fishR Vignette - Catch Curve Estimates of Mortality

Dr. Derek Ogle, Northland College December 16, 2013

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 catch curve regression (see Section 2.5) and Chapman-Robson (see Section 3) methods used to analyze this decline are collectively called catch-curve methods and form the topic of this vignette.

This vignette requires functions in the FSA package maintained by the author. This package is loaded into R with

```
> library(FSA)
```

1 Data Requirements

In a population that is closed to emigration or immigration, the annual mortality rate (A) between two times is equal to the number of deaths during the time period divided by the population size at the start of the time period, or

$$A = \frac{N_t - N_{t+1}}{N_t} = 1 - \frac{N_{t+1}}{N_t}$$

Unfortunately, it is usually not possible to know the number of fish in a population. However, if the catch of fish (C) is proportional to the size of the population, i.e., $C_t = vN_t$, then algebra quickly shows that

$$A = \frac{C_t - C_{t+1}}{C_t} = 1 - \frac{C_{t+1}}{C_t}$$

Thus, the mortality of a cohort of fish can be estimated from knowing the catches of fish at various times. In many fisheries, fisheries scientists "record time" by estimating the age of fish. Thus, the number of fish captured at varous ages, i.e., catch-at-age data, can be used to estimate mortality rates of fish populations.

In the statistical literature, longitudinal data is data that occurs when multiple samples are taken from the same group of individuals over time. The catches of fish from the same cohort over time is an example of longitudinal data. Longitudinal fisheries data takes many years to collect, which can be very costly and impart a long time-lag in management decisions.

Catch data from a single year across many cohorts of fish will be identical to longitudinal data of a single cohort if each cohort sampled began with the same number of fish (i.e., recruitment is constant) and if the mortality rate is constant across all ages and years. For example, the catches-at-age of the hypothetical 2002 and 2006 cohorts is shown on diagonals in Table 1 and are equal to the catch-at-age for fish in the 2009 capture year. The catch in a single year is called cross-sectional data because it "crosses" several cohorts of fish.

2 Catch Curve Regression Methods

Annual mortality can be estimated from catch data for two ages as shown above. However, fisheries scientists prefer estimates that are more synthetic, i.e., based on more ages. The two most common methods for computing synthetic estimates of mortality rates are the catch-curve regression and Chapman-Robson methods. The regression method is discussed in this section, whereas the Chapman-Robson method is discussed in Section 3.

Table 1. The hypothetical catch of fish by age and capture year. The longitudinal catch of the 2002 and the partial 2006 year-classes of fish are shown by the two sets of diagonal cells highlighted in dark grey. The cross-sectional catch in the 2009 capture year is shown by the column of cells highlighted in light grey. All data were modeled with (2) assuming that $N_0 = 500$, $Z = -log(0.7)$, and $v = 0.1$.

	Capture Year											
Age	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
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

2.1 Regression Model

The decline in individuals with age can be theoretically modeled with a modified continuous exponential population model. Because the population is closed with the exception of mortality, the instantaneous population growth rate parameter in the exponential population model is replaced with an instantaneous total mortality parameter (Z). Thus, the modified model is

$$N_t = N_0 e^{-Zt} \tag{1}$$

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

The catch of fish at age t is proportional to the number of fish of age t, or $C_t = vN_t$, as mentioned previously[1]. This is rearranged to show that the relationship between population size and catch is $N_t = \frac{C_t}{v}$ and substituted into (1) to reveal

$$\begin{aligned} \frac{C_t}{v} &= N_0 e^{-Zt} \\ C_t &= vN_0 e^{-Zt} \end{aligned} \tag{2}$$

In contrast to (1) the variables in (2) (catch-at-age, C_t, and age, t) are directly observable. The shape of (2) (Figure 1-Left) follows the expected exponential decline. As is typical with exponential models, natural logarithms of both sides of (2) yields

$$log(C_t) = log(vN_0) - Zt \tag{3}$$

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 (3) is that the negative of the slope is Z. Thus, the negative of the slope of the linear 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.
The (1) can also be recast by assuming that catch-at-age (C_t) is proportionately related to the number-at-age and the amount of effort expended to catch those fish (i.e., E_t). Thus, $C_t = qE_tN_t$, where q represents a

[1] The v represents a constant proportion of the population that is "vulnerable" to the fishery.

