shifted_fit[["parameters"]][["variance"]][["sigmasq"]]

stomach <- stomach%>%filter(!is.na(ppmr))

x_vals <- seq(min(log(stomach$ppmr)), max(log(stomach$ppmr)), length.out = 1000)

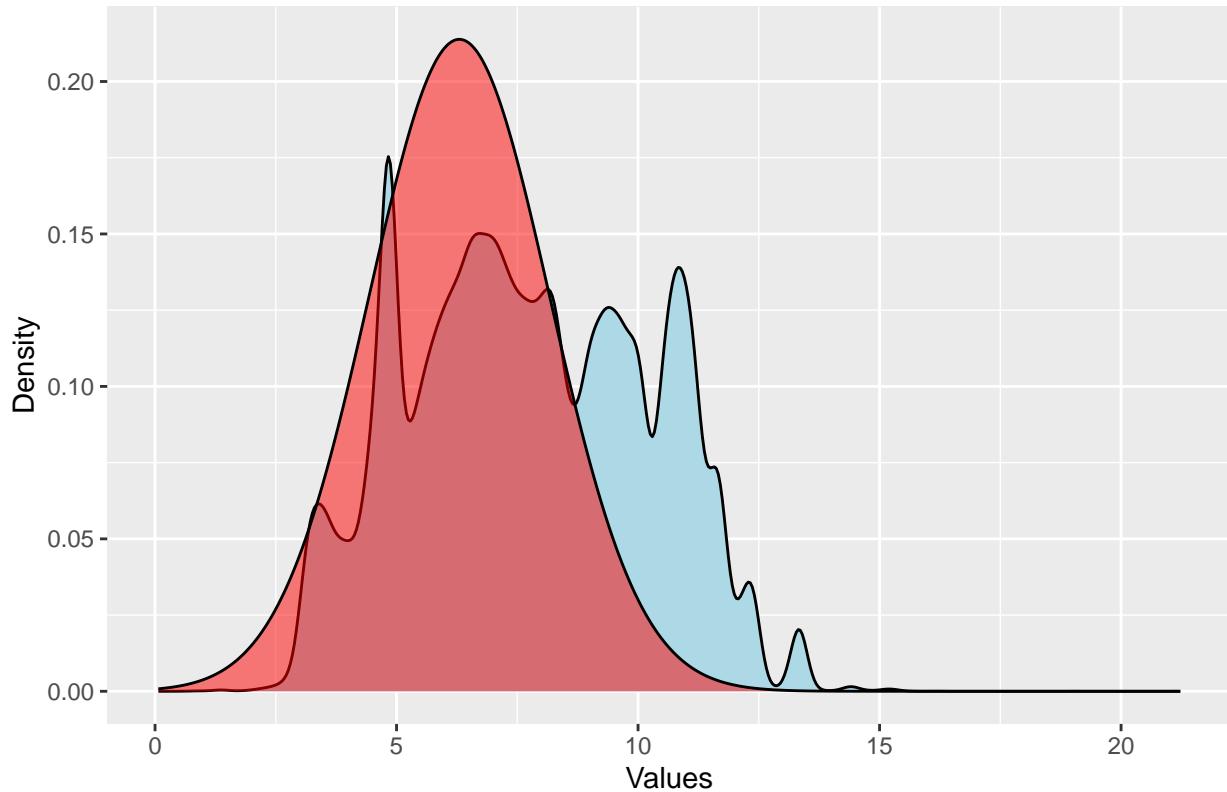
#generating the density values
shifted_pdf <- sapply(x_vals, function(x) {
  sum(shifted_fit$parameters$pro * dnorm(x, mean = shifted_fit$parameters$mean, sd = sqrt(shifted_fit$parameters$variance)))
})

plot_data <- data.frame(x = x_vals, shifted_pdf = shifted_pdf)

