

The **mixdist** package – MIX for the R Environment

Ichthus Data Systems is excited to release a new version of MIX for the R environment. The R package **mixdist** has most of the functionality of MIX but with greatly improved numerical methods and graphics. The list of distributions that can be fitted now includes Binomial, Poisson and Negative Binomial. R is truly cross-platform and will run in exactly the same way on Windows, Macintosh, Unix or Linux, or any platform to which the R environment has been ported. R provides convenient tools to organize and prepare your data, and the results of **mixdist** calculations are R objects which can be used as arguments to other R functions for further analysis.

The R environment is open code and can be downloaded at no cost from <http://cran.us.r-project.org>. Additional resources for learning R can be found on my McMaster University course web sites, for example http://www.math.mcmaster.ca/peter/s3n03/s3n03_0506.

The **mixdist** functions are fully documented online, with examples that can be run from the documentation pages. The examples use the same data sets as in the MIX documentation. The options for truncated data and the optimal design calculation are not yet programmed in **mixdist**.

Links for downloading **mixdist**, and installation instructions, are found at <http://www.math.mcmaster.ca/peter/mix/mix.html>.

To use Rmix, launch an R session and attach the **mixdist** library. To get help on a particular function, enter, e.g. **?mixparam**. To open the html help system in your browser, type **help.start()**. To get one of the sample data sets, type, e.g. **data(cassie)** then coerce it into a mixdata object with **cassie <- as.mixdata(cassie)**.

Here is a sample session fitting a mixture of gamma distributions to the cassie data, first with constant coefficient of variation, then unconstrained.

```
R : Copyright 2003, The R Development Core Team
Version 1.6.2 (2003-01-10)
```

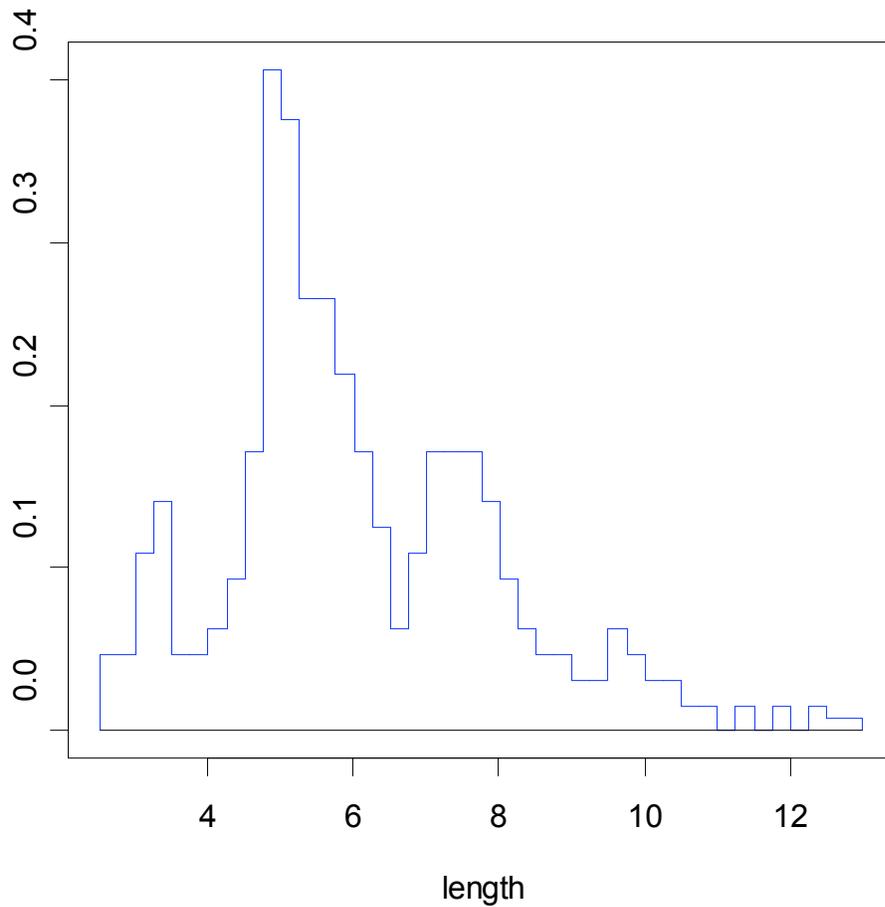
```
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type `license()' or `licence()' for distribution details.
```

```
R is a collaborative project with many contributors.
Type `contributors()' for more information.
```

```
Type `demo()' for some demos, `help()' for on-line help, or
`help.start()' for a HTML browser interface to help.
Type `q()' to quit R.
```

```
[Previously saved workspace restored]
```

```
> help.start()
updating HTML package listing
> library(mixdist)
> data(cassie)
> cassie <- as.mixdata(cassie)
> plot(cassie)
```



```
> fitcass1 <- mix(cassie,mixparam(c(3,5,7,9,11),.5),"gamma",mixconstr(consigma="CCV"))
> summary(fitcass1)
```