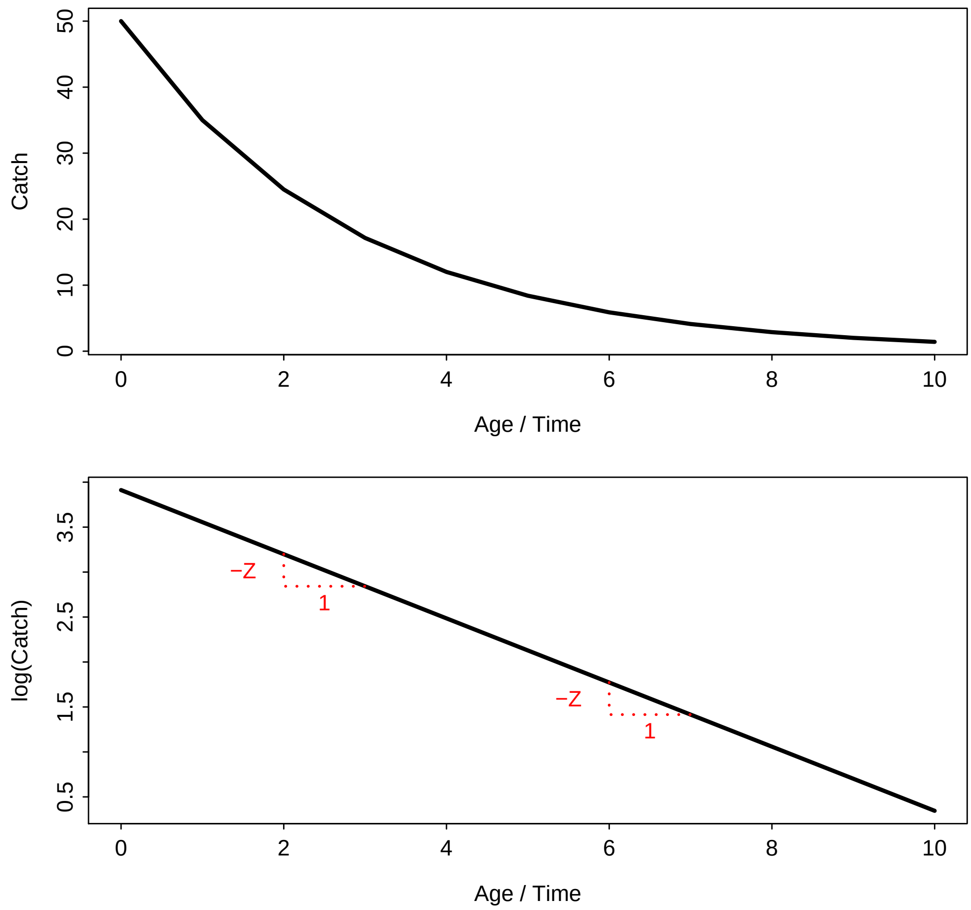

Figure 1. Ideal plots of catch versus age (Left) and the natural log of catch versus age (Right) for a single cohort of fish. The right graph is called a longitudinal catch curve. The change in $log(C_t)$ for a unit change in t is emphasized on the catch curve to reinforce the idea that the slope of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

constant proportion of the population captured by one unit of effort. This model can be rearranged to show that the relationship between population size and catch-per-unit-effort is $N_t = \frac{1}{q}\frac{C_t}{E_t}$. This is substituted into (1) and simplified to reveal

$$\begin{aligned} \frac{1}{q}\frac{C_t}{E_t} &= N_0 e^{-Zt} \\ \frac{C_t}{E_t} &= qN_0 e^{-Zt} \end{aligned} \tag{4}$$

Again, the variables in (4) (catch-per-unit-effort, $\frac{C_t}{E_t}$, and age, t) are directly observable. Furthermore, natural logarithms of both sides of (4) yields

$$log(\frac{C_t}{E_t}) = log(qN_0) - Zt \tag{5}$$

which again 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 variable can be either catch or catch-per-unit-effort data. The specifics of this regression methodology are discussed in Section 2.5.