(numbfitbio <- ggplot() +
  geom_density(data=stomach, aes(log(ppmr), weight=weight_numbers), fill="lightblue")+
  geom_density(data=plot_data, aes(x, weight=shifted_pdf), fill="red", alpha=0.5)+
  #geom_line(data=dplot, aes(x = x, y = density), color="red") +
  labs(x = "Values", y = "Density") +
  ggtitle("Diet Density Plot from Number Distribution"))

```

Diet Density Plot from Number Distribution



#different way

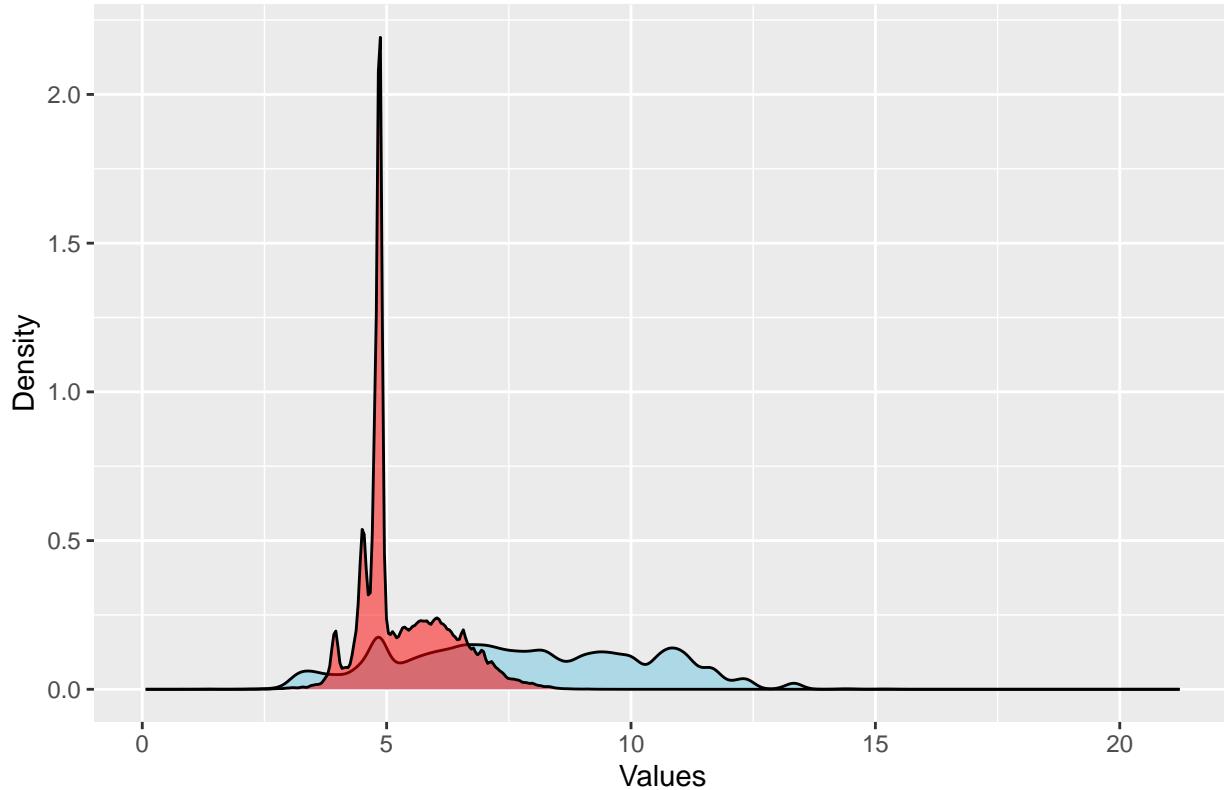
```
newspratdens <- gmm[["density"]]*exp(dig*gmm[["data"]])

shifted_pdf_normalized <- newspratdens / sum(newspratdens)

dplot <- data.frame(x=gmm[["data"]], density=shifted_pdf_normalized)

(numbfitbio <- ggplot() +
  geom_density(data=stomach, aes(log(ppmr), weight=weight_numbers), fill="lightblue")+
  geom_density(data=dplot, aes(x, weight=shifted_pdf_normalized), fill="red", alpha=0.5) +
  #geom_line(data=dplot, aes(x = x, y = density), color="red") +
  labs(x = "Values", y = "Density") +
  ggtitle("Diet Density Plot from Number Distribution"))
```

Diet Density Plot from Number Distribution



Okay, it doesn't shift well either here. I will try the exponential.

