Some reflections on the possibility of estimating predation and digestion parameters from stomach data

Knud P. Andersen

The Danish Institute for Fisheries and Marine Research, Charlottenlund Castle, DK-2920 Charlottenlund, Denmark

Summary

The possibility of estimating parameters related to predation and digestion has been examined by simulating stomach data sets adhibiting assumptions on predation, digestion and environmental prey distributions. All estimations were performed by using non-linear regression. Some light was thrown on certain difficulties in the estimation.

In Andersen 1982 an attempt was made to demonstrate the theoretical possibility of determining predation and digestion parameters from fish stomach data, when certain hypotheses on predation and digestion are assumed. The methods are however only valid for data without stochastic components, and it can be of some interest to give a summary of results obtained by studying some simulated data where underlying hypotheses are fulfilled.

In two examples we have one cod-like predator Pr sampled for 6 weight sizes W = 10 g, 40 g, 160 g, 640 g, 2560 g, and 10240 g. Pr is preying on several prey species: pr1, pr2, ... some of which have exactly known environmental weight distributions and relative abundances.

The underlying definitions and hypotheses can be summarized like this:

1. The environmental w (weight) distributions N(w, pr) are given by:

$$\begin{split} N(w, pr) * dw &= \widetilde{N}^*(pr) * \exp(-(\log(w) - \mu(pr))^2/2/\tau(pr)^2) * dw \\ SQRT(2 * \pi * \tau(pr)^2)/w \text{ (log-normal distributions).} \end{split}$$

2. The size preference functions prfu(w, pr, W) are given by:

 $prfu(w, pr, W) = \rho(pr) * \exp(-(\log(W/w) - \eta(pr))^2/2/\sigma(pr)^2)$

3. The 'available' food $\phi(W)$ is given as:

$$\phi(W) = \sum_{\text{all } pr} \bigg\{ \int_{0}^{\infty} [prfu(w, pr, W) * N(w, pr) * dw] \bigg\}.$$

4. The feeding level f(W) is determined by:

 $f(\mathbf{W}) = \phi(\mathbf{W})/(\phi(\mathbf{W}) + q * \mathbf{W}^r).$

5. The digestion parameter $\kappa(W)$ is of the form:

 $\kappa(W) = \kappa_0 * W^{\kappa_1}.$

6. Assumptions on $t_0(W)$, the timespan from ingestion, where a prey specimen can be recognized in the stomach contents:

 $\kappa(W) * t_0(W)$ is a constant independent of W and prey species.

7. Equilibrium conditions:

 $\kappa(W) * (\text{weight of recognizable stomach contents})/(1 - \exp(-\kappa(W) * t_0(W))) = \text{consumption} = f(W) * h * W^m$.

Here W is the weight of the predator, $\widetilde{N}_{st}^*(pr)$, $\mu(pr)$, and $\tau(pr)^2$ are prey species dependent constants, $\rho(pr)$, $\eta(pr)$ and $\sigma(pr)^2$ are predator species and prey species dependent constants, h, q, r, κ_0 , and κ_1 are predator species dependent constants, and m is the ingestion exponent.

As only one predator species is used the predator species dependence has been omitted.

All data sets are simulations of random encounters between predator and prey controlled by a Poisson process $NN(W, pr, t_1, t_2)$ giving the number of encounters in the time interval (t_1, t_2) . NN is specified by:

probability(NN(W, pr, t, t + dt) = 1) = $\widetilde{N}^*(pr) * h * W^m * dt/(\phi(W) + q * W')$, probability(NN(W, pr, t, t + dt) = 0) = 1 - probability(NN(W, pr, t, t + dt) = 1).

The probability for an encounter to result in an ingestion is determined by the preference function prfu(w(pr), pr, W), where w(pr) is a prey weight drawn at random from the given environmental distribution.

A simulation along these lines is a stochastic edition of formula (i), p.1 in Andersen 1982.

