

Aster Models for Life History Analysis

Charles J. Geyer
School of Statistics
University of Minnesota

<http://www.stat.umn.edu/geyer/aster/>

Aster Models

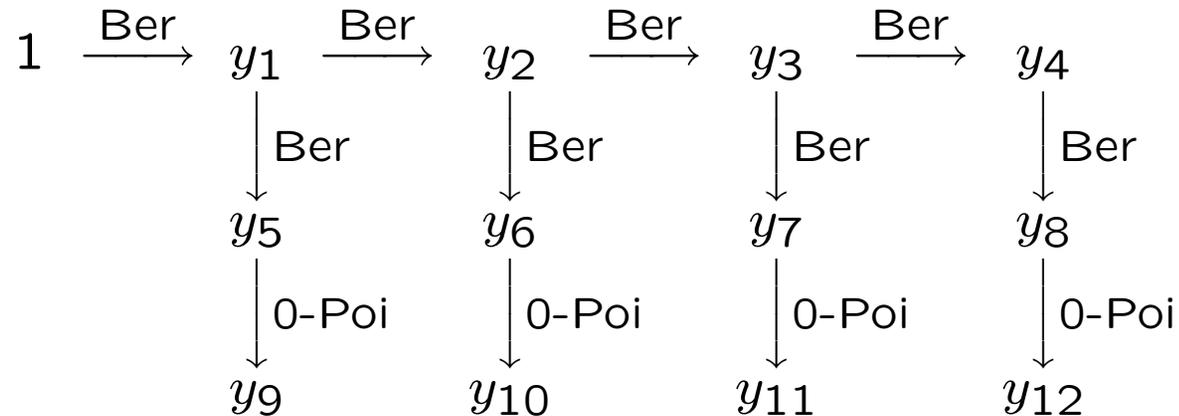
Aster models (named after the flowers) are a generalization of generalized linear models (GLM) and survival analysis.

As in linear models (LM) and GLM we model the distribution of the response vector y conditional on covariate data x .

Components of y are allowed to be dependent conditional on x .

Components of y are allowed to have different families (e. g., some Bernoulli, some Poisson, some normal) conditional on x and on each other.

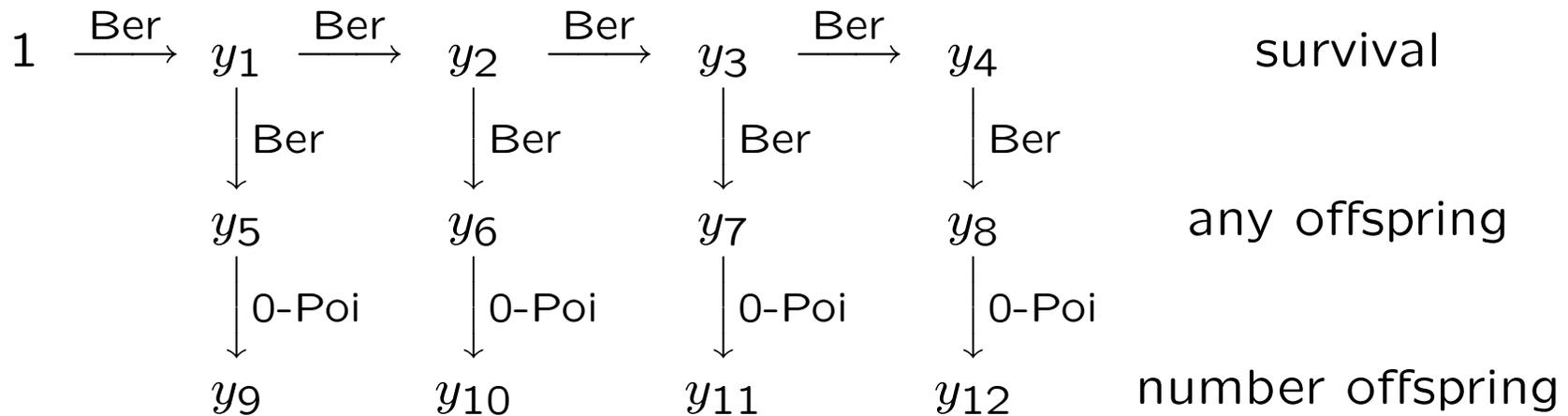
Typical Aster Data



Graph shows components of response vector for one individual. Every individual has isomorphic graph.

Joint distribution is product of conditional distributions, one for each arrow in graph. Labels indicate conditional distributions (Ber = Bernoulli, 0-Poi = zero-truncated Poisson).

