Using Natural Splines in Logistic Regression, with an Application to Tulip Germination Experiments

Mickey Warner

23 April 2014

Problem Statement

Tulips are significant to Netherlands:

- Symbol of the region
- Contribution to agricultural exports
- Tourist attraction

Tulips require a chilling time to bloom

Climate change in Netherlands

Research Questions

Is the effect of chilling time the same across all populations?

Which populations are the same/different?

Is there an “ideal” chilling time?

Does this ideal chilling time vary by population?

What effect will a decrease from 10 to 8 weeks of winter/chilling time have for tulips?

Statistical Model and Methods

Logistic regression

Model formulation:

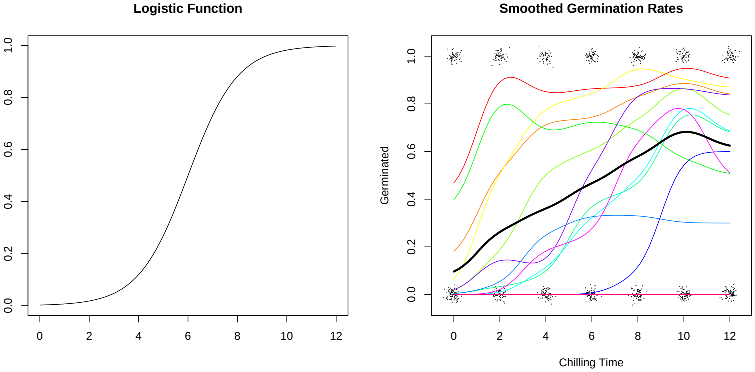
$$Y_i \sim \mathrm{Bern}(p_i)$$

$$\log\left(\frac{p_i}{1-p_i}\right) = \mathbf{x}_i'\boldsymbol{\beta}$$

p_i is the probability of germination given covariates $\mathbf{x}_i$

$\boldsymbol{\beta}$ is the vector of coefficient parameters

Non-monotonic effects



Natural splines (basis functions)

Advantages:

- Allows for non-“linear” responses
- Imposes smoothness constraints (knots)
- Provides nice tail behavior

Disadvantages:

- Increases the number of parameters
- Is difficult to interpret (without a figure)

Analysis

Likelihood ratio tests for differences in models ($\alpha = 0.05$)

- Compute full and reduced models
- Bonferroni correction
- Low p-value suggests a difference in population

Determine number of knots via cross-validation and AUC

Bootstrapping to compute ideal chilling times

Results

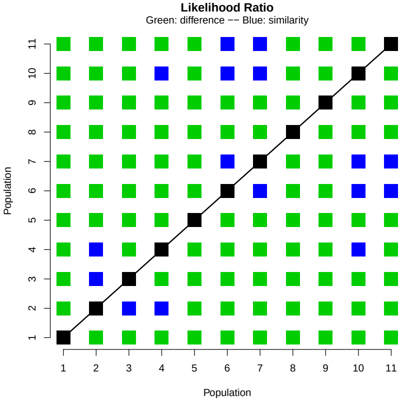
Effect of chilling time for populations

Not the same across all populations (small p-value)

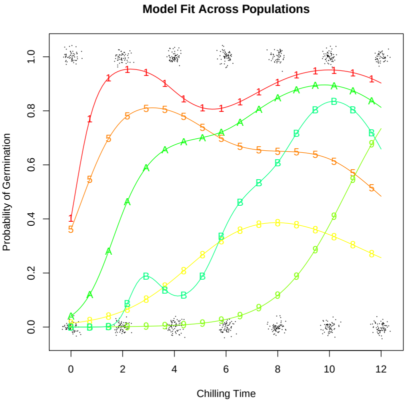
Populations with same chilling time effect:

- A: $2, 3, 4$
- B: $6, 7, 10, 11$

Effect of chilling time for populations



Models



Model Fit Across Populations
Probability of Germination
Chilling Time
0.0
0.2
0.4
0.6
0.8
1.0
0
2
4
6
8
10
12

Ideal chilling times and uncertainties

Chilling Time			
Pop.	Lower	Est.	Upper
1	1.28	10.06	12.00
5	2.40	2.86	9.68
8	6.75	7.65	12.00
9	12.00	12.00	12.00
A	8.92	9.81	12.00
B	9.64	10.15	10.55

Probabilities			
Pop.	Lower	Est.	Upper
1	0.93	0.99	1.00
5	0.69	0.84	0.94
8	0.26	0.41	0.55
9	0.54	0.77	0.93
A	0.83	0.90	0.95
B	0.73	0.84	0.92

Change from 10 to 8 week winters

	Change in Probability		
Pop.	Lower	Est.	Upper
1	-0.263	-0.008	0.090
5	-0.127	0.002	0.204
8	-0.069	0.037	0.185
9	-0.451	-0.230	-0.161
A	-0.108	-0.051	0.003
B	-0.379	-0.193	-0.065

Conclusion

Some populations experience same effect on chilling time

Most groups had ideal chilling times around 10 weeks, others had low times

Future research:

- Consider other knot locations for natural splines
- Other possible non-linear methods
- What's up with population 12?