```

fl <- function(l, alpha, ll, ul, lr, ur) {
  dl <- ll - l
  dr <- l - lr
  fl_values <- exp(alpha * l) / (1 + exp(ul * dl)) / (1 + exp(ur * dr))

  # Debugging output
  if (any(!is.finite(fl_values))) {
    print("Non-finite fl values found")
    print(fl_values)
  }

  return(fl_values)
}

## Define the truncated exponential PDF with debugging
dteexp <- function(l, alpha, ll, ul, lr, ur) {
  fl_values <- fl(l, alpha, ll, ul, lr, ur)

  integral_result <- tryCatch(
    integrate(fl, 0, 30, alpha = alpha, ll = ll, ul = ul, lr = lr, ur = ur),
    error = function(e) {
      print("Integration failed")
      print(e)
      return(NULL)
  })
}
```

```

        }

    }

    if (is.null(integral_result)) {
      return(rep(NA, length(l)))
    }

    d <- fl_values / integral_result$value

    # Debugging output
    if (any(!is.finite(d))) {
      print("Non-finite d values found")
      print(d)
    }

    return(d)
}

#Define the MLE function with debugging
mle_texp <- function(df) {
  loglik <- function(alpha, ll, ul, lr, ur) {
    L <- dtexp(df$l, alpha, ll, ul, lr, ur)

    # Debugging output
    if (any(!is.finite(L) | L <= 0)) {
      print("Non-finite or non-positive L values found")
      print(which(!is.finite(L) | L <= 0))
      return(Inf)
    }

    -sum(log(L) * df$weight_numbers)
  }

  result <- tryCatch(
    mle2(loglik, start = list(
      alpha = 0.5,
      ll = min(df$l),
      lr = max(df$l),
      ul = 5,
      ur = 5
    ), method = "L-BFGS-B", control = list(maxit = 10000)),
    error = function(e) {
      print("MLE fitting failed")
      print(e)
      return(NULL)
    }
  )

  return(result)
}

```

```
library(bbmle)
```

```

stomach <- stomach%>%mutate(l=log(ppmr))%>%filter(!is.na(l))

est <- mle_texp(stomach)