Typical Aster Data (cont.)



y_1, \dots, y_4 indicate survival in each of 4 years.

y_5, \dots, y_8 indicate one or more offspring in each year.

y_9, \dots, y_{12} count offspring in each year.

Conditional distribution of y_{j+8} given y_j is zero-inflated Poisson.

Typical Aster Data (cont.)

Lots of data like this — thousands of papers on evolution in wild or experimental populations of plants and animals.

Main interest is in lifetime number of offspring, $y_9 + \dots + y_{12}$ in example. Best surrogate of Darwinian fitness.

Marginal distribution of $y_9 + \dots + y_{12}$ is nonnegative, has large atom at zero, is multimodal. No LM or GLM appropriate. Nor any other procedures available before aster models.

Scientists made do with either analyzing separate variables or treating $y_9 + \dots + y_{12}$ as normal even though usually grossly wrong. Neither satisfactory. Much literature saying so.

General Aster Graphical Model

Arbitrary directed acyclic graph such that each node has at most one incoming arrow.

Let J indicate nodes with incoming arrow and F nodes without.

Each $j \in J$ associated with component y_j of response vector.
Each $j \in F$ associated with constant y_j .

For each $j \in J$ let $p(j)$ denote unique node such that there is arrow $p(j) \longrightarrow j$.

Joint Distribution

Graph determines factorization of joint distribution of y into product of conditionals

$$\text{pr}(y) = \prod_{j \in J} \text{pr}(y_j \mid y_{p(j)})$$

(conditioning on constant same as unconditional).

Conditional distributions are exponential family with canonical statistic y_j , canonical parameter θ_j , cumulant function c_j , and sample size $y_{p(j)}$. Log likelihood is

$$l(\theta) = \sum_{j \in J} \left[y_j \theta_j - y_{p(j)} c_j(\theta_j) \right]$$

Aster Transform

Joint distribution is exponential family because all components of y enter linearly in

$$\begin{aligned} l(\theta) &= \sum_{j \in J} \left[y_j \theta_j - y_{p(j)} c_j(\theta_j) \right] \\ &= \sum_{j \in J} y_j \left[\theta_j - \sum_{\substack{k \in J \\ p(k)=j}} c_k(\theta_k) \right] - \sum_{\substack{k \in J \\ p(k) \in F}} y_{p(k)} c_k(\theta_k) \end{aligned}$$

Coefficients of y_j (blue) are canonical parameters for the joint distribution

$$\varphi_j = \theta_j - \sum_{\substack{k \in J \\ p(k)=j}} c_k(\theta_k), \quad j \in J$$

This is invertible change-of-parameter, infinitely differentiable both ways.

Log Likelihood and Fisher Information

Log likelihood for canonical parameters is

$$l(\varphi) = y^T \varphi - c(\varphi)$$

where

$$c(\varphi) = \sum_{\substack{k \in J \\ p(k) \in F}} y_{p(k)} c_k(\theta_k)$$

the relation between φ and θ being the aster transform. Fisher information is

$$I(\varphi) = \nabla^2 c(\varphi)$$

Canonical Linear Submodels

Submodels parameterized

$$\varphi = M\beta$$

(model matrix M depends on covariate data) are themselves exponential families with canonical statistic vector $M^T y$, canonical parameter vector β , and mean value parameter vector

