

fishR Vignette - Catch Curve Estimates of Mortality

Dr. Derek Ogle, Northland College

September 29, 2011

Mortality is a key component to understanding the population dynamics of fish species. Total mortality is often estimated from the sequential decline observed in cohorts of fish. The methods used to analyze this decline are collectively called catch-curve methods and form the topic of this vignette.

A general background, including the required type of data, for catch curve analyses is described in Section 1. The catch curve regression analysis is described in Section 2 and the Chapman-Robson method is described in Section 3 with a discussion of the relative advantages of the two methods in Section 4. Methods for comparing the estimates of total mortality between groups is illustrated in Section 5. Finally, an appendix is given with a derivation for the Chapman-Robson method.

This vignette requires functions in the `FSA` and `NCStats` package maintained by the author. These packages are loaded into R with

```
> library(FSA)
> library(NCStats)
```

1 Background

1.1 Longitudinal Data

In the statistical literature, longitudinal data is data that occurs from extracting multiple samples from the same group of individuals over time. Thus, catches of fish from the same cohort over time is longitudinal data. Following a single cohort of fish over time is the same as following a group of fish in a closed population that is not subject to inclusion of “new” individuals through the birth process. Catch curve analysis is used to estimate total mortality by observing the regular decline of individuals in a cohort.

The decline in individuals in a cohort can be theoretically modeled with a modified continuous exponential “growth” model. This model assumes that the population is closed to emigration and immigration and it will be modified to assume that no fish can be added to the population. Thus, the population is only affected by the loss of individuals to mortality. Therefore, the instantaneous growth rate parameter in the model is replaced with an instantaneous total mortality parameter (Z) so that the modified model looks like

$$N_t = N_0 e^{-Zt} \quad (1)$$

where N_t is the population size at time t and N_0 is the initial population size.

One form of the “catch equation” states that the “catch” of fish at age t (i.e., C_t) is proportional to the number of fish of age t available, or

$$C_t = vN_t \quad (2)$$

where v represents a constant proportion of the population that was “vulnerable” to the fishery. This can be rearranged to show the relationship between population size and catch,

$$N_t = \frac{C_t}{v}$$

which is then substituted into (1) to reveal

$$\begin{aligned}\frac{C_t}{v} &= N_0 e^{-Zt} \\ C_t &= v N_0 e^{-Zt}\end{aligned}\tag{3}$$

The shape of this model is shown in Figure 1-Left.

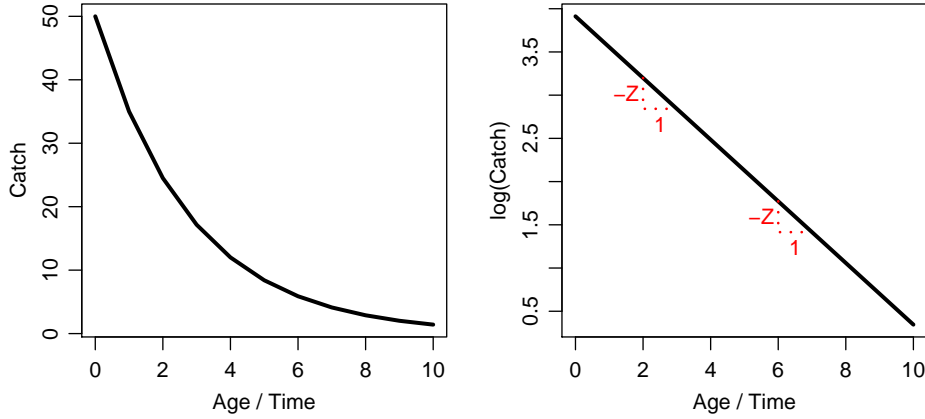


Figure 1. Ideal plots of catch versus age (Left) and the natural log of catch versus age (Right) for a single cohort of fish. The right graph is called a longitudinal catch curve. The change in $\log(C_t)$ for a unit change in t is emphasized on the catch curve to reinforce the idea that the slope of the idealized catch curve is Z .

Natural logarithms of both sides of (3) yields,

$$\log(C_t) = \log(vN_0) - Zt\tag{4}$$

which is in the form of a linear equation with $\log(C_t)$ on the y-axis and t on the x-axis (Figure 1-Right). Of great interest in (4) is that the negative of the slope in this model is Z . Thus, the negative of the slope of the regression between $\log(C_t)$ and t is an integrative measure of the instantaneous total mortality rate experienced by this cohort of fish over time.