A consequence of the assumptions is that the stomach prey weights at ingestion will be log-normally distributed:

$$N_{st}(w, pr, W) * dw = N_{st}^{*}(pr, W) * \exp(-(\log(w) - \mu(pr, W))^{2}/2/\tau(pr, W)^{2}) * dw/SQRT(2 * \pi * \tau(pr, W)^{2})/w$$

with parameters:

 \sim .

$$\begin{split} N_{st}^{*}(pr, W) &= \rho(pr) * \tilde{N}^{*}(pr) * h * W^{m-\kappa_{1}} * SQRT(\tau(pr, W)^{2})\kappa(W) * \\ t_{0}(W) * \exp(-(\mu(pr) + \eta(pr) - \log(W))^{2}/2(\sigma(pr)^{2} + \tau(pr)^{2}))/\kappa_{0}/(\phi(W) + qW^{r})/SQRT(\tau(pr)^{2}) \\ \tau(pr, W)^{2} &= \sigma(pr)^{2} * \tau(pr)^{2}/(\sigma(pr)^{2} + \tau(pr)^{2}) \\ \mu(pr, W) &= (\mu(pr) * \sigma(pr)^{2} - (\eta(pr) - \log(W)) * \tau(pr)^{2})/(\sigma(pr)^{2} + \tau(pr)^{2}). \end{split}$$

Example 1

This example is a stochastic version of example 4 *in* Andersen 1982. Pr is preying on 4 species pr1, pr2, pr3, and pr4 (and only these) with exactly known environmental parameters:

	μ	τ^2	\widetilde{N}^*
pr1	-5	1	8000
pr2	-2.5	1	4000
pr3	0	1	3000
pr4	2.5	1	250

Simulations used the parameter values:

$$\begin{split} m &= 0.56; \ h = 35 \ \mathrm{g}^{1-m}/\mathrm{year}; \ r = -0.15; \ q = 500 \ \mathrm{g}^{1-r}; \\ \kappa_1 &= -0.25; \ \kappa_0 = 1519 \ \mathrm{g}^{-\kappa_1}/\mathrm{year}; \\ \kappa(W) * t_0(W) &= 0.75; \\ \rho(pr1) &= 0.9; \ \sigma(pr1)^2 = 1 \quad ; \quad \eta(pr1) = 7; \\ \rho(pr2) &= 1 \quad ; \ \sigma(pr2)^2 = 2 \quad ; \quad \eta(pr2) = 6; \\ \rho(pr3) &= 0.8; \ \sigma(pr3)^2 = 0.8; \ \eta(pr3) = 5; \\ \rho(pr4) &= 0.6; \ \sigma(pr4)^2 = 0.4; \ \eta(pr4) = 4. \end{split}$$

500 stomach contents were simulated for each of the 6 W groups delivering 4 * 6 tables, each giving data relevant for the parameter estimation. As an example the table for W = 640 g and *pr4* is shown here:

Predator weight W = 640 g. 500 stomachs. Recognizable pr4 specimens.

		log(w) as found in stomachs							
		1.33	1.67	2.00	2.33	2.67	3.00	Total	
	1.67	1	3	0	0	0	0	4	
log(w) 2.00 at 2.33	0	3	2	0	0	0	5		
	0	0	5	3	0	0	8		
inges-	2.67	0	0	0	3	3	0	6	
tion 3.0	3.00	0	0	0	0	0	4	4	
	3.33	0	0	0	0	0	1	1	
	Total	1	6	7	6	3	5	28	
$Obs(\widetilde{N}_{st}^{**}(pr4, 640 g))$					g))				

Obs $(\widetilde{N}_{st}^*(pr4, 640 \text{ g})) = 0.056 = \text{Obs}(\widetilde{N}_{st}^{**}(pr4, 640 \text{ g}))/500 \text{Obs}(\mu(pr4, 640 \text{ g})) = 2.38095.$

Obs $(\tau(pr4, 640 \text{ g})^2) = 0.20238$ (with Sheppard's correction). Digested weight = 290.2888 g = 500 * Obs(stctr(pr4, 640 g)).

Obs(stctr(pr4, 640 g)) = 0.580578 g.

(stctr stands for recognizable stomach contents).

The directly measured weights are the weights of the prey specimens as found, but it is assumed that an exact method for transforming digested weight to weight at ingestion is available.

