Maximum Likelihood

Maximum likelihood estimation is a method for estimating model parameters based on data. For a parameter θ and data $x_1, x_2, ...x_n$, obtained from an independent sample with a probability density function $f()$, the likelihood is given by

$$L(\theta) = \prod_{i=1}^{n} f(x_1, x_2, ...x_n|\theta) \tag{1}$$

and the log likelihood by

$$L(\theta) = \sum_{i=1}^{n} f(x_1, x_2, ...x_n|\theta) \tag{2}$$

The best choice for θ is the value that maximises the likelihood, i.e. by choosing the value of θ that gives the greatest probability of obtaining the observations. This assumes that the distribution of the data is known and that the likelihood function can actually be evaluated for all the values of θ considered.

Example

To illustrate the maximum likelihood approach we use a stock assessment example based. Stock assessment use two main sources of information, i.e. trends in population abundance obtained from surveys and standardised catch per unit effort (CPUE) series and information on age structure [**?**]. If we have a model that predicts the stock biomass (B_i) and a relative index of stock abundance (I_i), then for all parameter values (θ) we choose the ones that minimises the difference between observations and predictions i.e.

$$\sum_i (I_i - qB_i)^2 \tag{3}$$

where B is scaled by the catchability (q the constant of proportionality).

The lognormal probability density is given by

$$\frac{1}{\sqrt{2\pi}} \sum_i \left(\frac{I_i/(qB_i)}{\sigma_i} \right)^2 + ln(\sigma_i) \tag{4}$$

where σ^2 is the residual variance.

The probability density is a function of θ so for a given set of observations finding the values of θ that maximises the density gives the maximum likelihood estimate of θ.

Maximum likelihood also allows you to test alternative hypotheses, since the maximum likelihood principle says that you should choose the hypothesis that gives the highest probability of occurring. Since x is continuous the probability of any particular value is infinitesimal. Instead the probability densities under the various hypotheses, as given by the height of the probability density function, are compared

Likelihoods

We document the various likelihood formulations as full (negative) log-likelihood functions (including constants) to allow the use of information criteria for model comparisons (Burnham and Anderson 2002).

Assuming that an abundance index is log-normally distributed about the model estimates, its negative log-likelihood contribution is given by

$$L(\theta) = \sum_{i=1} \frac{1}{2} W_i \sum_{y=Y_j} \left\{ log[2\pi(\sigma_{i,j}^2 + \lambda_i^2] + \frac{[log(X_{i,y} - log(q_i \hat{X}_{i,j})]^2}{\sigma_{i,j}^2 + \lambda_i^2} \right\} \quad (5)$$

where $I_{i,y}$ is the abundance index observation for year y and series i and $\hat{I}_{i,y}$ is the corresponding model estimate. q_i is the constant of proportionality associated with series i and $\sigma_{i,j}^2 + \lambda_i^2$ is the residual variance for year y and series i, where represents the sampling component of this variance associated with each observation (i.e. each year y) of series i, and represents the extent of additional variance (over and above that linked to sampling – Punt and Butterworth 2003, Porch 2003) associated with series i; $Y_{1,i}$ represents the years for which abundance observations are available for series i; and $W_{1,i}$ is the relative weight given to the abundance index component of the likelihood associated with series i.

Variances

There are several options for handling the variances. If all values for all indices are given equal weight, they can be set to 1. Alternatively an index could be weighted depending on how well it is fitted by the model, i.e. maximum likelihood weighting,

$$\sigma_i^2 = \frac{\sum_y (I_{iy} - \hat{I}_{iy})^2}{n_i} \quad (6)$$

σ_i^2 can be obtained by external to the procedure used to maximise the likelihood or as part of maximisation procedure.

Alternatively, the variances could be input for each value, based on external information, for example if a generalized linear model (GLM) has been used to standardise the CPUE.

The variance has two components σ_{iy}^2 and λ_{iy}^2 due to measurement and process error. The later represents the sampling component of the variance associated with each observation (y) of series i, and represents the extent of additional variance (over and above that linked to sampling (Punt and Butterworth 2003) associated with series i.