biomassco <- est@coef

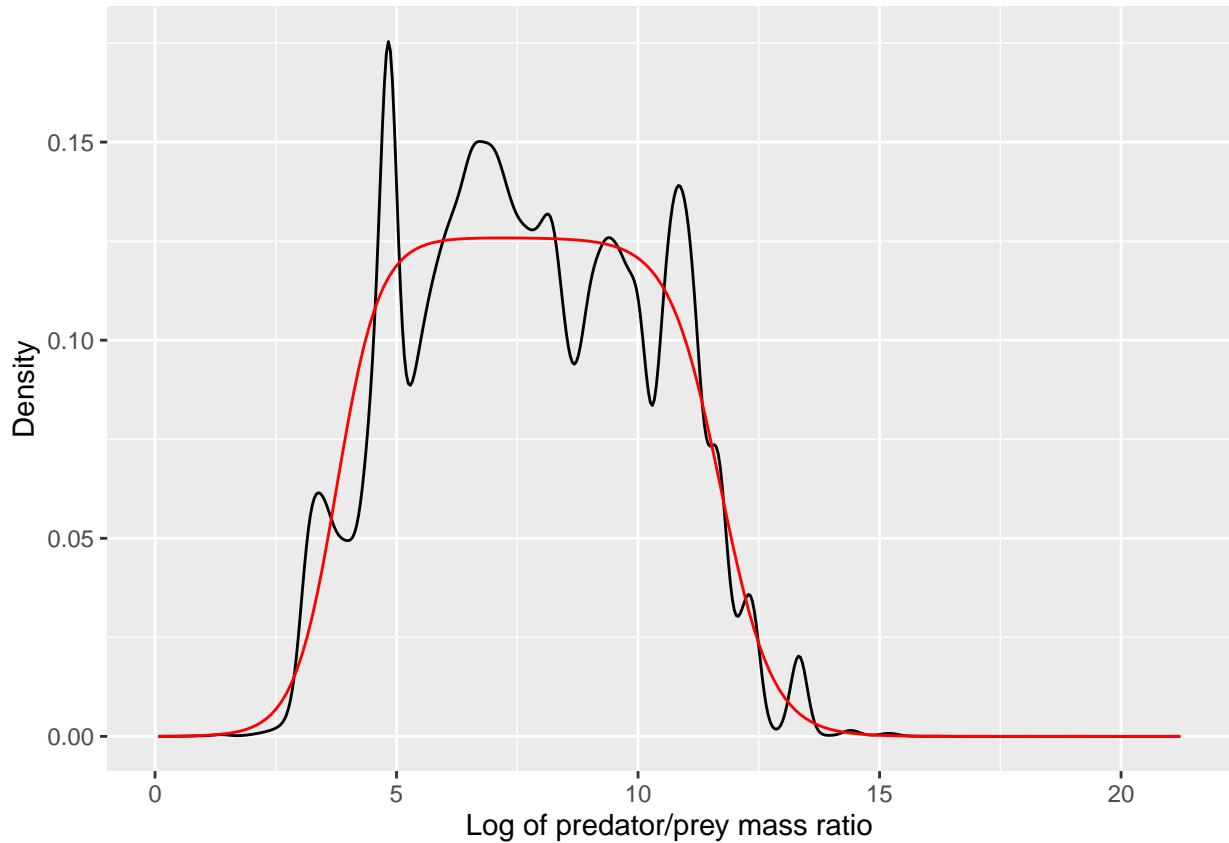
grid = seq(min(stomach$l), max(stomach$l), length.out = 200)

dist <- dtexp(grid, alpha = (biomassco[1]), ll = biomassco[2], ul = biomassco[3],
              lr = biomassco[4], ur = biomassco[5])

dist <- data.frame(l=grid, Density=dist)

ggplot(stomach) +
  geom_density(aes(l, weight=weight_numbers))+
  xlab("Log of predator/prey mass ratio") +
  geom_line(aes(l, Density), data = dist, color = "red")

```



Fits okay, does it shift.

```

numberestco <- biomassco

dist <- dtexp(grid, alpha = (numberestco[1]), ll = numberestco[2], ul = numberestco[3], lr = numberestco[4], ur = numberestco[5])
numberdist <- data.frame(l=grid, Density=dist)
dist <- dtexp(grid, alpha = (numberestco[1]-1), ll = numberestco[2], ul = numberestco[3], lr = numberestco[4], ur = numberestco[5])
numberdist <- data.frame(l=grid, Density=dist)

```

```

biomassdist <- data.frame(l=grid, Density=dist)
#now plot these two together

stomach <- stomach %>% mutate(biomass = nprey_perpred * prey_ind_weight_g)

