Next Generation SELECT Examples of Use

Russell Millar

#NB: remotes package needs to be installed  
remotes::install\_github("rbmillar/SELECT")

## Skipping install of 'SELECT' from a github remote, the SHA1 (452af299) has not changed since last install.  
## Use `force = TRUE` to force installation

require(SELECT)

## Loading required package: SELECT

require(tidyverse)

## Loading required package: tidyverse

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

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

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

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

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

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

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

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

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.4 ✔ readr 2.1.5  
## ✔ forcats 1.0.0 ✔ stringr 1.5.1  
## ✔ ggplot2 3.5.1 ✔ tibble 3.2.1  
## ✔ lubridate 1.9.3 ✔ tidyr 1.3.1  
## ✔ purrr 1.0.2

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

# Analysis of Clark’s covered codend haddock data

**Logistic and Richards curves are fitted. We see that the logistic is preferred**

Clark **is a dataframe included within the SELECT package.**

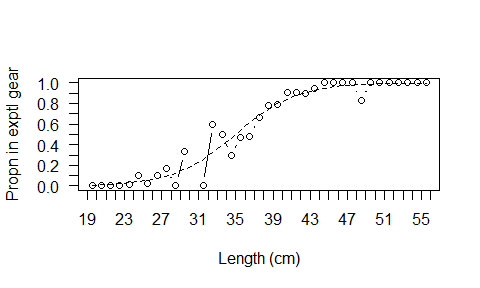
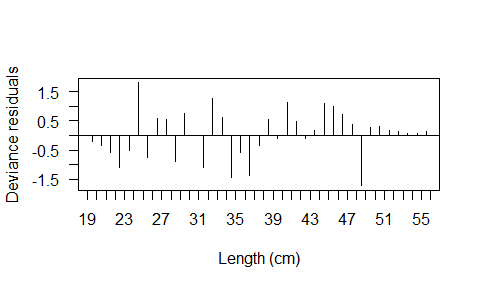
#Use dtype="cc" to specify that design type is covered codend  
data(Clark)  
head(Clark)

## lenclass cover codend  
## 1 19.5 2 0  
## 2 20.5 5 0  
## 3 21.5 11 0  
## 4 22.5 28 0  
## 5 23.5 53 1  
## 6 24.5 46 5

cc.names=c("lenclass","cover","codend")  
#Fit logistic selection curve  
fitL=SELECT(Clark,cc.names,dtype="cc")

## Log-likelihood is -173.4602 at x0= -16.76 0.41   
## Saturated log-likelihood is -26.15727   
## Fitting SELECT model with logistic selection curves to covered codend data.  
## Convergence code 0: Optimizer has converged  
## Pars= -10.63315 0.303986 , Deviance= 23.43561 , #len classes= 36

ModelCheck(fitL,minE=1,las=1)

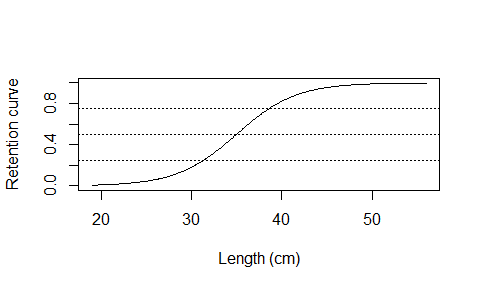


## Model fit:  
## null.l model.l full.l Deviance Pearson.chisq   
## -242.3891493 -37.8750779 -26.1572742 23.4356074 27.1011306   
## dof Deviance.CF Pearson.CF   
## 34.0000000 0.6892826 0.7970921   
##   
## Correction factors from cells with expected count > 1 :  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 12.7692810 13.9166294 14.0000000 0.9120915 0.9940450

Estimates(fitL)