The assumption of log-normality allows the sampling component of the residual standard deviation, σ_{ij} to be approximated by the CV (coefficient of variation) of the untransformed distributions of the Xi,y.

Estimation options for λ_i and qi, given σ_{ij} (a fixed input value) are as follows: (a) λ_i fixed (i) If $\lambda_i = 0$ then it is necessary that $\sigma_{ij} > 0$ for all y. Setting $\sigma_{ij} = 1$ for all y reduces equation [1] to a least-squares formulation.

(ii) If $\lambda_i > 0$ then $\sigma_{ij}\theta$ for all y. Setting $\sigma_{ij} = \sigma_j$ for all y (σ_{ij} θ) allows for equal weighting of all observations relative to one another within an abundance series.

If qi is to be estimated, then a closed form solution exists, as follows: If σ_{ij} $= \sigma_i$ for all y, then equation [2] reduces to the following:

b) λ_i estimated (i) $\sigma_{ij} = 0$ for all y: As for (a)(ii) above, this allows equal weighting for all observations within an abundance series. A closed-form solution for λ_i exists, as follows:

If qi is also estimated, then the closed form solution given in equation [3] applies.

(ii) $\sigma_{ij} > 0$ In this case, a closed form solution for λ_i is no longer straight forward. An estimate for λ_i can be obtained by treating it as an estimable parameter in the non-linear optimisation of equation [1], or by finding the value for λ_i that satisfies the following equati If qi is also to be estimated, then equation [2] can be used in conjunction with equation [5] to simultaneously solve for qi and λ_i using an iterative algorithm. Alternatively, both qi and λ_i could be treated as estimable parameters.

1 Abundance indices

Due to the difficulty of conducting fisheries independent surveys for highly migratory stock most indices of abundance in tuna stock assessments are based on commercial catch per unit effort (CPUE). Indices incorporate a range of ages depending on their vulnerability to the fishery. Therefore when using CPUE as a proxy for relative stock abundance it is necessary to specify how individual ages contribute to an index, For example when modelling CPUE in a stock assessments [Butterworth and Geromont(1999)] proposed weighting estimates of numbers-at-age by their relative vulnerability-at-age (v_{iay}) i.e.

$$v_{ia} = \frac{\sum_y C_{iay} F_{ay} / C_{ay}}{\max_a \{C_{iay} F_{ay} / C_{ay}\}} \tag{7}$$

where i is index, a age and y year and C_{iay} is the partial catch and F_{ay} fishing mortality-at-age.

The index predicted by the assessment model is then given by

$$I_{iy} = q_{iy}\delta_i \sum_a v_{iay} w_{iay} \tilde{N}_{ay} \tag{8}$$

Where δ_{iy} is the adjustment for time of year, q_{iy} is the catchability coefficient, w_{iay} mass-at-age and $\tilde{N}_{ay}$ the estimated of numbers-at-age.

Lognormal

$$\sum_i \sum_y 0.5 \left(\frac{I_{iy}/\hat{I}_{iy}}{\tilde{\sum}_{iy}} \right)^2 + ln(\tilde{\sum}_{iy}) \tag{9}$$

$$\tilde{\sum}_{iy} = \sqrt{ln\left\{\left(\frac{\sum_{iy}}{\hat{I}_{iy}}\right)^2 + 1\right\}} \tag{10}$$

Normal

$$\sum_i \sum_y 0.5\left(\frac{I_{iy} - \hat{I}_{iy}}{\sum_{iy}}\right)^2 + ln(\sum_{iy}) \tag{11}$$

Poisson

$$\sum_i \sum_y \hat{I}_{iy} - I_{iy} ln(\hat{I}_{iy}) \tag{12}$$

Chi-squared

$$\sum_i \sum_y \frac{(I_{iy} - \hat{I}_{iy})^2}{\sum_{iy}^2 (\hat{I}_{iy} + 1)} \tag{13}$$

