SELECT package demonstration  
Twin haul

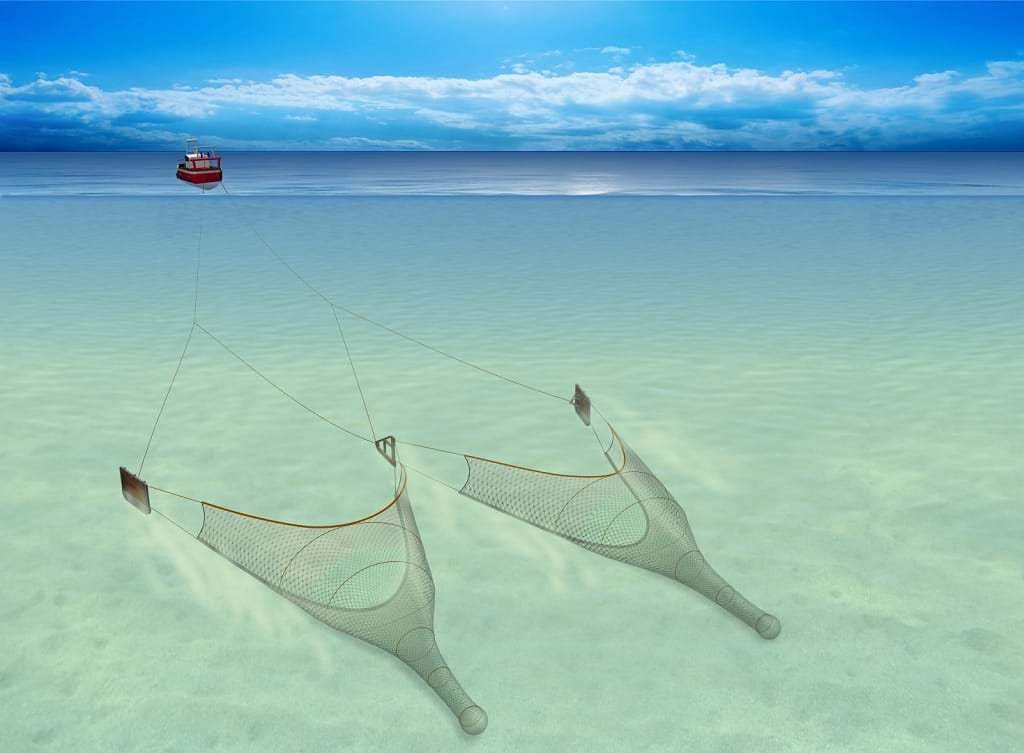
## Summary

**This case study demonstrates:**

* Estimating logistic and Richards selection curves from twin-haul data.
* Fitting to catch-share curves.
* Using combined individual-haul data to estimate over-dispersion (due to causes such as sub-sampling, between-haul variability and non-independent behaviour of fish).
* Using over-dispersion adjustments.

### Data source

The data are for Eastern king prawn selectivity in a 20 mm square-mesh penaeid trawl (Broadhurst et al., 2004. Selectivity of conventional diamond- and novel square-mesh codends in an Australian estuarine penaeid-trawl fishery. Fisheries Research, 67: 183-194).



require(SELECT)  
require(dplyr)

### Input the data

#Read in data and remove zero catch lengths  
sqKP.df=read.csv("../Data/SqNorKP.csv",header=T) #500 by 4  
sqKP.df=subset(sqKP.df,control+codend>0) #192 by 4  
#Quick peek at data  
head(sqKP.df)

## lgth codend control PairID  
## 10 10 0 6 1  
## 11 11 5 6 1  
## 12 12 11 11 1  
## 13 13 22 6 1  
## 14 14 13 14 1  
## 15 15 13 6 1

### Define variable names

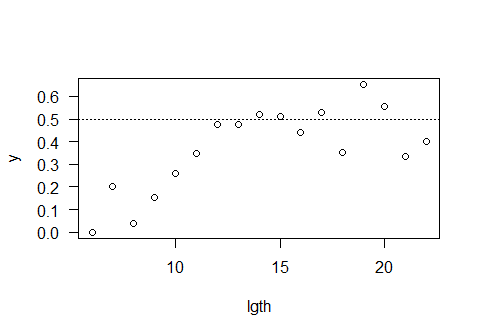
v.names=c("lgth","control","codend")