If we define the relative stomach abundance of pr, R(pr, W), as:

$$R(pr, W) = \widetilde{N}_{st}^{**}(pr, W) * \widetilde{N}^{*}(pr2) / \widetilde{N}_{st}^{**}(pr2, W) / \widetilde{N}^{*}(pr),$$

and

$$stctr(W) = stctr(pr1, W) + stctr(pr2, W) + stctr(pr3, W) + stctr(pr4, W)$$

the stomach contents give us the observations:

Obs
$$(\tau(pr1, 10 \text{ g})^2)$$
, Obs $(\tau(pr1, 40 \text{ g})^2)$,
Obs $(\tau(pr1, 160 \text{ g})^2)$, Obs $(\tau(pr1, 640 \text{ g})^2)$,
Obs $(\tau(pr1, 2560 \text{ g})^2)$, Obs $(\tau(pr1, 10240 \text{ g})^2)$,,
Obs $(\tau(pr4, 10 \text{ g})^2)$, Obs $(\tau(pr4, 40 \text{ g})^2)$,
Obs $(\tau(pr4, 160 \text{ g})^2)$, Obs $(\tau(pr4, 640 \text{ g})^2)$,
Obs $(\tau(pr4, 2560 \text{ g})^2)$, Obs $(\tau(pr4, 10240 \text{ g})^2)$,
Obs $(\mu(pr1, 10 \text{ g}))$, Obs $(\mu(pr1, 40 \text{ g}))$, ..., Obs $(\mu(pr1, 10240 \text{ g}))$,,
Obs $(\mu(pr4, 10 \text{ g}))$, Obs $(\mu(pr4, 40 \text{ g}))$, ..., Obs $(\mu(pr4, 10240 \text{ g}))$,
Obs $(R(pr1, 10 \text{ g}))$, ..., Obs $(R(pr1, 10240 \text{ g}))$,
Obs $(R(pr3, 10 \text{ g}))$, ..., Obs $(R(pr3, 10240 \text{ g}))$,
Obs $(R(pr4, 10 \text{ g}))$, ..., Obs $(R(pr4, 19240 \text{ g}))$,
Obs $(R(pr4, 10 \text{ g}))$, ..., Obs $(R(pr4, 19240 \text{ g}))$,

i.e. a set of 72. The assumptions make it possible to express each observation as a non-linear regression in the independent variables $\tau(pr1)^2, ..., \tau(pr4)^2, \mu(pr1), ..., \mu(pr4), \tilde{N}^*(pr1), ..., \tilde{N}^*(pr4)$, and W with the regression coefficients:

 $\begin{array}{l} \sigma(pr1)^2, \, \sigma(pr2)^2, \, \sigma(pr3)^2, \, \sigma(pr4)^2, \\ \eta(pr1), \, \eta(pr2), \, \eta(pr3), \, \eta(pr4), \\ \rho(pr1), \, \rho(pr3), \, \rho(pr4), \\ \kappa_0/h/(1 - \exp(-\kappa(W) * t_0(W))) \; (= C_1), \\ \kappa_1 - m \; (= C_2), \\ q * \kappa_0/h/(1 - \exp(-\kappa(W) * t_0(W))) \; (= C_3), \\ r + \kappa_1 - m \; (= C_4). \end{array}$

 $(\rho(pr2)$ is not among the regression coefficients. The reason is that one ρ can be set arbitrarily, and $\rho(pr2)$ is taken as 1, or expressed in another way: $\rho(pr)$ stands for $\rho(pr)/\rho(pr2)$). E.g.:

 $\begin{aligned} & Obs(\tau(pr1, 160 \text{ g})^2) = \sigma(pr1)^2 * \tau(pr1)^2 / (\sigma(pr1)^2 + \tau(pr1)^2) + \varepsilon_1(pr1, 160 \text{ g}) \\ & Obs(\mu(pr3, 40 \text{ g})) = (\mu(pr3) * \sigma(pr3)^2 - (\eta(pr3) - \log(40)) * \tau(pr3)^2) / (\sigma(pr3)^2 + \tau(pr3)^2) + \varepsilon_2(pr3, 40 \text{ g}) \\ & Obs(R(pr4, 640 \text{ g})) = \rho(pr4) * SQRT[(1 + \tau(pr2)^2/\sigma(pr2)^2) / (1 + \tau(pr4)^2 / \sigma(pr4)^2)] * \exp[(\mu(pr2) + \eta(pr2) - \log(640))^2 / 2 / (\sigma(pr2)^2 + \tau(pr2)^2)) \\ & - (\mu(pr4) + \eta(pr4) - \log(640))^2 / 2 / (\sigma(pr4)^2 + \tau(pr4)^2)] + \varepsilon_3(pr4, 640 \text{ g}) \end{aligned}$