Another form of the “catch equation” states that the “catch” of fish at age t is proportionately related to the number of fish of age t available and the amount of effort expended to catch those fish (i.e., E_t), or

$$C_t = qE_t N_t\tag{5}$$

where q represents a constant proportion of the population captured by one unit of effort. This catch equation can be rearranged to show the relationship between population size and catch-per-unit-effort,

$$N_t = \frac{1}{q} \frac{C_t}{E_t}$$

which is then substituted into (1) to reveal

$$\begin{aligned}\frac{1}{q} \frac{C_t}{E_t} &= N_0 e^{-Zt} \\ \frac{C_t}{E_t} &= q N_0 e^{-Zt}\end{aligned}\tag{6}$$

Natural logarithms of both sides of (6) yields,

$$\log\left(\frac{C_t}{E_t}\right) = \log(qN_0) - Zt \quad (7)$$

which is in the form of a linear equation with $\log(\frac{C_t}{E_t})$ on the y-axis and t on the x-axis. Thus, the negative of the slope of the regression between $\log(\frac{C_t}{E_t})$ and t is also an integrative measure of the instantaneous total mortality rate experienced by this cohort of fish over time. In other words, the y-axis of the catch-curve can either represent the catch or the catch-per-unit-effort of the cohort. The specifics of this regression methodology are discussed in Section 2.

1.2 Cross-Sectional Data

Let's broaden our perspective a bit and explore several cohorts of fish at once. In doing so, let's assume, in addition to the usual assumptions of a closed population and a constant Z , that each cohort starts with the same number of fish (i.e., N_0 is the same for each cohort). For example, the catch of the hypothetical 1990 year-class arranged by age and capture year is shown on a diagonal in Table 1. The catches of other year-classes is shown in the same table.

Table 1. The hypothetical catch of fish by age and capture year. The longitudinal catch of the 1990 and the partial 1994 year-classes of fish are shown by the two sets of diagonal cells highlighted in dark grey. The cross-sectional catch in the 1997 capture year is shown by the column of cells highlighted in light grey. All data were modeled with (3) assuming that $N_0 = 500$, $Z = -\log(0.7)$, and $v = 0.1$.

Age	Capture Year											
	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
1	35.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0
2	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5
3	17.1	17.1	17.1	17.1	17.1	17.1	17.1	17.1	17.1	17.1	17.1	17.1
4	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0	12.0
5	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4	8.4
6	5.9	5.9	5.9	5.9	5.9	5.9	5.9	5.9	5.9	5.9	5.9	5.9
7	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1	4.1
8	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9
9	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
10	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4

The number of fish of each age in any given capture year is shown within a column of Table 1. Upon closer examination it can be seen that the vector of catches at each age within a given capture year (i.e., a column in Table 1) is the same as the vector of catches at each age (or time) of a given year-class of fish (i.e., a diagonal in Table 1). Thus, with the added assumption of a constant initial number of fish for a year-class, the aged cross-sectional sample from a given capture year will yield the same estimate of Z as the longitudinal sample of a cohort.

1.3 Characteristics

All catch curves, regardless of whether they come from longitudinal or cross-sectional samples, have three regions of interest: an ascending left limb, a domed middle portion, and a descending right limb (Figure 2). The ascending left limb represents age-classes of fish that are not fully vulnerable to the gear used in the fishery. Fish in these age-classes are said to have "not fully recruited to the fishery." The catches of fish in these age-classes are not useful for estimating the total mortality rate.

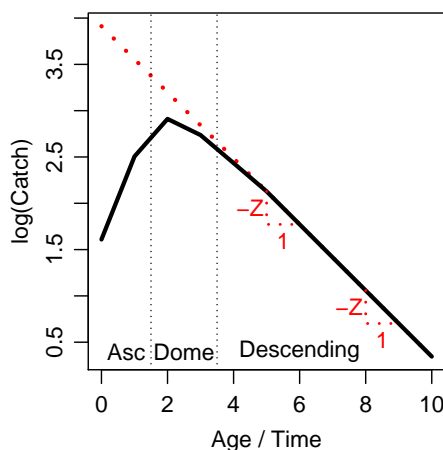


Figure 2. Idealized catch curve (plot of the natural log of catch versus age) illustrating the ascending, domed, and descending portions. The red dotted line represents the idealized catch curve if all age-classes were fully recruited to the fishery.

The domed portion of the catch-curve generally consists of age-classes of fish that are nearly, but incompletely, recruited to the fishery. The relative width of the domed portion provides some insight into the rate of recruitment. For example, a very sharply pointed dome indicates that the fish recruit rather “quickly”¹. In contrast, a relatively rounded dome shows that fish recruit to the exploited phase of the population more slowly, perhaps requiring several years before the mean size of fish in that year-class is sufficiently large to ensure capture upon encounter with the gear. Fish in age-classes in the domed portion of the catch curve are also excluded from use when estimating Z . Despite the exclusion of age-classes in the ascending limb and domed portion of the catch curve it is, however, imperative to have some animals from these age-classes in your sample, so that you can identify the important descending limb of the catch curve.

The descending left limb of the catch curve represents the regular decline of fully-recruited individuals in the fishery. Thus, Z can be estimated by applying the concept of (4) to the catches of fish in the ages corresponding only to the descending portion of the catch curve.

1.4 Assumptions

As with any model, the analysis of catch curves for estimating instantaneous total mortality rate depends on a series of assumptions being met. The longitudinal and cross-sectional methods share the following assumptions,

- “Closed Population” – there is no immigration or emigration to the population.
- “Constant Mortality” – The instantaneous total mortality rate is independent of age and year (i.e., constant) for ages on the descending limb of the catch curve.
- “Constant Vulnerability” – The vulnerability (*if catch data is used*) and catchability (*if CPUE data is used*) of the fish to the fishery, for ages on the descending limb of the catch curve, is independent of age and year (i.e., constant).
- “Unbiased Sample” – The sample is not biased regarding any specific age-group(s).