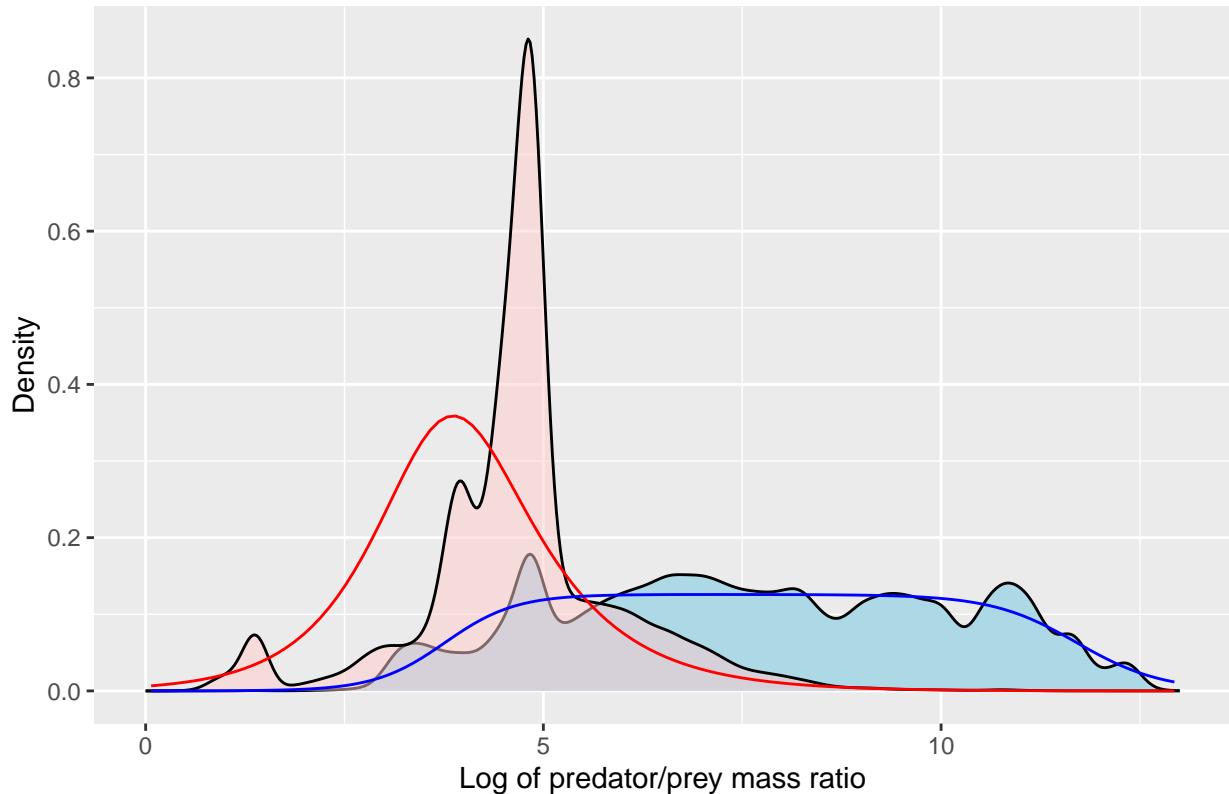
ggplot(stomach) +
  geom_density(aes(l, weight=weight_numbers), fill="lightblue")+
  geom_density(aes(l, weight=biomass), fill="#ffcccb", alpha=0.5)+ 
  xlab("Log of predator/prey mass ratio") +
  geom_line(aes(l, Density), data = biomassdist, color = "red")+
  geom_line(aes(l, Density), data = numberdist, color = "blue")+
  xlim(0,13) +
  ggtitle("Fit to number, shift to diet")

## Warning: Removed 80 rows containing non-finite outside the scale range
## ('stat_density()').
## Removed 80 rows containing non-finite outside the scale range
## ('stat_density()').

## Warning: Removed 78 rows containing missing values or values outside the scale range
## ('geom_line()').
## Removed 78 rows containing missing values or values outside the scale range
## ('geom_line()').

```

Fit to number, shift to diet



This doesn't fit well. Its okay, but it is still off.

Will try it the other way around and compare. (fitting number from diet)

```
stomach <- stomach%>%mutate(weight_numbers=nprey_perpred*stomach$prey_ind_weight_g^dig)