2.2 Characteristics

All catch curves 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 yet 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.

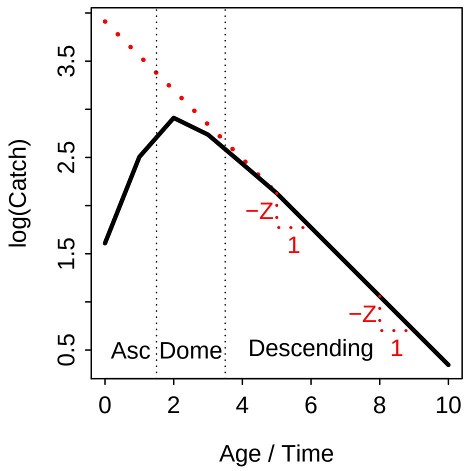

Figure 2. Idealized catch curve (plot of the natural log of catch versus age) illustrating the ascending, domed, and descending portions. The red dotted line represents the idealized catch curve if all age-classes were fully recruited to the fishery.

The domed portion of the catch-curve generally consists of age-classes of fish that are nearly, but not completely, 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"[2]. 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 (3) to the catches of fish in the ages corresponding only to the descending portion of the catch curve. There is some debate about how the descending limb is defined in practice. Additionally, the portion of the descending limb corresponding to the older ages is often poorly represented in the catch data because these individuals are relatively rare. At times, some catch data for older ages may be ignored. Specific guidelines for identifying the start of the descending limb and which, if any, older ages should be excluded will be given in Section 2.5 and Section 3.

2.3 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 regression method using longitudinal and cross-sectional data 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,

- "Accurate Ages" – The fish in a sample can be accurately assigned an age. In longitudinal data, this means that you can follow a cohort through time.

Additionally, if cross-sectional data is being used then it is assumed that there is constant recruitment, i.e., the initial number of individuals is the same for each cohort of fish.

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).

2.4 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 "millisecond" of time? Fortunately, the instantanous 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

[2]This 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.

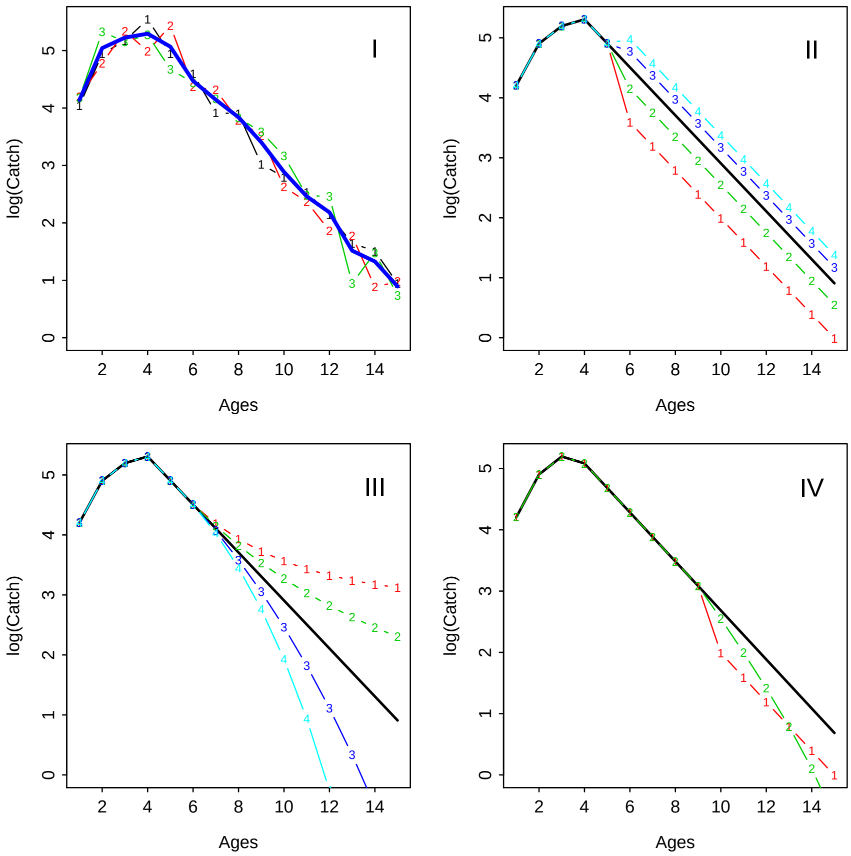

Figure 3. Simulated catch curves to illustrate shapes when assumptions are violated. Each simulation (i.e., plot), unless otherwise, noted uses $N_0 = 1000$ for each year-class, $Z = 0.40$, and incomplete recruitment until age-4 and then constant recruitment for subsequent ages. In simulation I, a coefficient 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.