## par s.e.  
## L50 34.979066 0.4246542  
## SR 7.228046 0.5618577

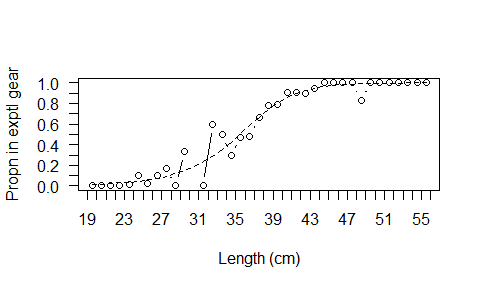
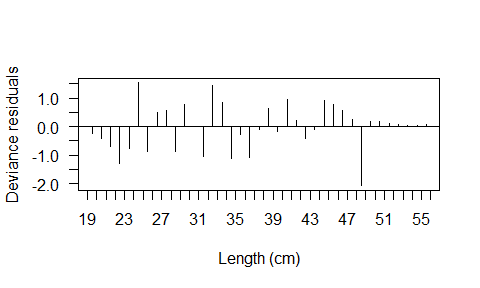
PlotCurves(fitL,plotlens=seq(19,56,0.1))



#Fit Richards selection curve  
fitR=SELECT(Clark,cc.names,dtype="cc",stype="richards")

## Log-likelihood is -173.4602 at x0= -16.76 0.41 0   
## Saturated log-likelihood is -26.15727   
## Fitting SELECT model with richards selection curves to covered codend data.  
## Convergence code 0: Optimizer has converged  
## Pars= -14.49815 0.3882429 0.4855092 , Deviance= 22.30068 , #len classes= 36

ModelCheck(fitR,minE=1,las=1)

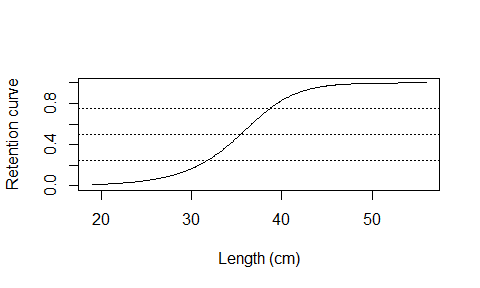


## Model fit:  
## null.l model.l full.l Deviance Pearson.chisq   
## -242.3891493 -37.3076135 -26.1572742 22.3006786 34.9757576   
## dof Deviance.CF Pearson.CF   
## 33.0000000 0.6757781 1.0598714   
##   
## Correction factors from cells with expected count > 1 :  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 10.9179499 11.5869505 12.0000000 0.9098292 0.9655792

Estimates(fitR);

## par s.e.  
## L25 31.826655 0.6453290  
## L50 35.451142 0.5478466  
## L75 38.676063 0.4111978  
## SR 6.849408 0.6009597  
## delta 1.625002 0.6468334

PlotCurves(fitR,plotlens=seq(19,56,0.1))



#Likelihood ratio test for H0 that logistic mode is adequate  
LRTstat=2\*(deviance(fitL)-deviance(fitR))  
cat("\n p-value for H0:logistic is",1-pchisq(LRTstat,1))

##   
## p-value for H0:logistic is 0.131912

**The ModelCheck shows no evidence of overdispersion so we can use twice the difference in deviance as a likelihood ratio test statistic. There in no significant evidence against the logistic curve.**

# Analysis of Pope’s alternate haul haddock data

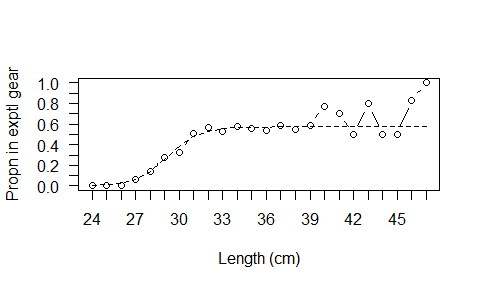
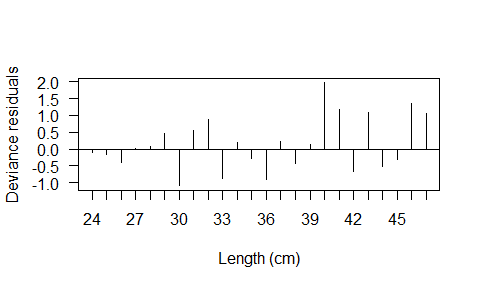
#Use dtype="ph" to specify that design type is paired haul  
data(Pope)  
head(Pope)