### Produce plot of catch-share proportions

Tots=Raw2Tots(sqKP.df,v.names)  
Tots=transform(Tots,n=codend+control,y=codend/(codend+control))   
Tots

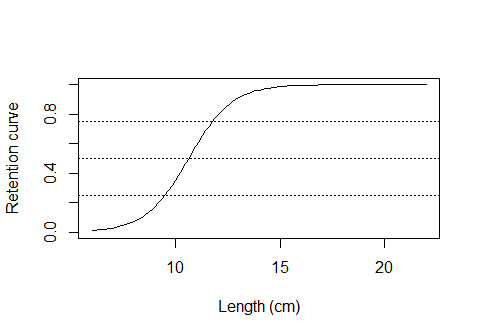
## lgth control codend n y  
## 1 6 1 0 1 0.0000000  
## 2 7 4 1 5 0.2000000  
## 3 8 24 1 25 0.0400000  
## 4 9 33 6 39 0.1538462  
## 5 10 66 23 89 0.2584270  
## 6 11 104 55 159 0.3459119  
## 7 12 111 100 211 0.4739336  
## 8 13 116 105 221 0.4751131  
## 9 14 89 96 185 0.5189189  
## 10 15 73 76 149 0.5100671  
## 11 16 60 47 107 0.4392523  
## 12 17 34 38 72 0.5277778  
## 13 18 26 14 40 0.3500000  
## 14 19 8 15 23 0.6521739  
## 15 20 4 5 9 0.5555556  
## 16 21 4 2 6 0.3333333  
## 17 22 3 2 5 0.4000000

plot(y~lgth,data=Tots,las=1)  
abline(h=0.5,lty=3)

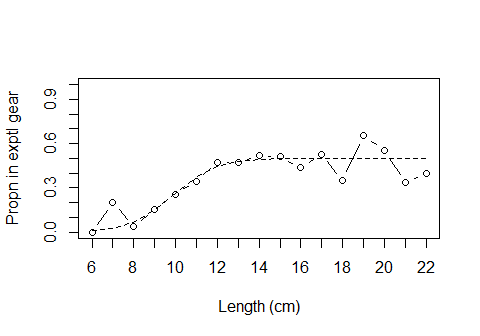
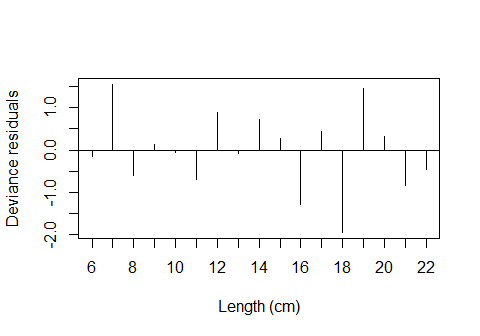


### Fit logistic selection curve

Logist.fit=SELECT(Tots,var.names=v.names,dtype="ec")  
plot(Logist.fit)



ModelCheck(Logist.fit)



## Model fit:  
## null.l model.l full.l npar AIC   
## -69.63410 -38.95222 -32.29383 3.00000 83.90443   
## GOF:  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 13.3167657 16.2189932 14.0000000 0.9511976 1.1584995

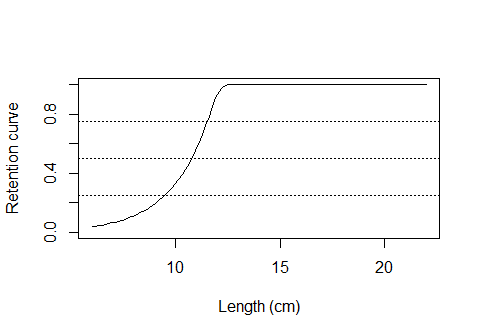
Estimates(Logist.fit)