The longitudinal method has the following additional assumption,

- “Follow A Cohort” – A cohort can be accurately examined at multiple times.

¹This may be something similar to the specification of ‘knife-edge’ selection that may suggest that the fish grow rapidly relative to the time or scale of the x-axis.

Finally, the cross-sectional method has the following additional assumptions,

- “Constant Recruitment” – The initial number of individuals is the same for each year-class of fish.
- “Accurate Ages” – The fish in a sample can be accurately assigned an age.

Violations of these assumptions often lead to catch curves that are “bumpy”, convex, concave, or offset rather than linear in the right descending limb (Figure 3).

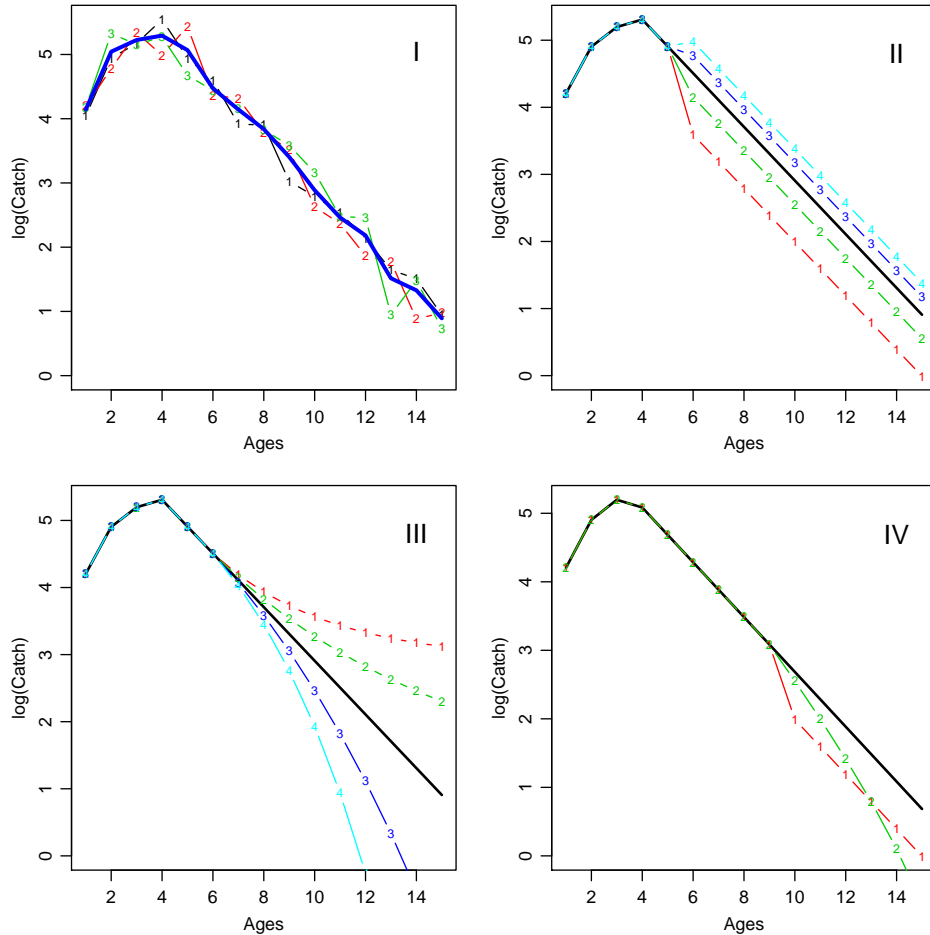


Figure 3. Simulated catch curves to illustrate shapes when assumptions are violated. Each simulation (i.e., plot), unless otherwise, noted uses $N_0 = 1000$ for each year-class, $Z = 0.40$, and incomplete recruitment until age-4 and then constant recruitment for subsequent ages. In simulation I, a coefficient of variation for N_0 of 0.3 was used. In simulation II, constant multipliers of change in recruitment of 0.4, 0.7, 1.3, and 1.6 were applied at age-6. In simulation III, geometric multipliers of Z by age of 0.8, 0.9, 1.1, and 1.2 were applied at age-6. In simulation IV, the vulnerability of age-10 and older fish was cut in half (in run 1) and decreased by 0.1 for each age (in run 2). In each plot, the catch curve with no assumption violations is shown as a solid black line. In simulation I, the average of the three runs is shown as a solid blue line.

