Combining Survival Analysis and a Linear Animal Model,

to estimate genetic parameters for social effects on survival in layers

Esther Ellen[1], Vincent Ducrocq[2], Bart Ducro[1], Roel Veerkamp[3] and Piter Bijma[1]

[1]Animal breeding and Genomics Centre, Wageningen University, The Netherlands
[2] Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, The Netherlands
[3] UMR, INRA, France

WAGENINGEN UNIVERSITY
WAGENINGEN UR

NWO
Netherlands Organisation for Scientific Research

ISA
A Hendrix Genetics Company

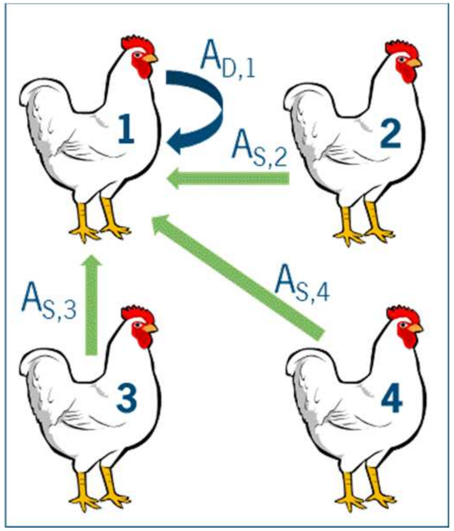
Animal Breeding &
Genomics Centre

Survival in laying hens

Picture provided by Bill Muir

- Substantial mortality due to cannibalism
- Currently beak-trimming is used as a “solution”
- Genetic solution is desired

The Genetic Model

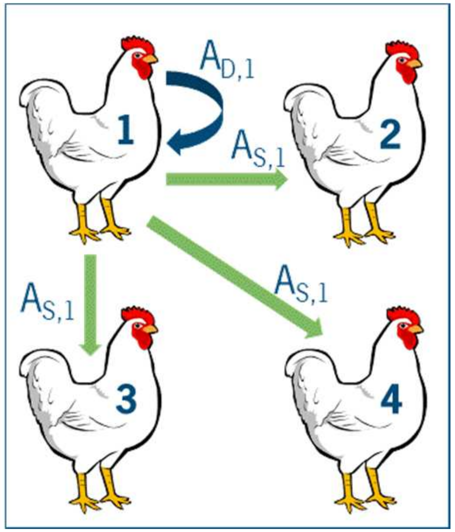


Previous Results:

33-87% of Var(A_T) in survival time

originates from social effects.

But: from Linear Models



The Phenotype

$$P_i = A_{D,i} + \sum_{\text{group mates}} A_{S,j} + e$$

The Total Breeding Value

$$A_{T,i} = A_{D,i} + 3A_{S,i}$$

Parameters of interest:

Direct genetic variance: $\sigma^2_{A_D}$

Direct-social covariance $\sigma_{A_{DS}}$

Social genetic variance: $\sigma^2_{A_S}$

Total heritable variance: $\sigma^2_{A_T} = \sigma^2_{A_D} + 6\sigma_{A_{DS}} + 9\sigma^2_{A_S}$

Genetic analysis of survival time in laying hens

- Issues:
  - Censoring: Many animals still alive at end of study period
  - Social genetic effects: Survival time depends on group mates

- Use survival analysis (survival kit (Ducrocq and Sölkner, 1998))
- Problem: Social effects cannot (yet?) be included
- Solution: Combine survival analysis and a linear animal model including social effects
  - → two-step approach

Objectives

- To estimate direct and social genetic parameters for survival time in cannibalistic layers
- Using a two-step approach (2STEP)
- Compare 2STEP with a 1-step linear animal model (LAM)

Material

- Survival data from ISA (Hendrix Genetics)
- 2 purebred White Leghorn layer lines: W1 and WB
- ~13,000 hens
- 4 birds/cage
- Randomly composed cages
- Intact beaks

- **Trait:** survival days

Survival days = number of days from start of laying till either death or end of study (max = 447 days)