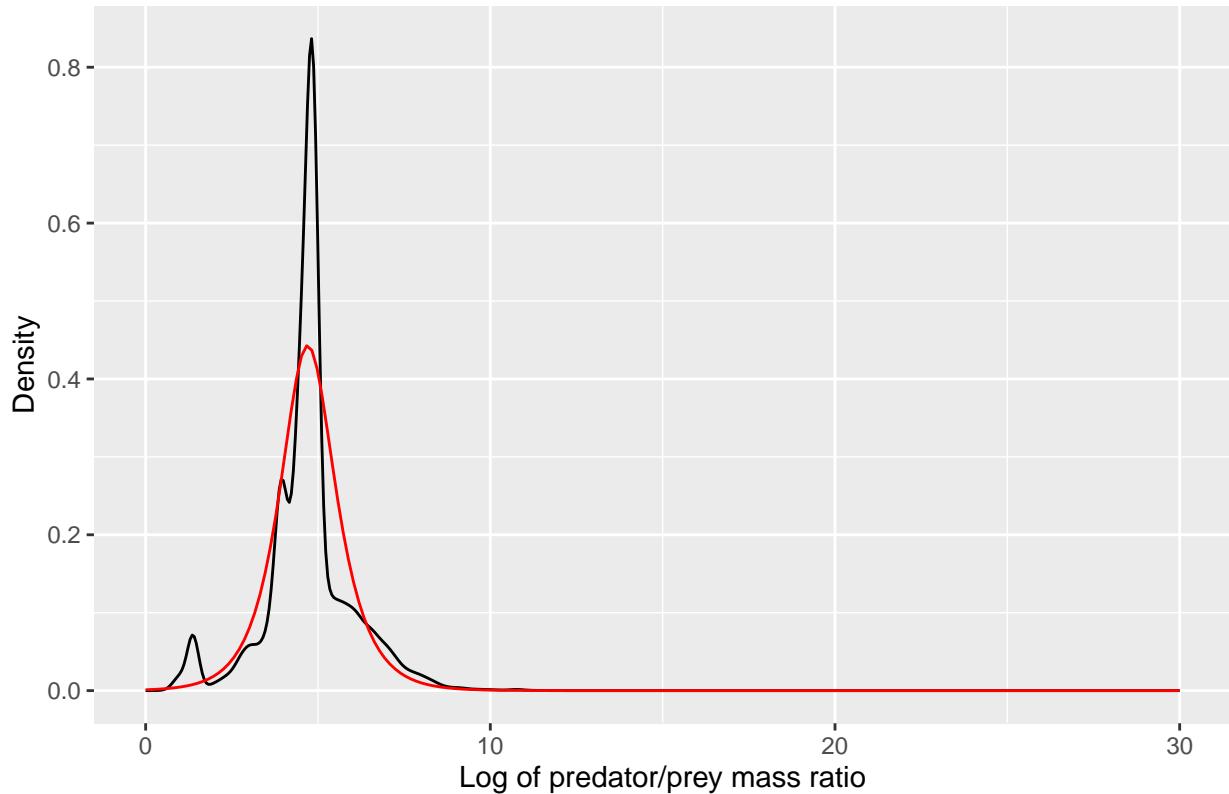
est <- mle_texp(stomach)

estco <- est@fullcoef
grid = seq(0, 30, length.out = 200)
dist <- dtexp(grid, alpha = estco[1], ll = estco[2], ul = estco[3], lr = estco[4], ur = estco[5])
dist <- data.frame(l=grid, Density=dist)

ggplot(stomach) +
  geom_density(aes(l, weight=nprey_perpred*stomach$prey_ind_weight_g^dig))+
  xlab("Log of predator/prey mass ratio") +
  geom_line(aes(l, Density), data = dist, color = "red")+
  ggtitle("fit to diet")