Parameters:

	pi	mu	sigma
1	0.10819	3.277	0.3363
2	0.50550	5.243	0.5380
3	0.29085	7.396	0.7589
4	0.08331	9.682	0.9936
5	0.01215	12.126	1.2443

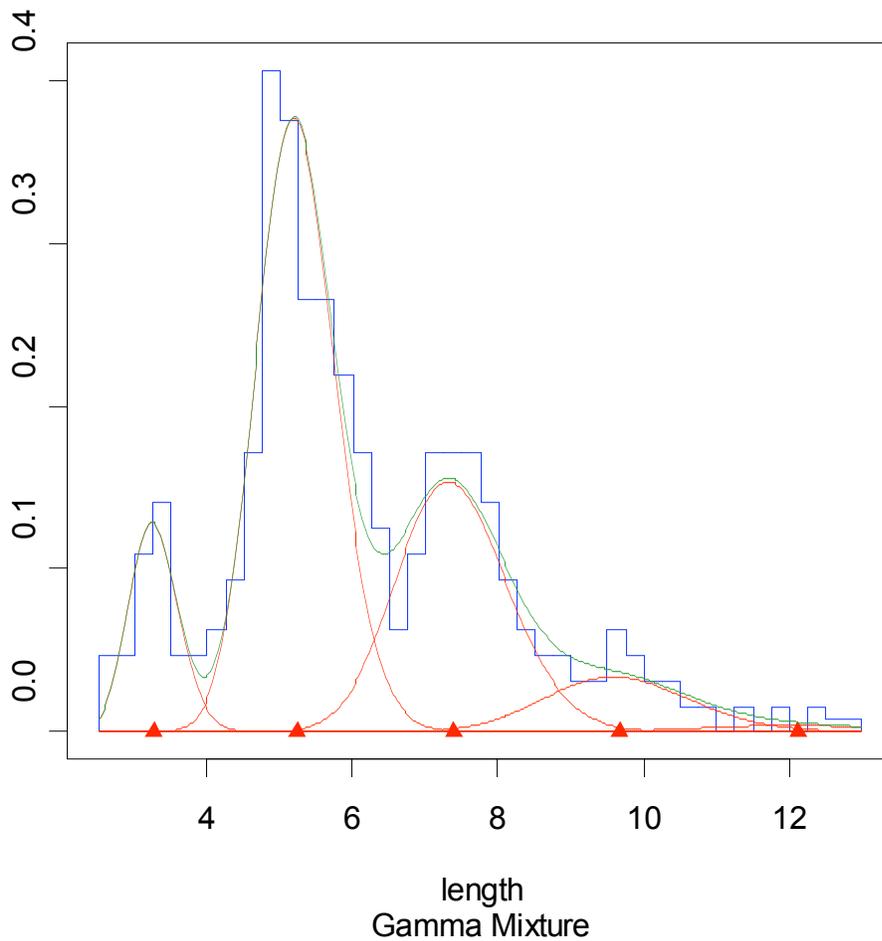
Standard Errors:

	pi.se	mu.se	sigma.se
1	0.02022	0.07738	0.02277
2	0.03580	0.06081	NA
3	0.03523	0.15335	NA
4	0.02508	0.47385	NA
5	0.01399	1.47106	NA

Analysis of Variance Table

	Df	Chisq	Pr(>Chisq)
Residuals	29	16.515	0.9692

```
> plot(fitcass1)
```



```
> fitcass2 <- mix(cassie,coef(fitcass1),"gamma")
```