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

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.

2.5 Basic Regression Method in R

The regression method for estimating Z will be demonstrated in this section using the catches-at-age for Brook Trout (*Salvelinus fontinalis*) captured by the U.S. Fish and Wildlife Service from 1996-1998 in Tobin Harbor of Isle Royale (data shown in (Table 2); Quinlan (1999.)). The Tobin Harbor Brook Trout catch-at-age data[3] were entered into a data frame and the natural log of the catch data was computed with[4]

```
> ( 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
> 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 ...
```

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 catch-curve plot (Figure 4) was produced with `plot()` as follows[5]

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

From the initial catch curve plot it was determined that the descending limb consisted of ages 2 through 6. Thus, a new data frame which contains only the information for these ages on the descending limb should be extracted with `Subset()`, which requires the original data frame as the first argument and a conditioning statement as the second argument. Thus, the data for ages 2 to 6 were isolated with

[3]These data can also be obtained with `data(BrkTrtTH)`.
[4]The "extra" parentheses around this command simply force the result that is saved to an object to be printed to the console.
[5]Note that `pch=19` simply forces `plot()` to use filled circles as points.

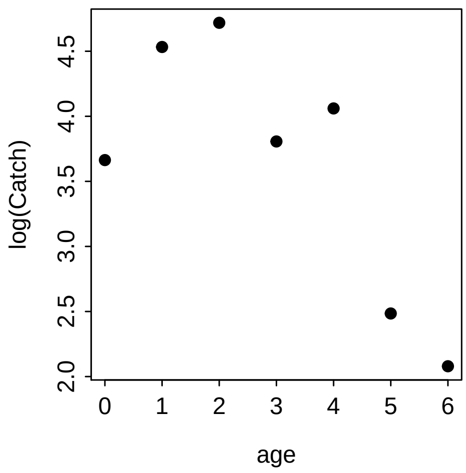

Figure 4. Catch curve for the Tobin Harbor Brook Trout data.

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

  age  ct logct
3   2 112 4.718
4   3  45 3.807
5   4  58 4.060
6   5  12 2.485
7   6   8 2.079
```

A linear regression model is fit with `lm()` using a formula of the form `y~x` as the first argument and the appropriate data frame in `data=`. The result should be saved to an object so that other information can be extracted from the results. For example, the fitted coefficients are extracted with `coef()` and corresponding confidence intervals with `confint()`. Thus, the linear regression of log catch-at-age on age for only the descending limb is fit and then seen with

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

(Intercept)         age
       6.07       -0.66

> confint(cc)

             2.5 %  97.5 %
(Intercept)  4.224  7.9162
age         -1.095 -0.2248
```

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. Thus, from these results it is seen that $\hat{Z}$ is 0.66 with a 95% confidence interval between 0.22 and 1.10.

The estimated annual mortality rate ($\hat{A}$), with confidence interval, can be computed by extracting the slope information from the previous results and using $A = 1 - e^{-Z}$. For example,

```
> ( A <- 1-exp(coef(cc)[2]) )

   age
0.4831

> ( A.ci <- 1-exp(confint(cc)[2,]) )
```

```
 2.5 % 97.5 %
0.6655 0.2013
```

Thus, the estimated annual mortality rate is 48.3% with an approximate 95% confidence interval between 20.1% and 66.6%.