Laplace

$$\sum_i \sum_y \left(\frac{\sqrt{2}|I_{iy} - \hat{I}_{iy}|}{\sum_{iy}}\right)^2 + ln(\sum_{iy}) \tag{14}$$

Gamma

$$\sum_i \sum_y C ln(\beta) - (\alpha - 1) ln(I_{iy}) - \frac{I_{iy}}{\beta} - ln\Gamma(\alpha) \tag{15}$$

$$\alpha = (\hat{I}_{iy} / \sum_{iy})^2 \tag{16}$$

$$\beta = \hat{I}_{iy} / \alpha \tag{17}$$

References

[Butterworth and Geromont(1999)] D. Butterworth and H. Geromont. Some aspects of adapt vpa as applied to north atlantic bluefin tuna. *Int. Commn Cons. Atl. Tunas, Coll. Vol. Sci. Pap*, 49(2):233–241, 1999.

References

Burnham, K.P. and D.R. Anderson 2002 – Model selection and multimodel inference: a practical information-theoretic approach. Second Edition. Springer-Verlag, New York: [xxvi] + 488pp.

De Oliveira, J.A.A. 2003 – The development and implementation of a joint management procedure for the South African pilchard and anchovy resources. Ph.D. thesis, University of Cape Town, Cape Town: [iv] + 319pp.

Ernst, B. 2002 – An investigation on length-based models used in quantitative population modeling. Ph.D. thesis, University of Washington, Seattle: 150pp.

Fournier, D.A., Sibert, J.R., Majkowski, J. and J. Hampton 1990 – MULTIFAN a likelihood-based method for estimating growth parameters and age composition from multiple length frequency data sets illustrated using data for southern bluefin tuna (Thunnus maccoyii). Can. J. Fish. Aquat. Sci. 47: 301-317.

Geromont, H.F. and D.S. Butterworth 1999 – Operating models proposed for south coast Merluccius capensis management procedure robustness trials. Unpublished Report, Marine and Coastal Management, South Africa WG/11/99/D:H:46: 33pp.

Hilborn, R., Maunder, M.[N.], Parma, A.[M.], Ernst, B., Payne, J.[] and P.[J.] Starr 2003 – Coleraine. A generalized age-structured stock assessment model. User's Manual Version 2.0 (Revised May 2003): 58pp.

Johnson, N.L. and Kotz,S. (1972). Distributions in Statistics: Continuous Multivariate Distribution. Wiley, New York.

Methot, R.D. 1989 – Synthetic estimates of historical abundance and mortality for northern anchovy. Am. Fish. Soc. Symp. 6: 66-82.

McAllister, M.K. and J.N. Ianelli 1997 – Bayesian stock assessment using catch-age data and the sampling – importance resampling algorithm. Can. J. Fish. Aquat. Sci. 54: 284-300.

Porch, C.E. 2003 – VPA-2Box. Version 3.01. User's Guide.

Punt, A.E. and D.S. Butterworth 2003 – Specifications and clarifications regarding the ADAPT VPA assessment/projection computations carried out during the September 2000 ICCAT West Atlantic bluefin tuna stock assessment session. Col. Vol. Sci. Pap. ICCAT, 55(3): 1028-1040.

Punt, A.E. and R. Hilborn 1997 – Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries 7: 35-63.

Punt, A.E. and R.B. Kennedy 1997 – Population modelling of Tasmanian rock lobster, Jasus edwardsii, resources. Mar. Freshwater Res. 48: 967-980.

Smith, A.D.M and A.E. Punt 1998 – Stock assessment of gemfish (Rexea solandri) in eastern Australia using maximum likelihood and Bayesian methods. In Fishery Stock Assessment Models. F. Funk, T.J. Quinn II, J. Heifetz, J.N. Ianelli, J.E. Powers, J.F. Schweigert, P.J. Sullivan, and C.I. Zhang (Eds). Alaska

Sea Grant College Program Report No. AK-SG-98-01, University of Alaska Fairbanks: 245-286.