## Lengths nfine nwide  
## 1 24 1 0  
## 2 25 1 0  
## 3 26 3 0  
## 4 27 14 1  
## 5 28 30 5  
## 6 29 49 19

ph.names=c("Lengths","nfine","nwide")  
#FIt logistic seln cure  
fitL=SELECT(Pope,ph.names,dtype="ph")

## Log-likelihood is -406.4186 at x0= -24.07 0.7 0   
## Saturated log-likelihood is -39.95578   
## Fitting SELECT model with logistic selection curves to paired haul data.  
## Convergence code 0: Optimizer has converged  
## Pars= -27.64604 0.9161796 0.2935352 , Deviance= 14.83721 , #len classes= 24

ModelCheck(fitL,minE=1,las=1)

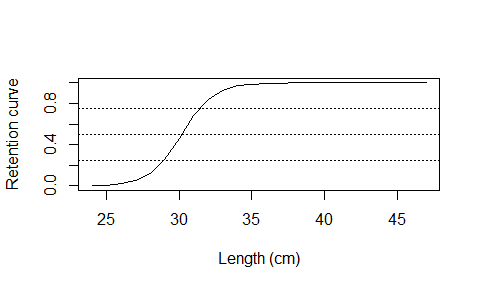


## Model fit:  
## null.l model.l full.l Deviance Pearson.chisq   
## -87.5336729 -47.3743864 -39.9557805 14.8372118 13.7355791   
## dof Deviance.CF Pearson.CF   
## 21.0000000 0.7065339 0.6540752   
##   
## Correction factors from cells with expected count > 1 :  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 13.5204486 12.8869874 16.0000000 0.8450280 0.8054367

Estimates(fitL)

## par s.e.  
## L50 30.1753462 0.35526986  
## SR 2.3982465 0.51419914  
## p 0.5728614 0.01721912

PlotCurves(fitL)



# Analysis of Pope’s alternate haul haddock data with fixed split (of 0.5)

**For fixed split, the user must specify start values for parameters of the selection curve.**

**The fixed split is specified using relative power, in this case, rel.power=c(1,1) since the experimental and control have the same fishing power when the split is 0.5.**

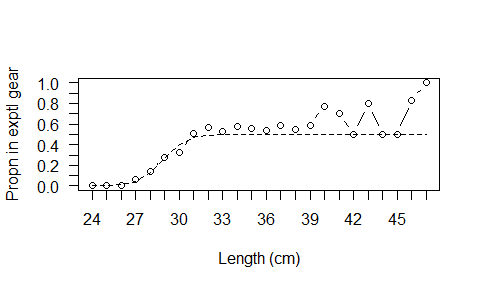
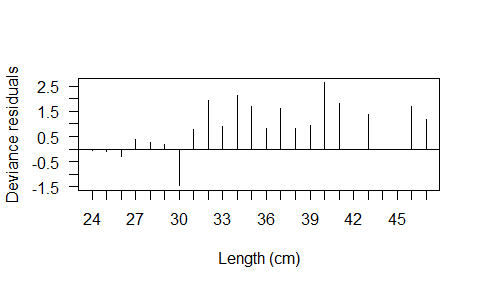
**Use dtype=re for relative (power) fits.**

#Fixed split, psplit=0.5  
fitR2=SELECT(Pope,ph.names,dtype="re",x0=c(-30,1),rel.power=c(1,1))

## Log-likelihood is -60.84633 at x0= -30 1   
## Saturated log-likelihood is -39.95578   
## Fitting SELECT model with logistic selection curves to relative data.  
## Convergence code 0: Optimizer has converged  
## Pars= -36.29322 1.233239 , Deviance= 36.02905 , #len classes= 24

ModelCheck(fitR2,minE=1,las=1)