$$\begin{aligned} \text{Obs}(stctr(10 \text{ g})) &= 1/[\kappa_0 * 10^{\kappa_1 - m} / h / (1 - \exp(-\kappa(10 \text{ g}) * t_0(W))) + q * \kappa_0 * \\ 10^{(r+\kappa_1 - m)} / \phi(10 \text{ g}) / h / (1 - \exp(-\kappa(10 \text{ g}) * t_0(W)))] + \varepsilon_4(10 \text{ g}). \end{aligned}$$

The ε 's are random components approximately normally distributed with approximate zero means. The variances of the ε 's are however not equal, and the covariances not zero. The covariance matrix for the observations is therefore needed for the estimation of the regression coefficients. The assumptions make it possible to find a good estimate of the covariance matrix using observations and independent variables only. A few examples are given here.

$$\begin{aligned} \operatorname{var}(\operatorname{Obs}(\tau(pr, W)^2)) &= 2 * (\operatorname{Obs}(\tau(pr, W))^2))^2 / (\operatorname{Obs}(N_{st}^{**}(pr, W)) - 1) \\ \operatorname{var}(\operatorname{Obs}(R(pr, W))) &= \widetilde{N}^*(pr2)^2 * \operatorname{Obs}(\widetilde{N}_{st}^{**}(pr, W)) / \widetilde{N}^*(pr)^2 / \\ \operatorname{Obs}(\widetilde{N}_{st}^{**}(pr2, W))^2 + \widetilde{N}^*(pr2)^2 * \operatorname{Obs}(\widetilde{N}_{st}^{**}(pr, W))^2 / \operatorname{Obs}(\widetilde{N}_{st}^{**}(pr2, W))^3 / \widetilde{N}^*(pr)^2 \\ \operatorname{cov}(\operatorname{Obs}(\mu(pr, W')), \operatorname{Obs}(stctr(W''))) &= \operatorname{Obs}(\tau(pr, W')^2) * \\ \operatorname{Obs}(stctr(pr, W')) / \operatorname{Obs}(\widetilde{N}_{st}^{**}(pr, W')) \\ \operatorname{if} W' &= W'' \text{ else } 0 \\ \operatorname{cov}(\operatorname{Obs}(R(pr1, W)), \operatorname{Obs}(R(pr3, W))) &= \operatorname{Obs}(\widetilde{N}_{st}^{**}(pr1, W)) * \\ \operatorname{Obs}(\widetilde{N}_{st}^{**}(pr3, W)) * \widetilde{N}^*(pr2)^2 / \operatorname{Obs}(\widetilde{N}_{st}^{**}(pr2, W))^3 / \widetilde{N}^*(pr1) / \widetilde{N}^*(pr3) \end{aligned}$$

The estimation produced results like this:

The estimates are correlated, some of them heavily:

$\operatorname{cor}(\operatorname{Est}(\sigma(pr1)^2), \operatorname{Est}(\eta(pr1))) = -0.6$	5806.
$cor(Est(C_1), Est(C_2)) = -0.9183$	$cor(Est(C_2), Est(C_3)) = 0.6865$
$cor(Est(C_1), Est(C_3)) = -0.5350$	$cor(Est(C_2), Est(C_4)) = -0.5528$
$cor(Est(C_1), Est(C_4)) = 0.3051$	$cor(Est(C_3), Est(C_4)) = -0.9015.$