1.5 Instantaneous vs. Annual Mortality Rates

The instantaneous mortality rate (Z) that is estimated via the catch curve method is a measure of (i) how much the natural log of number of individuals declines annually or (ii) how much the actual number of individuals declines in an imperceptibly short period of time (i.e., in an “instant”). The instantaneous mortality rate has some very useful mathematical properties, but providing a practical interpretation of its

meaning is difficult – e.g., what does it mean if the log number of individuals declines by 0.693 or if the population changes by 0.693 in a “second” of time? Fortunately, the instantaneous mortality rate can be easily converted to an annual mortality rate (A), the proportion of the population that suffers mortality in a given year, with

$$A = 1 - e^{-Z}$$

Thus, a Z of 0.693 corresponds to an A of $1 - e^{-0.693}$ or 0.500. Thus, this largely uninterpretable value of Z corresponds to an annual mortality rate of 50.0%. In other words, an average of 50.0% of the population dies on an annual basis.

2 Catch Curve Analysis

2.1 Step-by-Step Regression Method

The required data for estimating the instantaneous total mortality rate consists of the catches or CPUEs determined for each age of fish in the fishery (Table 2). As shown previously, these data can be obtained by following a cohort through time or from a sample from one year in the fishery (although the assumptions are different for each type of data).

Table 2. Cross-sectional total catch-at-age of Tobin Harbor brook trout in fyke nets, 1996-1998.

Age	Catch
0	39
1	93
2	112
3	45
4	58
5	12
6	8

The instantaneous total mortality rate can be estimated by plotting the natural log of catches versus age to identify the descending limb of the curve and then fitting a regression model to only the data on the descending limb. The negative of the slope from fitting this model is the estimate of Z , or \hat{Z} . Because \hat{Z} is essentially the slope of this model, the confidence interval for the slope is also the confidence interval for Z . This process is illustrated below with catch-at-age data for coaster brook trout (*Salvelinus fontinalis*) captured by the U.S. Fish and Wildlife Service from 1996-1998 Tobin Harbor of Isle Royale (data shown in (Table 2); Quinlan (1999.)).

The Tobin Harbor brook trout catch-at-age data² were entered into a data frame with³

```
> ( bkt <- data.frame(age=0:6, ct=c(39,93,112,45,58,12,8)) )
```

```
  age  ct
1    0  39
2    1  93
3    2 112
4    3  45
5    4  58
6    5  12
7    6   8
```

²These data can also be obtained with `data(BrkTrtTH)`.

³The “extra” parentheses around this command simply force the result that is saved to an object to be printed to the console.

```
> str(bkt)

'data.frame':      7 obs. of  2 variables:
 $ age: int  0 1 2 3 4 5 6
 $ ct : num  39 93 112 45 58 12 8
```

A variable that contains the natural log of the catches was then appended to the data frame with

```
> bkt$logct <- log(bkt$ct)
> str(bkt)

'data.frame':      7 obs. of  3 variables:
 $ age  : int  0 1 2 3 4 5 6
 $ ct   : num  39 93 112 45 58 12 8
 $ logct: num  3.66 4.53 4.72 3.81 4.06 ...
```

The catch-curve plot was produced with

```
> plot(logct~age,data=bkt,pch=19,ylab="log(Catch)")
```

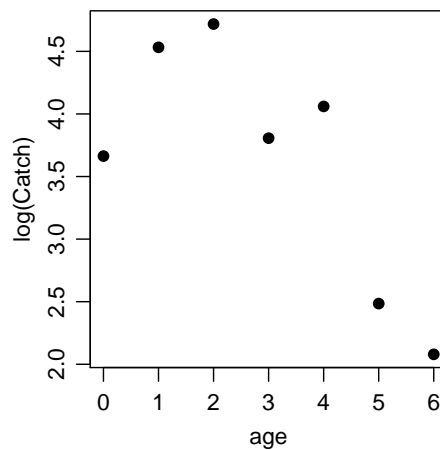


Figure 4. Catch curve for the Tobin Harbor brook trout data.

From this plot (Figure 4) it was determined that the descending limb consisted of catches at ages 2 through 6. A new data frame containing just the information for ages on the descending limb was extracted with

```
> ( bkt.d <- Subset(bkt,age >= 2) )

  age  ct  logct
3   2 112 4.718499
4   3  45 3.806662
5   4  58 4.060443
6   5  12 2.484907
7   6   8 2.079442
```

The linear regression model applied to the descending limb is fit and then seen with

```
> cc <- lm(logct~age,data=bkt.d)
> summary(cc)
```

```

Call:
lm(formula = logct ~ age, data = bkt.d)

Residuals:
    3      4      5      6      7
-0.03147 -0.28332  0.63045 -0.28510 -0.03057

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.0699     0.5801  10.463  0.00186
age         -0.6600     0.1367  -4.827  0.01695

Residual standard error: 0.4324 on 3 degrees of freedom
Multiple R-squared:  0.8859,    Adjusted R-squared:  0.8479
F-statistic: 23.3 on 1 and 3 DF,  p-value: 0.01695

```

From these results it is seen that \hat{Z} is -0.66 with a $SE_{\hat{Z}}$ of 0.14. The corresponding estimated annual mortality rate (\hat{A}) is 48.3% as obtained with (note the negative sign in the Z line),

```

> ( Z <- -coef(cc)[2] )

age
0.659987

> ( A <- 1-exp(-Z) )

age
0.483142