## Warning in FUN(X[[i]], ...): NaNs produced  
## Warning in FUN(X[[i]], ...): NaNs produced

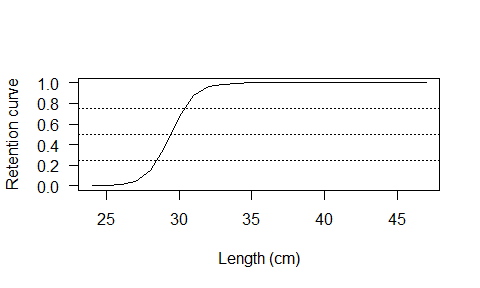


## Model fit:  
## null.l model.l full.l Deviance Pearson.chisq   
## -87.533673 -57.970303 -39.955780 36.029045 34.683422   
## dof Deviance.CF Pearson.CF   
## 22.000000 1.637684 1.576519   
##   
## Correction factors from cells with expected count > 1 :  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 34.409548 33.479306 17.000000 2.024091 1.969371

Estimates(fitR2)

## par s.e.  
## L50 29.42919 0.2506562  
## SR 1.78167 0.3385577

PlotCurves(fitR2,las=1)



# Analysis of Holt’s gillnet data

**Several alternative curve types are fitted. The bi-lognormal looks preferable, evidenced by a decrease in deviance of about 160.** **Even taking in to account the over-dispersion correction factor of about 9, this is still highly significant.** **The starting parameters for the gear with smallest mesh size need to be specified.**

data(Holt)  
Meshs=Holt$Meshsize   
Counts=Holt$Counts  
names(Counts)=c("lgth",paste0("M",Meshs))  
re.names=names(Counts)  
#Equal fishing power of the eight meshsizes  
pwr=rep(1,8)  
#Or use this for fishing power proportional to meshsize  
#pwr=Meshsize  
  
par(mfrow=c(3,2),mar=c(4.1,4.1,1,1))  
fit=SELECT(Counts,re.names,dtype="re",stype="norm.loc",Meshsize=Meshs,x0=c(60,4),rel.power=pwr)

## Log-likelihood is -7040.14 at x0= 60 4   
## Saturated log-likelihood is -171.3226   
## Fitting SELECT model with norm.loc selection curves to relative data.  
## Convergence code 0: Optimizer has converged  
## Pars= 54.60056 5.148722 , Deviance= 862.9243 , #len classes= 11

ModelCheck(fit,minE=1); Estimates(fit); PlotCurves(fit,plotlens=seq(40,90,0.1))

## Model fit:  
## null.l model.l full.l Deviance Pearson.chisq   
## -1867.55849 -602.78470 -171.32257 862.92426 2577.14061   
## dof Deviance.CF Pearson.CF   
## 75.00000 11.50566 34.36187   
##   
## Correction factors from cells with expected count > 1 :  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 778.94965 847.51159 64.00000 12.17109 13.24237

## par s.e.  
## Mode(mesh1) 54.600565 0.07307949  
## Std dev.(mesh1) 5.148722 0.06713026

fit=SELECT(Counts,re.names,dtype="re",stype="norm.sca",Meshsize=Meshs,x0=c(60,4),rel.power=pwr)

## Log-likelihood is -3709.341 at x0= 60 4   
## Saturated log-likelihood is -171.3226   
## Fitting SELECT model with norm.sca selection curves to relative data.  
## Convergence code 0: Optimizer has converged  
## Pars= 55.28193 4.343832 , Deviance= 772.7614 , #len classes= 11

ModelCheck(fit,minE=1); Estimates(fit); PlotCurves(fit,plotlens=seq(40,90,0.1))

## Model fit:  
## null.l model.l full.l Deviance Pearson.chisq   
## -1867.55849 -557.70325 -171.32257 772.76137 1679.84763   
## dof Deviance.CF Pearson.CF   
## 75.00000 10.30348 22.39797   
##   
## Correction factors from cells with expected count > 1 :  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 655.25169 672.44117 63.00000 10.40082 10.67367