Methods – 2STEP approach

- Step 1: Calculate pseudo-records using survival analysis
  - “Record” on the hazard scale
  - Cox animal model
  - Fixed effects:
    - Laying house x row x level
    - Effect of back neighbours
  - Pseudo record (y_i^*): $y_i^* = \frac{\delta_i}{\omega_i} + a_i - 1$
    - δ_i = censoring code of individual i (1 = uncensored; 0 = censored)
    - a_i = direct breeding value of individual i
    - ω_i = weight for individual i

Methods – 2STEP approach

- Step 2: Estimate genetic parameters with linear model

$$\mathbf{y}^* = \mathbf{Z}_\mathrm{D}\mathbf{a}_\mathrm{D} + \mathbf{Z}_\mathrm{S}\mathbf{a}_\mathrm{S} + \mathbf{e}$$ (Muir, 2005)

- $\mathbf{y}^*$ = vector of pseudo-records
- $\mathbf{a}_\mathrm{D}$ = direct breeding values
- $\mathbf{a}_\mathrm{S}$ = social breeding values
- $\mathbf{e}$ = vector of residuals, where $Var(e_i) = \frac{1}{\omega_i}\sigma_e^2$

Methods – 1step Linear Animal Model (LAM)

- Fit linear model directly to observed survival days

$$\mathbf{y} = \mathbf{Z}_D\mathbf{a}_D + \mathbf{Z}_S\mathbf{a}_S + \mathbf{e}$$ (Muir, 2005)

  - $\mathbf{y}$ = vector of observed survival days
  - $\mathbf{a}_D$ = direct breeding values
  - $\mathbf{a}_S$ = social breeding values
  - $\mathbf{e}$ = residuals

Comparison of Methods: Cross-validation

- 20% of phenotypes were set to missing
- 5 subsets
- Missing phenotypes were predicted: $\hat{P}_i = \hat{A}_{D,i} + \sum_{\text{group mates}} \hat{A}_{S,j}$
- Validation is difficult:
  - Scale difference (survival days *vs.* hazard scale)
  - 50-70% of data were censored – no observed phenotype
- 2 approaches to compare 2STEP and LAM:
  - Rank-correlation between observed and predicted phenotypes
    - Random rank for censored records
  - “Response to selection” approach:
    - Observed difference between 25% best and worst predicted individuals

Results – Genetic parameters

- Both methods yielded significant social genetic variance
  - Estimates are not comparable due to scale difference (not shown)

- Both methods yielded a similar relative contribution of social effects to total genetic variance
  - ~ 50% for line W1
  - ~ 60% for line WB

Results – Cross-validation: rank-correlation

	Rank Corr(P, $P_{predicted}$)		
Line	2STEP	LAM	2STEP-LAM
W1	0.149 ± 0.011	0.144 ± 0.010	0.954 ± 0.003
WB	0.174 ± 0.020	0.170 ± 0.020	0.962 ± 0.004

Both method predict missing phenotypes equally well

Predicted phenotypes of both methods were highly correlated

Results – Cross validation: “response to selection”

	W1		WB	
	2STEP	LAM	2STEP	LAM
Mean	354 ± 2		326 ± 2	
Predicted best	377 ± 3	377 ± 3	359 ± 2	357 ± 3
Predicted worst	327 ± 2	327 ± 3	292 ± 6	290 ± 6
Difference	**50**	**50**	**67**	**67**

Both methods yield the same “response to selection”

Discussion

- Surprise: No difference between 2STEP and LAM
  - Reason: All individuals were censored at same time
  - With variation in censoring time, 2STEP was superior
- Assumptions
  - In the cross-validation, observed phenotypes were precorrected for fixed effects using LAM
  - Rank-correlation approach assumed random ranks for censored records
  - But “response to selection” approach had no assumptions
- Conclusion: 2STEP-approach is useful when censoring time varies

Combining Survival Analysis and a Linear Animal Model,

to estimate genetic parameters for social effects on survival in layers

Esther Ellen[1], Vincent Ducrocq[2], Bart Ducro[1], Roel Veerkamp[3] and Piter Bijma[1]

[1]Animal breeding and Genomics Centre, Wageningen University, The Netherlands
[2] Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, The Netherlands
[3] UMR, INRA, France

WAGENINGEN UNIVERSITY
WAGENINGEN UR

NWO
Netherlands Organisation for Scientific Research

ISA
A Hendrix Genetics Company

Animal Breeding & Genomics Centre