```

Confidence intervals for the slope and intercept of the regression fit are obtained by sending the saved `lm()` object to `confint()` as follows,

```

> confint(cc)

            2.5 %      97.5 %
(Intercept)  4.223665  7.9162122
age         -1.095158 -0.2248162

```

However, there is no need to have a confidence interval for the intercept, the confidence interval for the slope should have the negatives removed, and it would be preferable to have the slope values reversed. These modifications are accomplished, though the labels are backwards, with

```

> ( Z.ci <- -confint(cc)[2,2:1] )

    97.5 %    2.5 %
0.2248162 1.0951579

```

Approximate 95% confidence interval for A are then obtained with

```

> ( A.ci <- 100*(1-exp(-Z.ci)) )

    97.5 %    2.5 %
20.13370 66.55132

```

Thus, from this sample of brook trout, the estimated instantaneous total mortality rate is between 0.22 and 1.10 and the estimated annual mortality rate is between 20.1% and 66.6%.

2.2 catchCurve() Function

The catch curve analysis depicted in the previous section can be more efficiently obtained with `catchCurve()`. This function requires three arguments. The first argument is a formula of the form `catch ~ age` where `catch` and `age` generically represent the variables containing the catches and ages for the catch curve. A `data=` argument set to the data frame containing the catch and age variables is also required. This data frame does NOT have to consist of only the descending limb of the catch curve. The required `ages2use=` argument is a vector identifying the ages corresponding to the descending limb of the catch curve. The results of `catchCurve()` should be saved to an object which is then submitted to `summary()` to obtain the estimated Z and A values⁴, to `confint()` to obtain the confidence intervals for Z and A , and to `plot()` to return the catch-curve with the descending limb highlighted, the regression model superimposed, and the mortality rate estimates printed.

The catch curve analysis for the Tobin Harbor brook trout using ages two through six, with illustrative plot Figure 5, is obtained with

```
> thcc <- catchCurve(ct~age,data=bkt,ages2use=2:6)
> summary(thcc)
```

```
      Estimate Std. Error t value Pr(>|t|)
Z  0.659987    0.136741  4.826549 0.01695159
A 48.314197         NA         NA         NA
```

```
> confint(thcc)
```

```
      95% LCI  95% UCI
Z  0.2248162  1.095158
A 20.1337012 66.551321
```

```
> plot(thcc)
```

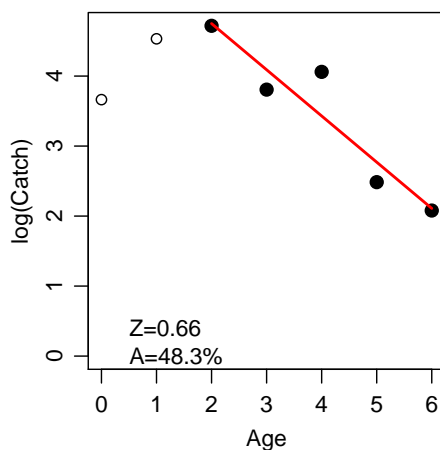


Figure 5. Catch curve for the Tobin Harbor brook trout data.

2.3 Weighted Catch-Curve Regression

Maceina and Bettoli (1998) suggested that a weighted regression should be used with the catch-curve method in order to reduce the relative impact of older ages with fewer fish. They suggested that an unweighted

⁴The `summary()` and `confint()` functions can also take an optional `type="lm"` argument to return the summary and confidence intervals for the slope and intercept of the linear model.

regression should be fit to the descending limb of the catch curve and that the resultant model should be used to predict the natural log number of fish in each age class. These predictions can then serve as the weights to a second regression of the natural log of catches on age. The methodology of [Maceina and Bettoli \(1998\)](#) is implemented with `catchCurve()` by including the `use.weights=TRUE` argument. This is illustrated below for the Tobin Harbor brook trout data,

```
> thcc2 <- catchCurve(ct~age,data=bkt,ages2use=2:6,use.weights=TRUE)
> summary(thcc2)
```

```
      Estimate Std. Error t value Pr(>|t|)
Z  0.6430183  0.1417433  4.5365 0.02004993
A 47.4296703          NA      NA      NA
```

```
> confint(thcc2)
```

```
      95% LCI  95% UCI
Z  0.191928  1.094109
A 17.463369 66.516206
```

3 Chapman-Robson Method

3.1 Background

[Chapman and Robson \(1960\)](#) (and [Robson and Chapman \(1961\)](#)) provided an alternative method for estimating the total annual survival rate (S), and thus the annual (A) and instantaneous (Z) total mortality rates, from catch curve data. Their method was based on understanding that the catches at each age on the descending limb of the catch curve followed a geometric probability distribution and using this to derive a maximum likelihood estimator for the survival parameter of the distribution. Their method, called the Chapman-Robson method, is outlined below and the derivation of their method is shown in [Appendix A](#).