2.6 `catchCurve()` Convenience Function

The catch curve analysis depicted in the previous section can be more efficiently obtained with `catchCurve()`, which 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 values[6], 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     0.66     0.1367   4.827  0.01695
A    48.31         NA      NA       NA

> confint(thcc)

  95% LCI 95% UCI
Z  0.2248   1.095
A 20.1337  66.551

> plot(thcc)
```

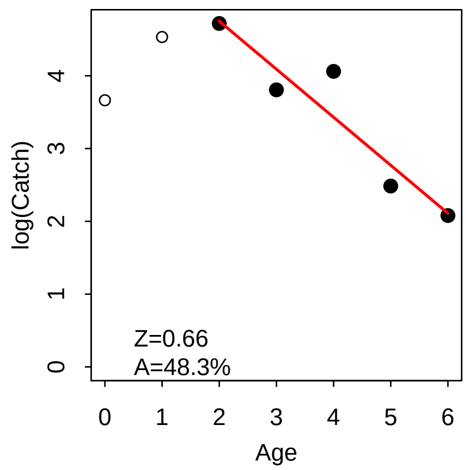

Figure 5. Catch curve for the Tobin Harbor Brook Trout data.

[6]The `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.

2.7 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 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    0.643     0.1417   4.536  0.02005
A   47.430         NA      NA       NA

> confint(thcc2)

  95% LCI 95% UCI
Z  0.1919   1.094
A 17.4634  66.516
```

2.8 Which Ages?

Smith *et al.* (2012), in a simulation study, suggested that the first age to be included on the descending limb is the age where the peak catch occurred (in contrast to one year after the peak catch). Chapman and Robson (1960) proposed that the regression methods should exclude 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. Smith *et al.* (2012), however, showed that the weighted regression method was less biased than the unweighted regression method no matter what decision rule was used for the oldest age in the analysis. Thus, if the regression method is used (see Section 4) it is suggested that a weighted regression that uses all age-classes older than and including the age with the maximum catch should be used.

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

$$\hat{S} = \frac{T}{n + T - 1} = \frac{\bar{T}}{1 + \bar{T} - \frac{1}{n}} \tag{6}$$

where n is the total number of fish observed on the descending limb of the catch curve, T is the total recoded age of fish on the descending limb of the catch curve, and $\bar{T}$ is the mean recoded age of fish on the descending limb of the catch curve (i.e., $\bar{T} = \frac{T}{n}$). 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

$$SE_{\hat{S}} = \sqrt{\frac{T}{n+T-1}\left(\frac{T}{n+T-1} - \frac{T-1}{n+T-2}\right)} = \sqrt{\hat{S}\left(\hat{S} - \frac{T-1}{n+T-2}\right)} \tag{7}$$

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

$$SE_{\hat{S}} = \sqrt{\frac{\hat{S}(1-\hat{S})^2}{n}} \tag{8}$$

The Chapman-Robson estimate of S can be transformed into an estimate of Z through the relationship $S = e^{-Z}$, i.e., $\hat{Z} = -log(\hat{S})$ with a large sample approximation of $SE_{\hat{Z}}$ (Jensen 1985) as

$$SE_{\hat{Z}} = \frac{SE_{\hat{S}}}{\hat{S}} \tag{9}$$

Hoenig *et al.* (1983), however, have shown that these estimates are slightly biased and that an unbiased estimate of Z is obtained with

$$\hat{Z} = -log(\hat{S}) - \frac{(n-1)(n-2)}{n(T+1)(N+T-1)} \tag{10}$$

with the large sample approximation of $SE_{\hat{Z}}$ as

$$SE_{\hat{Z}} = \frac{1-e^{-\hat{Z}}}{\sqrt{ne^{-\hat{Z}}}} \tag{11}$$

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 shown in (Table 3). From that, it is seen that $n = 235$ and $T = 229$. Thus,

$$\hat{S} = \frac{229}{235+229-1} = 0.4946004 \tag{12}$$

$$SE_{\hat{S}} = \sqrt{0.4946004\left(0.4946004 - \frac{229-1}{235+229-2}\right)} = 0.02326041 \tag{13}$$

An approximate 95% confidence interval for S is $0.4946 \pm 1.96(0.0233)$ or (0.4490,0.5402). Furthermore,

$$\hat{Z} = -log(0.4946004) - \frac{(235-1)(235-2)}{235(229+1)(235+229-1)} = 0.7018264 \tag{14}$$

$$SE_{\hat{Z}} = \frac{1-e^{-0.7018264}}{\sqrt{235 * e^{-0.7018264}}} = 0.04672751 \tag{15}$$

An approximate 95% confidence interval for Z is $0.7018 \pm 1.96(0.0467)$ or (0.6102,0.7934).