## par s.e.  
## Mode(mesh1) 55.281929 0.06986836  
## Std dev.(mesh1) 4.343832 0.05269769

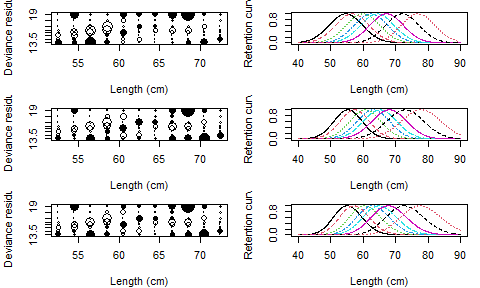
#Needs wide seln curve for x0  
fit=SELECT(Counts,re.names,dtype="re",stype="gamma",Meshsize=Meshs,x0=c(10,5),rel.power=pwr)

## Log-likelihood is -2721.729 at x0= 10 5   
## Saturated log-likelihood is -171.3226   
## Fitting SELECT model with gamma selection curves to relative data.  
## Convergence code 0: Optimizer has converged  
## Pars= 158.2611 0.3499446 , Deviance= 719.2918 , #len classes= 11

ModelCheck(fit,minE=1); Estimates(fit); PlotCurves(fit,plotlens=seq(40,90,0.1))

## Model fit:  
## null.l model.l full.l Deviance Pearson.chisq   
## -1867.558490 -530.968476 -171.322572 719.291810 1629.020063   
## dof Deviance.CF Pearson.CF   
## 75.000000 9.590557 21.720268   
##   
## Correction factors from cells with expected count > 1 :  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 635.842277 675.648540 64.000000 9.935036 10.557008

## par s.e.  
## Mode(mesh1) 55.032682 0.07014009  
## Std dev.(mesh1) 4.402369 0.05446245



fit=SELECT(Counts,re.names,dtype="re",stype="logistic",Meshsize=Meshs,x0=c(-60,1),rel.power=pwr)

## Log-likelihood is -2121.887 at x0= -60 1   
## Saturated log-likelihood is -171.3226   
## Fitting SELECT model with logistic selection curves to relative data.  
## Convergence code 0: Optimizer has converged  
## Pars= -37.15865 0.5363473 , Deviance= 1571.981 , #len classes= 11

ModelCheck(fit,minE=1); Estimates(fit); PlotCurves(fit,plotlens=seq(40,90,0.1))

## Model fit:  
## null.l model.l full.l Deviance Pearson.chisq   
## -1867.55849 -957.31295 -171.32257 1571.98076 1956.95185   
## dof Deviance.CF Pearson.CF   
## 75.00000 20.95974 26.09269   
##   
## Correction factors from cells with expected count > 1 :  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 1542.56782 1630.70802 67.00000 23.02340 24.33893

## par s.e.  
## L50 69.280951 0.1273376  
## SR 4.096645 0.1297834

fit=SELECT(Counts,re.names,dtype="re",stype="richards",Meshsize=Meshs,x0=c(-60,1,0),rel.power=pwr)

## Log-likelihood is -2121.887 at x0= -60 1 0   
## Saturated log-likelihood is -171.3226   
## Fitting SELECT model with richards selection curves to relative data.  
## Convergence code 0: Optimizer has converged  
## Pars= -152.9469 2.12612 1.862147 , Deviance= 1466.833 , #len classes= 11

ModelCheck(fit,minE=1); Estimates(fit); PlotCurves(fit,plotlens=seq(40,90,0.1))

## Model fit:  
## null.l model.l full.l Deviance Pearson.chisq   
## -1867.55849 -904.73914 -171.32257 1466.83313 1626.10480   
## dof Deviance.CF Pearson.CF   
## 74.00000 19.82207 21.97439   
##   
## Correction factors from cells with expected count > 1 :  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 1454.0023 1593.0720 69.0000 21.0725 23.0880

## par s.e.  
## L25 67.739704 0.1233384  
## L50 69.843837 0.1308425  
## L75 71.146350 0.1603286  
## SR 3.406646 0.1404361  
## delta 6.437546 1.7037686

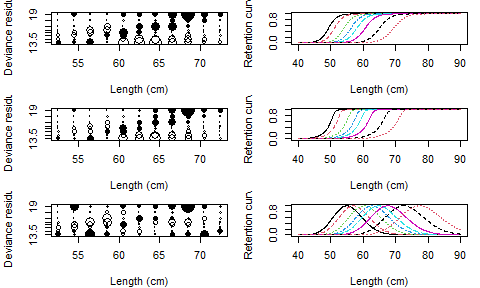
fit=SELECT(Counts,re.names,dtype="re",stype="lognorm",Meshsize=Meshs,x0=c(log(60),0.1),rel.power=pwr)