The Chapman-Robson estimate of the annual survival rate is,

$$\hat{S} = \frac{T}{n + T - 1} \quad (8)$$

where n is the total number of fish observed on the descending limb of the catch curve and T is the total recoded age of fish on the descending limb of the catch curve. It should be noted that the ages are “recoded” such that the first fully-recruited age on the descending limb of the catch-curve is set to 0. The total recoded age is calculated as a weighted sum of the recoded ages where the weights are the catches at each age. The standard error of this estimate is,

$$SE_{\hat{S}} = \sqrt{\frac{T}{n + T - 1} \left(\frac{T}{n + T - 1} - \frac{T - 1}{n + T - 2} \right)} = \sqrt{\hat{S} \left(\hat{S} - \frac{T - 1}{n + T - 2} \right)} \quad (9)$$

If n is large, as it often is in fisheries catch data, then $SE_{\hat{S}}$ can be estimated by,

$$SE_{\hat{S}} = \sqrt{\frac{\hat{S}(1 - \hat{S})^2}{n}} \quad (10)$$

The Chapman-Robson estimates of S can easily be transformed into estimates of Z through the relationship $S = e^{-Z}$. Thus, the Chapman-Robson estimate of Z is,

$$\hat{Z} = -\log(\hat{S}) = -\log\left(\frac{T}{n+T-1}\right) \quad (11)$$

and the large sample approximation of $SE_{\hat{Z}}$ (Jensen 1985) is

$$SE_{\hat{Z}} = \frac{SE_{\hat{S}}}{\hat{S}} \quad (12)$$

These calculations are illustrated below with the Tobin Harbor brook trout (assuming that the fish were fully recruited to the fyke nets at age-2). The original data (Table 2) were modified for calculation of the Chapman-Robson estimator of S as such,

Age	Recoded Age	Catch	Recode*Catch
2	0	112	0
3	1	45	45
4	2	58	116
5	3	12	36
6	4	8	32
sum		235	229

From this, it is seen that $n = 235$ and $T = 229$. Thus, $\hat{S} = \frac{229}{235+229-1} = 0.4946004$ and $SE_{\hat{S}} = \sqrt{0.4946004 \left(0.4946004 - \frac{229-1}{235+229-2}\right)} = 0.02326041$. Thus, an approximate 95% confidence interval for S is $0.4946 \pm 2(0.0233)$ or $(0.4480, 0.5412)$. Furthermore, $\hat{Z} = -\log(0.4946004) = 0.7040051$ and $SE_{\hat{Z}} = \frac{0.02326041}{0.4946004} = 0.04702869$.

3.2 chapmanRobson() Function

More efficient calculation of the Chapman-Robson estimator can be made with `chapmanRobson()`. The function takes the exact same arguments and with information extracted with the same functions as described previously for `catchCurve()`. The Chapman-Robson estimator of S for the Tobin Harbor brook trout is obtained with

```
> thcr <- chapmanRobson(ct~age, data=bkt, ages2use=2:6)
> summary(thcr)
```

```
Intermediate Statistics
n=235; T=229
```

```
Estimates with Standard Errors
  Estimate Std. Err.
S 49.4600432 2.32607488
Z  0.7040051 0.04702937
```

```
> confint(thcr)
```

```
      95% LCI      95% UCI
S 44.9010202 54.0190662
Z  0.6118292  0.7961809
```

4 Catch Curve vs Chapman-Robson

Dunn *et al.* (2002) provided an excellent review of past examinations of the regression and Chapman-Robson methods and their own examination of the precision and bias properties of these two methods in the face of stochastic errors related to Z , number of fish at time of recruitment to the fishery, sampling, and ageing. Overall, they found that the Chapman-Robson estimator was most precise and least biased; however, the advantage over the regression method declined somewhat with increasing amounts of stochastic error and increasing values of Z .

Chapman and Robson (1960) proposed that the regression methods should exclude from the analysis all age-classes above the age where the catches fall below five individuals. Dunn *et al.* (2002) considered a modification of this suggestion where the cutoff catch value is one individual. In their analyses, Dunn *et al.* (2002) found that their modified suggestion generally performed better than the regression using all available age-classes, but that the suggestion of Chapman and Robson (1960) actually performed worse. Thus, if the regression method is used it is suggested that all age-classes beyond the first age-class where one or fewer individuals is observed should be excluded from the regression analysis.

The work of Dunn *et al.* (2002) also showed that, in the face of only stochastic sampling variability, the Chapman-Robson estimator was very slightly positively biased, primarily for larger values of Z , but only on the order of approximately 2-3%. In contrast, the regression estimator had a strong negative bias on the order of 20%. The modified (excluding all age-classes beyond where one or fewer individuals were observed) regression estimator had a negative bias on the order of 2-5% with the larger values occurring when Z was larger. These results suggest that estimates of Z with the regression method may be serious underestimates.

Finally, other methods for estimating total mortality rates have been proposed (e.g., Heincke (1913), Jackson (1939), Ssentengono and Larkin (1973)). However, various studies have shown that these methods perform less well than the Chapman-Robson and regression methods described here.

5 Comparing Z Between Groups

A common problem in fisheries research is to determine if two groups or populations of fish have the same instantaneous total mortality rate. For example, a researcher may want to determine if Z differs between sexes, between species, between lakes, or between treatments. These types of questions reduce to determining whether the slope computed from the catch curve differs between the groups. This type of question can be answered by including an indicator and interaction variable to the catch curve regression analysis. This section shows, through an example, how to use those methods to determine if the instantaneous mortality rate differs between fish captured from two different areas.