Warning message:

The optimization process terminated because iteration limit exceeded in: mix(cassie, coef(fitcass1), "gamma")

```
> fitcass2 <- mix(cassie,coef(fitcass1),"gamma",iterlim=150)
```

Warning message:

The optimization process terminated because either the estimates are approximate local optimal solution or steptol is too small in: mix(cassie, coef(fitcass1), "gamma", iterlim = 150)

```
> summary(fitcass2)
```

Parameters:

	pi	mu	sigma
1	0.10313	3.248	0.3101
2	0.55401	5.317	0.6256
3	0.24471	7.536	0.5482
4	0.08462	9.669	0.7225
5	0.01353	12.225	0.5687

Standard Errors:

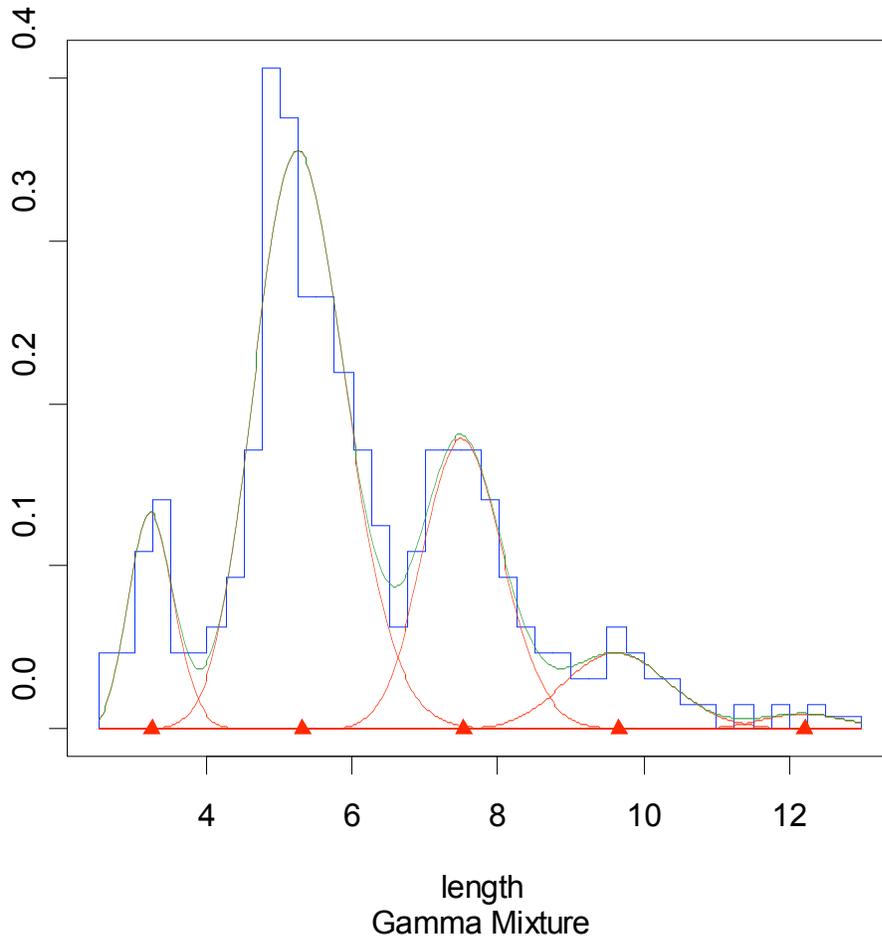
	pi.se	mu.se	sigma.se
1	0.02102	0.08157	0.08283
2	0.04301	0.07614	0.07569
3	0.05792	0.12432	0.17524

```
4 0.04151 0.38549 0.40752
5 0.01051 0.54573 0.49759
```

Analysis of Variance Table

	Df	Chisq	Pr(>Chisq)
Residuals	25	12.589	0.9812

```
> plot(fitcass2)
```



The theory and functionality of Rmix are documented in Du (2002), however, some of the functions have been reorganized since then.

Du, Juan (2002) *Combined Algorithms for Fitting Finite Mixture Distributions*, McMaster University unpublished M.Sc. project.

Peter Macdonald
Ichthus Data Systems
pdmmac@mcmaster.ca
2003-01-24