## Log-likelihood is -1441.164 at x0= 4.09 0.1   
## Saturated log-likelihood is -171.3226   
## Fitting SELECT model with lognorm selection curves to relative data.  
## Convergence code 0: Optimizer has converged  
## Pars= 4.012147 0.08026262 , Deviance= 704.2809 , #len classes= 11

ModelCheck(fit,minE=1); Estimates(fit); PlotCurves(fit,plotlens=seq(40,90,0.1))

## Model fit:  
## null.l model.l full.l Deviance Pearson.chisq   
## -1867.558490 -523.463042 -171.322572 704.280940 2285.670627   
## dof Deviance.CF Pearson.CF   
## 75.000000 9.390413 30.475608   
##   
## Correction factors from cells with expected count > 1 :  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 622.44927 660.50732 64.00000 9.72577 10.32043

## par s.e.  
## Mode(mesh1) 54.910506 0.07053896  
## Std dev.(mesh1) 4.457232 0.05826788



fit=SELECT(Counts,re.names,dtype="re",stype="binorm.sca",Meshsize=Meshs,x0=c(55,4,65,4,2),rel.power=pwr)

## Log-likelihood is -483.6976 at x0= 55 4 65 4 2   
## Saturated log-likelihood is -171.3226   
## Fitting SELECT model with binorm.sca selection curves to relative data.  
## Convergence code 0: Optimizer has converged  
## Pars= 54.67054 3.598595 80.34086 15.75706 1.490215 , Deviance= 563.8325 , #len classes= 11

ModelCheck(fit,minE=1); Estimates(fit); PlotCurves(fit,plotlens=seq(40,90,0.1))

## Model fit:  
## null.l model.l full.l Deviance Pearson.chisq   
## -1867.558490 -453.238845 -171.322572 563.832547 624.786768   
## dof Deviance.CF Pearson.CF   
## 72.000000 7.831008 8.677594   
##   
## Correction factors from cells with expected count > 1 :  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 554.606738 612.419948 67.000000 8.277713 9.140596

## Warning in sqrt(diag(varpars)): NaNs produced

## par s.e.  
## Mode1(mesh1) 54.6705386 0.06597046  
## Std dev.1(mesh1) 3.5985947 0.06432625  
## Mode2(mesh1) 80.3408633 NaN  
## Std dev.2(mesh1) 15.7570589 NaN  
## P(mode1) 0.8161106 NaN

fit=SELECT(Counts,re.names,dtype="re",stype="bilognorm",Meshsize=Meshs,x0=c(4,0.2,4.2,0.1,2),rel.power=pwr)

## Log-likelihood is -1935.07 at x0= 4 0.2 4.2 0.1 2   
## Saturated log-likelihood is -171.3226   
## Fitting SELECT model with bilognorm selection curves to relative data.  
## Convergence code 0: Optimizer has converged  
## Pars= 4.002056 0.06584885 4.226652 0.2145388 2.076375 , Deviance= 545.5346 , #len classes= 11

ModelCheck(fit,minE=1); Estimates(fit); PlotCurves(fit,plotlens=seq(40,90,0.1))

## Model fit:  
## null.l model.l full.l Deviance Pearson.chisq   
## -1867.558490 -444.089890 -171.322572 545.534636 619.161453   
## dof Deviance.CF Pearson.CF   
## 72.000000 7.576870 8.599465   
##   
## Correction factors from cells with expected count > 1 :  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 537.678287 605.836126 68.000000 7.907034 8.909355

## par s.e.  
## Mode1(mesh1) 54.4738204 0.08838661  
## Std dev.1(mesh1) 3.6143623 0.10510480  
## Mode2(mesh1) 65.4067212 4.79566934  
## Std dev.2(mesh1) 15.2099756 4.40707578  
## P(mode1) 0.8885856 0.01701247