Moshenko and Low (1980) examined the statistics from the lake whitefish fishery in Great Slave Lake, Northwest Territories in 1979. The catch statistics were grouped into catches from five areas. In particular, they wanted to determine if the instantaneous total mortality differed between the two areas labeled as II and V. The catches on the **descending limb** were entered into a data frame and a variable with the natural log of catch was added with

```
> wf <- data.frame(age=c(10:13,12:16),
  ct = c(220,90,39,14,156,59,23,12,5),
  area = c(rep("II",4),rep("V",5)))
> wf$logct <- log(wf$ct)
> str(wf)

'data.frame':      9 obs. of  4 variables:
 $ age  : int  10 11 12 13 12 13 14 15 16
 $ ct   : num  220 90 39 14 156 59 23 12 5
 $ area : Factor w/ 2 levels "II","V": 1 1 1 1 2 2 2 2 2
 $ logct: num  5.39 4.5 3.66 2.64 5.05 ...
```

```
> wf
```

```
  age  ct area  logct
1  10 220  II 5.393628
2  11  90  II 4.499810
3  12  39  II 3.663562
4  13  14  II 2.639057
5  12 156  V 5.049856
6  13  59  V 4.077537
7  14  23  V 3.135494
8  15  12  V 2.484907
9  16   5  V 1.609438
```

The full model used to determine if the instantaneous mortality rate (i.e., slope) differs between the two areas is fit and analyzed with

```
> cc.comp <- lm(logct~age*area,data=wf)
> anova(cc.comp)
```

Analysis of Variance Table

```
Response: logct
      Df Sum Sq Mean Sq  F value    Pr(>F)
age      1 10.4119  10.4119 1326.3434 2.939e-07
area      1  2.2390   2.2390  285.2196 1.331e-05
age:area  1  0.0131   0.0131   1.6666  0.2532
Residuals 5  0.0393   0.0079
```

The p-value for the interaction term ($p=0.2532$) suggests that the slope and, thus, the instantaneous mortality rate does not differ significantly between the two areas. The fit of this model (Figure 6) can be observed with

```
> fitPlot(cc.comp,legend="topright")
```

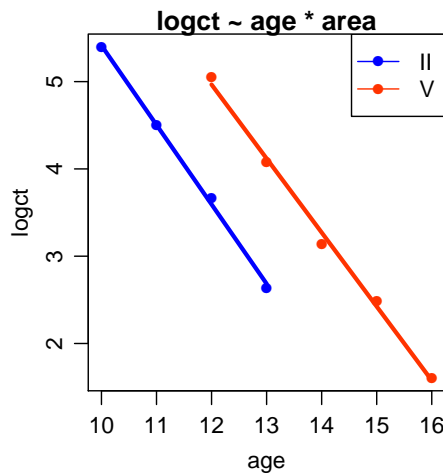


Figure 6. Catch curve comparison for Great Slave Lake lake whitefish captured in area-II and area-V.

It should be remembered that the sample size in these calculations is the number of ages on the descending limb of the catch curve. Samples size, and thus power, for most analyses is small. Therefore, relatively large differences in Z must be observed before statistical differences will be identified.

References

- Chapman, D. G. and D. S. Robson. 1960. The analysis of a catch curve. *Biometrics* 16:354–368. [10](#), [12](#), [16](#)
- Dunn, A., R. I. C. C. Francis, and I. J. Doonan. 2002. Comparison of the Chapman-Robson and regression estimators of Z from catch-curve data when non-sampling stochastic error is present. *Fisheries Research* 59:149–159. [12](#)
- Heincke, F. 1913. Investigations on the plaice. General report 1: The plaice fishery and protective measures. Preliminary brief summary of the most important points of the report. *Rapports et Proces-Verbaux des Reunions, Conseil International pour l'Exploration de la Mer*. 16:1–67. [12](#)
- Jackson, C. H. N. 1939. The analysis of an animal population. *Journal of Animal Ecology* 8:238–246. [12](#)
- Jensen, A. L. 1985. Comparison of catch-curve methods for estimation of mortality. *Transactions of the American Fisheries Society* 114:743–747. [11](#)
- Maceina, M. J. and P. W. Bettoli. 1998. Variation in largemouth bass recruitment in four mainstream impoundments on the tennessee river. *North American Journal of Fisheries Management* 18:998–1003. [9](#), [10](#)
- Moshenko, R. W. and G. Low. 1980. Data from the commercial fishery for lake whitefish, *Coregonus clupeaformis* (Mitchill), on Great Slave Lake, Northwest Territories, 1979. *Canadian Data Report of Fisheries and Aquatic Sciences*, No. 194 . [12](#)
- Quinlan, H. 1999. Biological characteristics of coaster brook trout at Isle Royale National Park, Michigan, 1996-98, U.S. Fish and Wildlife Service, Ashland Fishery Resources Office, Ashland, WI Office Report. [6](#)
- Robson, D. S. and D. G. Chapman. 1961. Catch curves and mortality rates. *Transactions of the American Fisheries Society* 90:181–189. [10](#)
- Ssentengono, G. W. and P. A. Larkin. 1973. Some simple methods of estimating mortality rates of exploited fish populations. *Journal of the Fisheries Research Board of Canada* 30:695–698. [12](#)