The feeding levels, their covariances, and correlations can be estimated from

$$\begin{split} f(W) &= C_1 * W^{C_2} / (C_1 * W^{C_2} + C_3 * W^{C_4} / \phi(W)); \\ \text{Est}(f(10 \text{ g})) &= 0.5323. \text{ var}(\text{Est}(f(10 \text{ g}))) = 0.0436^2 \\ \text{Est}(f(640 \text{ g})) &= 0.9642. \text{ var}(\text{Est}(f(640 \text{ g}))) = 0.0149^2 \\ \text{Est}(f(10240 \text{ g})) &= 0.8885. \text{ var}(\text{Est}(f(10240 \text{ g}))) = 0.0729 \\ \text{cov}(\text{Est}(f(10 \text{ g})), \text{Est}(f(640 \text{ g}))) = 4.81\text{E} - 5 \\ \text{cor}(\text{Est}(f(10 \text{ g})), \text{Est}(f(640 \text{ g}))) = 0.074 \\ \text{cov}(\text{Est}(f(640 \text{ g})), \text{Est}(f(10240 \text{ g}))) = 1.08\text{E} - 3 \\ \text{cor}(\text{est}(f(640 \text{ g})), \text{Est}(f(10240 \text{ g}))) = 0.987 \end{split}$$

Example 2

Example 1 illustrates a rather seldom met situation where all prey species are known and furnished with data. If the mathematical form of the available food $\phi(W)$ is known it is however possible to tackle a situation with stomach data for only a subset of the prey species. To illustrate this, data have been simulated for the example 1 set of predators, but only two prey species pr5 and pr6 with environmental parameters:

τ ²		μ	\widetilde{N}^*_{st}	
pr5	1	-2.3	2299	
pr6	1	-4.7	16960	

are used in the estimation.

The available food $\phi(W)$ is given by:

 $\phi(W) = \gamma * W^{\lambda}$

and in the simulations were used

 $\begin{array}{l} \gamma = 100 \\ \lambda = 0.4 \\ \rho(pr5) = 0.5. \ \sigma(pr5)^2 = 1.5. \ \eta(pr5) = 6. \\ \rho(pr6) = 1. \ \sigma(pr6)^2 = 3. \ \eta(pr6) = 7 \end{array}$

together with the relevant parameters from example 1.

The data set delivers the observations:

Obs $(\tau(pr5, 10 g)^2)$, ..., Obs $(\tau(pr5, 10240 g))$, Obs $(\tau(pr6, 10 g)^2)$, ..., Obs $(\tau(pr6, 10240 g)^2)$, Obs $(\mu(pr5, 10 g))$, ..., Obs $(\mu(pr5, 10240 g))$, Obs $(\mu(pr6, 10 g))$, ..., Obs $(\mu(pr6, 10240 g))$, Obs(stctr(10 g)), ..., Obs(stctr(10240 g))

as before but instead of the Rs quantities R' are introduced:

$$\frac{R'(pr, W) = N_{st}^{**}(pr, W) * SQRT(\tau(pr)^2) * \exp(b^2/4/a - c)/stctr(W)/N^*(pr)/SQRT(\tau(pr, W)^2)$$

where

$$\begin{aligned} a &= (1/\tau (pr, W)^2 - 1/\tau (pr)^2)/2 \\ b &= (\mu (pr, W) - \log(W))/\tau (pr, W)^2 - (\mu (pr) - \log(W))/\tau (pr)^2 \end{aligned}$$

and

$$c = (\mu(pr, W) - \log(W))^2 / 2/\tau(pr, W)^2 - (\mu(pr) - \log(W))^2 / 2/\tau(pr)^2.$$

Obs(R') can be written as:

$$\begin{aligned} \operatorname{Obs}(R'(pr,w)) &= \rho(pr) * W^{-\lambda} * \kappa(W) * t_0(W) * 500/\\ \gamma/(1-\exp(-\kappa(W) * t_0(W))) + \varepsilon_5(pr,W) \\ &(= C_1' * \rho(pr) * W^{-\lambda} + \varepsilon_5(pr,W)). \end{aligned}$$

The ε 's still have the normality properties and a covariance matrix that can be estimated from observations and independent variables only.