## Warning: Use of 'stomach$prey_ind_weight_g' is discouraged.
## i Use 'prey_ind_weight_g' instead.
```

fit to diet



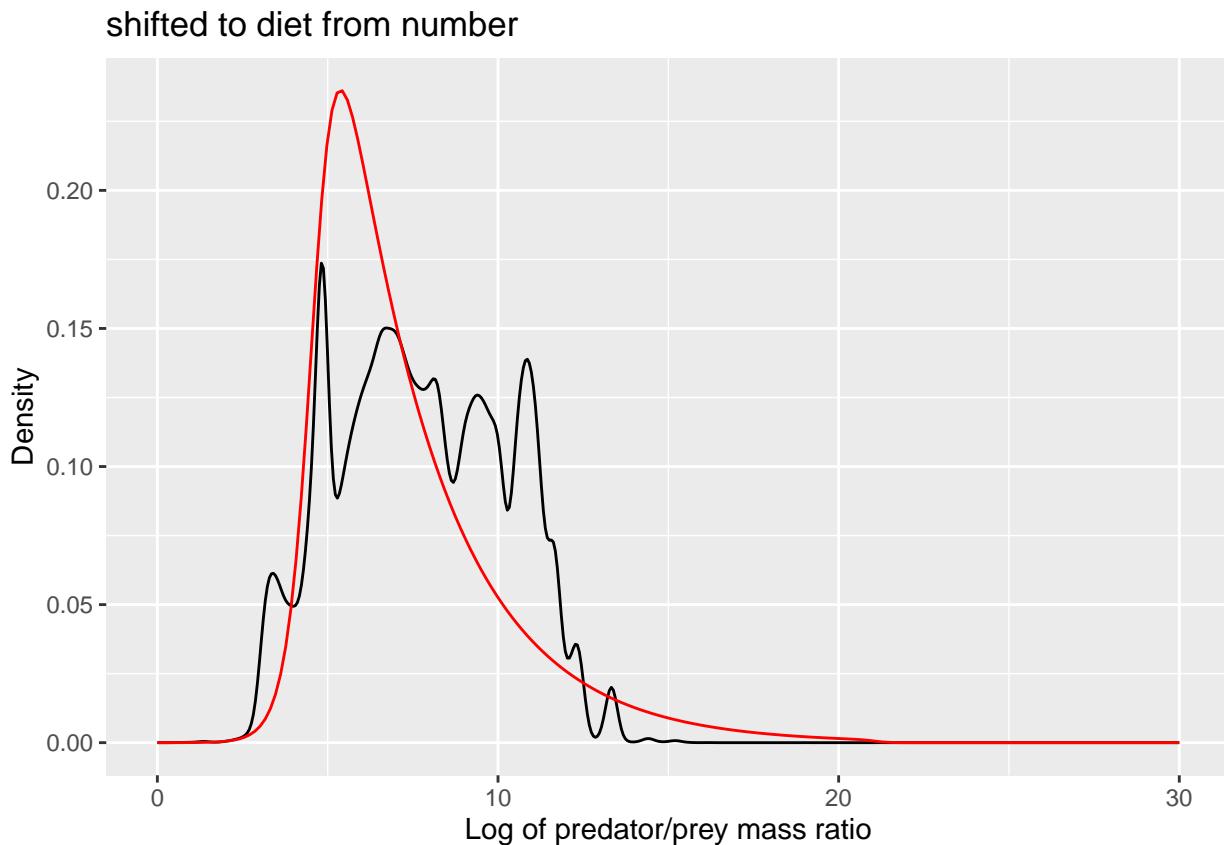
```
dist <- dtexp(grid, alpha = estco[1]+1, ll = estco[2], ul = estco[3], lr = estco[4], ur = estco[5])
dist <- data.frame(l=grid, Density=dist)

ggplot(stomach) +
  geom_density(aes(l, weight=nprey_perpred))+
```

```

xlab("Log of predator/prey mass ratio") +
  geom_line(aes(1, Density), data = dist, color = "red")+
  ggtitle("shifted to diet from number")

```



It doesn't fit well for both.

But shifting across number to diet works the best I think.

```
print(numberestco)
```

```

##           alpha          ll          ul          lr          ur
## -0.0005523561  3.7634229256  2.2665059865 11.7093958978  1.8644950972

```