## par raw s.e.  
## L50 10.6340407 0.3181545  
## SR 2.2629652 0.4569826  
## p 0.5020243 0.0200655

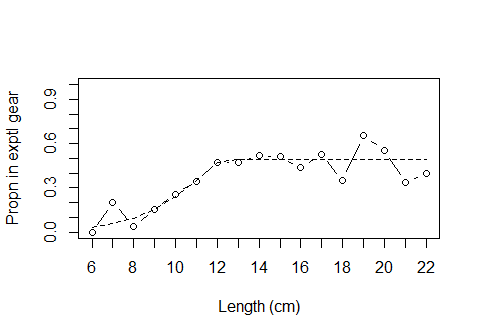
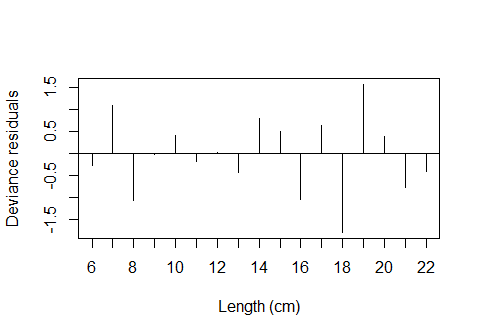
The logistic fit looks good.

### Fit Richards curve

Rich.fit=SELECT(Tots,var.names=v.names,dtype="ec",stype="richards")  
plot(Rich.fit)



ModelCheck(Rich.fit)



## Model fit:  
## null.l model.l full.l npar AIC   
## -69.63410 -38.14618 -32.29383 4.00000 84.29237   
## GOF:  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 11.7046981 12.0196464 13.0000000 0.9003614 0.9245882

Estimates(Rich.fit)

## par raw s.e.  
## L50 10.7689854 0.26819278  
## SR 2.0174087 0.48179535  
## delta 12.4539883 36.91844960  
## p 0.4899005 0.01749559

Even without correcting for possible over-dispersion it is clear that the extra complexity of the Richard curve is not needed. The logistic is good enough.

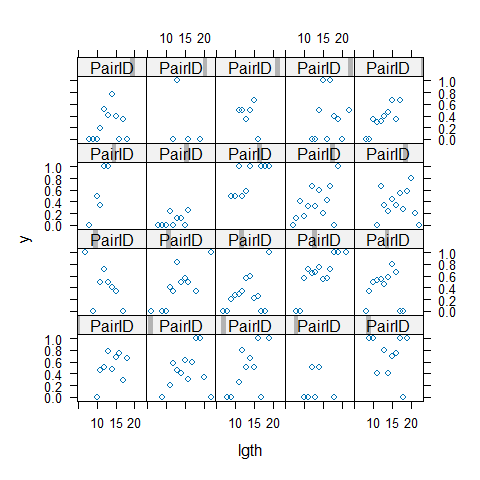
## Correcting for over-dispersion

### Produce a lattic plot of catch-share proportions by haul

HaulCatch=Raw2Tots(sqKP.df,v.names,sumHauls=F)  
HaulCatch=transform(HaulCatch,n=codend+control,y=codend/(codend+control))   
require(lattice) #For xyplot

## Loading required package: lattice

xyplot(y~lgth | PairID,data=HaulCatch)



The retention proportions are certainly highly variable, but it is hard to distinguish BHV from within-haul variability.

### Fit a logistic selection curve to the combined individual pair data

Logist.fit2=SELECT(sqKP.df,var.names=v.names,dtype="ec",sumHauls=F)  
ModelCheck(Logist.fit2, minE=5, plots=F)

## Model fit:  
## null.l model.l full.l npar AIC   
## -323.3478 -292.6659 -142.0254 3.0000 591.3318   
## GOF:  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 301.280993 282.166530 189.000000 1.594079 1.492945   
##   
## Correction factors from cells with expected count > 5 :  
## Deviance Pearson.chisq dof Deviance.CF Pearson.CF   
## 82.540482 74.852420 36.000000 2.292791 2.079234

The variance correction factor Pearson.CF shows that the actual variabilty is about twice that assumed under the binomial model. This is relatively low, and a bootstrap would probably over-estimate the standard errors.

In this case it is preferable to do a simple over-dispersion adjustment to the raw standard errors.

Estimates(Logist.fit,OD=2.08)

## par raw s.e. adj s.e.  
## L50 10.6340407 0.3181545 0.45884892  
## SR 2.2629652 0.4569826 0.65906975  
## p 0.5020243 0.0200655 0.02893888