Appendices

A Chapman-Robson Estimate of S

The probability that a fish survives from one age-class to the next age-class is A . Thus, the probability that a fish lives to age-0, meaning that it died in its first year of life, is A . The probability that a fish lives to age-1, meaning that it survived its first year of life but died in its second year, is SA , where S is the annual survival rate, or $(1 - A)A$. Similarly, the probability that a fish lives to age-2 (survived each of its first two years of life but died in the third year) is SSA or $(1 - A)^2A$. Thus, the probability that a fish lives to the generic age x (survived x years to die in the last year) is given by $f(x) = S^x A$ or $f(x) = (1 - A)^x A$ for $x = 0, 1, 2, \dots$. Thus, the probability that an individual survives to age- x is a random variable that follows a geometric probability distribution with $p = A$.

If the fish in a sample can be thought of as independent then the likelihood function for the sample of n individuals is given by,

$$\begin{aligned} L(A) &= \prod_{i=1}^n (1 - A)^x A \\ &= \prod_{i=1}^n (1 - A)^x \prod_{i=1}^n A \\ &= (1 - A)^{\sum_{i=1}^n x} A^n \end{aligned}$$

The log-likelihood function is then,

$$l(A) = \left(\sum_{i=1}^n x \right) \log(1 - A) + n \ln(A)$$

The derivative of the log-likelihood function is,

$$\frac{dl(A)}{dA} = -\frac{\sum_{i=1}^n x}{1 - A} + \frac{n}{A}$$

This derivative is then set equal to zero and solved for A . The first steps in this are

$$\begin{aligned} \frac{dl(A)}{dA} &= 0 \\ -\frac{\sum_{i=1}^n x}{1 - A} + \frac{n}{A} &= 0 \\ \frac{n}{A} &= \frac{\sum_{i=1}^n x}{1 - A} \\ \frac{1 - A}{A} &= \frac{\sum_{i=1}^n x}{n} \end{aligned}$$

The usual solution for a geometric distribution is to substitute $\bar{x} = \frac{\sum_{i=1}^n x}{n}$ on the RHS and solve for A . However, to more quickly get to the solution given by [Chapman and Robson \(1960\)](#) it is beneficial to substitute $T = \sum_{i=1}^n x$ on the RHS and $S = 1 - A$ and $A = 1 - S$ on the LHS and then solve for S ,

$$\begin{aligned}\frac{S}{1-S} &= \frac{T}{n} \\ Sn &= T - ST \\ Sn + ST &= T \\ S(n+T) &= T \\ S &= \frac{T}{n+T}\end{aligned}$$

Thus, the maximum likelihood estimator for S is $\frac{T}{n+T}$.

It is beyond the scope of this vignette to prove this but this estimator of S is slightly biased and the estimator derived by [Chapman and Robson \(1960\)](#), which subtracts one from the denominator, provides an unbiased estimator of S that is also minimum variance. Typically the values of n and T will be very large relative to 1 and, thus, the subtraction of 1 in the denominator has very little effect on the estimator.

Reproducibility Information

Version Information

- **Compiled Date:** Thu Sep 29 2011
- **Compiled Time:** 9:43:13 PM

Files

- [NoWeb Source \(.Rnw\) file](#)
- [R Script \(.R\) file](#)

R Information

- **R Version:** R version 2.13.1 (2011-07-08)
- **System:** Windows, i386-pc-mingw32/i386 (32-bit)
- **Base Packages:** base, datasets, graphics, grDevices, methods, splines, stats, tcltk, utils
- **Other Packages:** ascii_2.0, doBy_4.4.0, FSA_0.2-6, FSAdata_0.1-1, gdata_2.8.2, Hmisc_3.8-3, lattice_0.19-33, lme4_0.999375-41, MASS_7.3-14, Matrix_0.9996875-3, miscOgle_0.1-0, multcomp_1.2-7, mvtnorm_0.9-9991, NCStats_0.2-6, nlstools_0.0-11, plotrix_3.2-4, plyr_1.6, quantreg_4.71, R2HTML_2.2, reshape_0.8.4, sciplot_1.0-9, snow_0.3-7, SparseM_0.89, survival_2.36-9, svSocket_0.9-51, TinnR_1.0.3
- **Loaded-Only Packages:** car_2.0-11, cluster_1.14.0, gplots_2.10.1, grid_2.13.1, gtools_2.6.2, nlme_3.1-102, stats4_2.13.1, svMisc_0.9-63, TeachingDemos_2.7, tools_2.13.1
- **Required Packages:** FSA, NCStats and their dependencies (car, FSAdata, gdata, gplots, gtools, Hmisc, MASS, nlme, nlstools, plotrix, quantreg, reshape, sciplot, stats, tcltk, TeachingDemos)