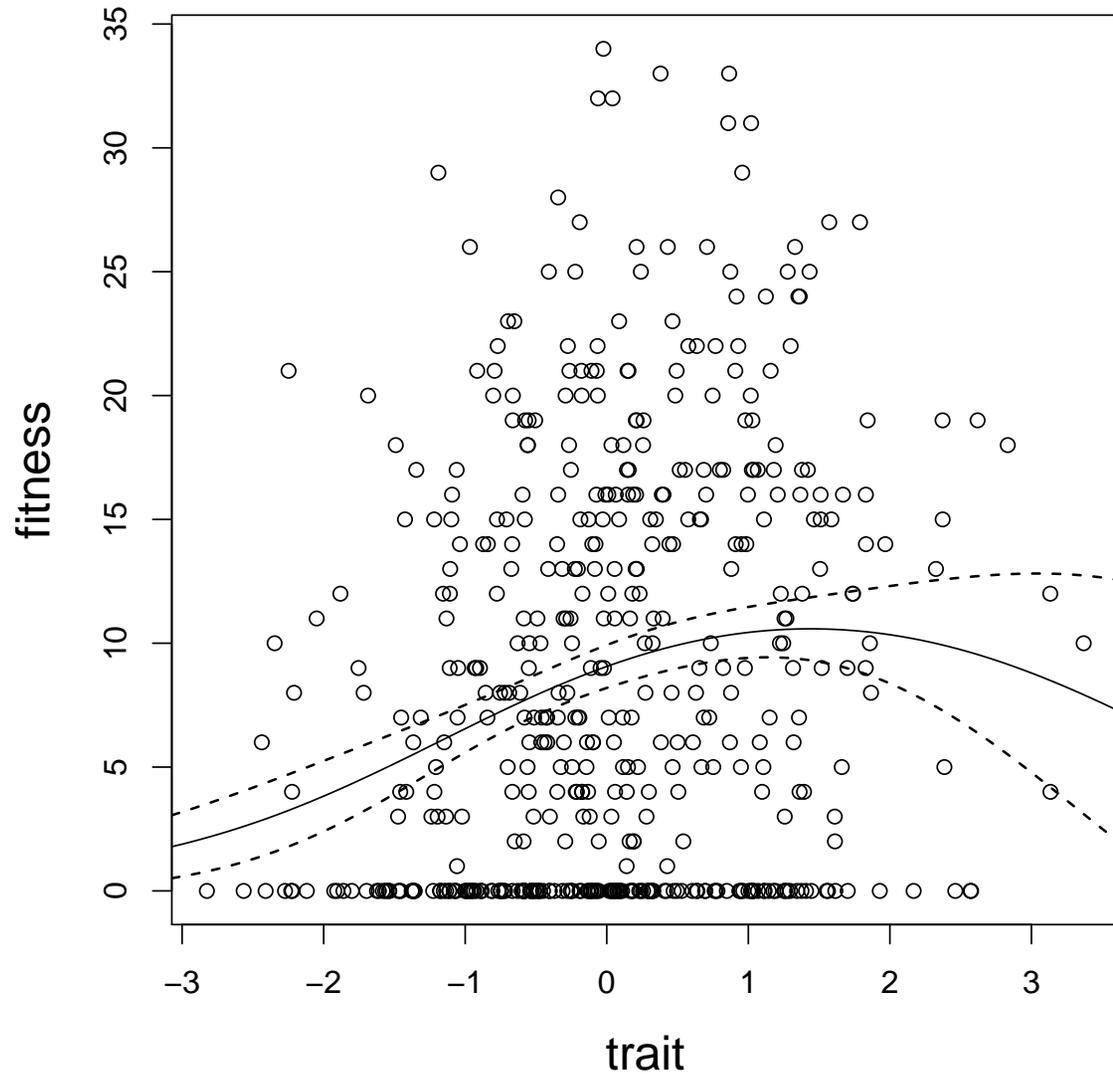
$$\tau = E_{\varphi}(M^T y)$$

Exponential family theory gives interpretability of submodels. “Observed equals expected” property of maximum likelihood estimates

$$\hat{\tau} = M^T y$$

and

$$\partial\tau_j/\partial\beta_j > 0, \quad j \in J.$$



Fitness Landscapes

In plot vertical coordinate “fitness” is $y_9 + \dots + y_{12}$ in aster graphical model (lifetime number of offspring) and horizontal coordinate “trait” is some measured characteristic of organism, like height or weight.

Biologists want to know the relationship between traits and fitness, the “fitness landscape” .

To statisticians this is just a regression function. Solid curve in plot is

$$E_{\hat{\beta}}(y_9 + \dots + y_{12} | x)$$

where x is the trait, plotted as a function of x . Dashed curves indicate pointwise asymptotic 95% confidence intervals from the aster model.

Fitness Landscapes (cont.)

R contributed package `aster` from CRAN does aster models. After setting up data structures

```
out <- aster(resp ~ varb + x + I(x^2),  
             pred, fam, varb, id, root, data = mydata)  
summary(out)  
pout <- predict(out, newdata = newdata, varb, id, root,  
                amat = amat, se.fit = TRUE)
```

fits aster model, shows fitted regression coefficients with standard errors etc., and calculates predicted values and standard errors for regression function.

Literature

Geyer, C. J., Wagenius, S. and Shaw, R. G. (2007).
Aster models for life history analysis.
Biometrika, **94** 415–426.

Shaw, R. G., Geyer, C. J., Wagenius, S., Hangelbroek, H. H.,
and Etterson, J. R. (2008).
Unifying life history analysis for inference of fitness and
population growth.
American Naturalist, **172**, E35–E47.

Shaw, R. G. and Geyer, C. J. (2010).
Inferring fitness landscapes.
To appear in *Evolution*.

Reproducible Research

All analyses in these papers in technical reports available at

`http://www.stat.umn.edu/geyer/aster/`

done using R function `Sweave` so exactly reproducible by anyone who has R.

Voluminous examples with all details.