Table 3. Cross-sectional total catch-at-age of Tobin Harbor Brook Trout in fyke nets, 1996-1998, modified to illustrate the calculations of the Chapman-Robson method.

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

3.2 `chapmanRobson()` Convenience 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  49.4600   2.32607
Z   0.7018   0.04673
> confint(thcr)
  95% LCI 95% UCI
S 44.9010 54.0191
Z  0.6102  0.7934
```

3.3 Which Ages?

Smith *et al.* (2012) suggest that the Chapman-Robson method should use all ages after the age where the peak catch occurred (i.e., all ages beginning with the first age after the age with the peak catch).

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.

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.

Smith *et al.* (2012) provided another study of precision between the regression and Chapman-Robson methods with the addition of consideration of different definitions of the descending limb of the catch curve. They found that the Chapman-Robson method using all ages after the age with the peak catch and the weighted regression using all ages after and including the age with the peak catch performed similarly. However, they suggest using the Chapman-Robson method because it is based on a statistical foundation and has a generally smaller variance, whereas the weighting procedure in the regression method is *ad hoc*. Finally, Smith *et al.* (2012) conclude that the unweighted regression should not be used.

Finally, other methods for estimating total mortality rates have been proposed (e.g., Heincke (1913), Jackson (1939), Ssentengono and Larkin (1973)). However, various studies (including Smith *et al.* (2012)) have shown that these methods perform less well than the Chapman-Robson and regression methods described here.

References

Chapman, D. G. and D. S. Robson. 1960. The analysis of a catch curve. Biometrics 16:354–368. 10, 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. 10, 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. 13

Hoenig, J. M., W. D. Lawing, and N. A. Hoenig. 1983. Using mean age, mean length and median length data to estimate the total mortality rate. Ices-cm-1983/d:23, International Council for the Exploration of the Sea, Copenhagen, Denmark. 11

Jackson, C. H. N. 1939. The analysis of an animal population. Journal of Animal Ecology 8:238–246. 13

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. 10

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. 7

Robson, D. S. and D. G. Chapman. 1961. Catch curves and mortality rates. Transactions of the American Fisheries Society 90:181–189. 10

Smith, M. W., A. Y. Then, C. Wor, G. Ralph, K. H. Pollock, and J. M. Hoenig. 2012. Recommendations for catch-curve analysis. North American Journal of Fisheries Management 32:956–967. 10, 12, 13

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. 13

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 $f(x) = S^xA$ or $f(x) = (1-A)^xA$ for $x = 0, 1, 2, \ldots$. 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,

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

The log-likelihood function is then,

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

The derivative of the log-likelihood function is,

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

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

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

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 substitue $T = \sum_{i=1}^{n} x$ 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:** Mon Dec 16 2013
- **Compiled Time:** 9:45:22 PM
- **Code Execution Time:** 2.07 s

R Information

- **R Version:** R version 3.0.2 (2013-09-25)
- **System:** Windows, i386-w64-mingw32/i386 (32-bit)
- **Base Packages:** base, datasets, graphics, grDevices, methods, stats, utils
- **Other Packages:** FSA_0.4.3, knitr_1.5.15
- **Loaded-Only Packages:** bitops_1.0-6, car_2.0-19, caTools_1.16, cluster_1.14.4, evaluate_0.5.1, formatR_0.10, Formula_1.1-1, gdata_2.13.2, gplots_2.12.1, grid_3.0.2, gtools_3.1.1, highr_0.3, Hmisc_3.13-0, KernSmooth_2.23-10, lattice_0.20-24, MASS_7.3-29, multcomp_1.3-1, mvtnorm_0.9-9996, nlme_3.1-113, nnet_7.3-7, plotrix_3.5-2, quantreg_5.05, sandwich_2.3-0, sciplot_1.1-0, SparseM_1.03, splines_3.0.2, stringr_0.6.2, survival_2.37-4, tools_3.0.2, zoo_1.7-10
- **Required Packages:** FSA and its dependencies (car, gdata, gplots, Hmisc, knitr, multcomp, nlme, plotrix, quantreg, sciplot)