The set of observations now consists of 42 figures, and all of them are regressions in the $\mu(pr)$'s, $\tau(pr)^2$'s, $\widetilde{N}^*(pr)$'s and W's with regression coefficients: $\sigma(pr5)^2$, $\sigma(pr6)^2$, $\eta(pr5)$, $\eta(pr6)$, $\rho(pr5)$, C'_1 , λ , $\kappa_0/h/(1 - \exp(-\kappa(W) * t_0(W))) (= C'_2)$, $-m + \kappa_1 (= C'_3)$, $q * \kappa_0/h/(1 - \exp(-\kappa(W) * t_0(W)))/C'_1 (= C'_4)$, $-\lambda - m + \kappa_1 + r$ (= C'_5). The attempt to estimate all regression coefficients was not successful, and as the reason probably was heavy correlation between C'_2 , ..., C'_5 the *stctr*'s were left out and an estimation of $\sigma(pr5)^2$, ..., λ was undertaken using the 36 observations left. Here are some of the results:

$$\begin{split} & \operatorname{Est}(\sigma(pr5)^2) = 1.362. \operatorname{var}(\operatorname{Est}(\sigma(pr5)^2)) = 0.1007^2 \\ & \operatorname{Est}(\eta(pr5)) = 6.148. \operatorname{var}(\operatorname{Est}(\eta(pr5))) = 0.0717^2 \\ & \operatorname{Est}(\rho(pr5)) = 0.558. \operatorname{var}(\operatorname{Est}(\rho(pr5))) = 0.0371^2 \\ & (\rho(pr6) \text{ is set to } 1). \\ & \operatorname{Est}(C'_1) = 6.847. \operatorname{var}(\operatorname{Est}(C'_1)) = 0.3773^2 \\ & \operatorname{Est}(\lambda) = 0.455. \operatorname{var}(\operatorname{Est}(\lambda)) = 0.0118^2 \\ & \operatorname{cor}(\operatorname{Est}(\sigma(pr5)^2), \operatorname{Est}(\eta(pr5))) = -0.523 \\ & \operatorname{cor}(\operatorname{Est}(\eta(pr5)), \operatorname{Est}(\lambda)) = 0.831 \end{split}$$

The form of the 'available' food function $\phi(W)$ implies that

 $log \left[w * N_{st}(w, pr, W) * d(log(w))/R'(pr, W) / (w * N(w, pr) * d(log(w)))/stctr(W) \right]$

should fit

$$P(\log(W/w)) = -(\log(W/w) - \eta(pr))^2/2/\sigma(pr)^2 = -\eta(pr)^2/2/\sigma(pr)^2 + \eta(pr) * \log(W/w)/\sigma(pr)^2 - \log(W/w)^2/2/\sigma(pr)^2,$$

a parabola in $\log(W/w)$. This gives a visual illustration of the adequacy of the model. The plot can be produced by using the R''s calculated from the observations, but a better fit will be obtained by using estimated R''s.

The problem with the estimation of C'_2 , ..., C'_5 can be overcome by increasing the number of stomachs, but simulations indicated that an unrealistic huge data set is necessary in this model for $\phi(W)$. A simulation with 50000 fish in each W-group gave the correlations:

$$cor(Est(C'_2), Est(C'_3)) = -0.994. cor(Est(C'_2), Est(C'_4)) = 0.624.$$

 $cor(Est(C'_2), Est(C'_5)) = -0.947. cor(Est(C'_3), Est(C'_4)) = -0.586.$
 $cor(Est(C'_3), Est(C'_5)) = 0.921. cor(Est(C'_4), Est(C'_5)) = -0.817.$

These figures show clearly that an estimation of the C''s, and therefore the feeding levels, is practically impossible when the 'available' food is a power function of W.

Concluding remarks

The purpose of this paper was to have a look at the possibilities for estimating predation and digestion parameters from stomach data. The examples show that this is practicable, at least for some specified models, if detailed and complete data is at hand. The indicated estimation procedure delivers the covariance matrix for the estimated parameters, and this means that one is able to test hypotheses on the parameters. An estimate of the sampling variance is also obtained, and as an estimate of the covariance matrix can be calculated from observations and independent variables, this means that it is possible to test the validity of the chosen model by an F-test. All testing results have however to be taken with a grain of salt as the regression functions are non-linear, and this means that the usual test procedures are only approximate.

The explicit way to follow when handling stomach data depends evidently of the underlying model, and as the models used in the examples are by no means sacrosanct, the procedures should be properly revised when applied to real data. It is however hoped that the sketched handling of some simulated data sets can be of some help when planning and evaluating experiments involving stomach contents.

Reference

Andersen, Knud P. 1982: An interpretation of the stomach contents of fish in relation to prey abundance. - Dana 2: 1-50.