

Penalized Likelihood Logistic Regression for Sparse Data Using Data Priors

by Martin Dörr

PhUSE US Connect 2020, March 8th-11th, Orlando, FL

At Clinipace, it's personal.

Penalized likelihood regression

Frequentists	$\log L^{\text{pen}}(\beta) = \log L(\beta) + A(\beta)$
Bayesians	$p(\theta \mid X) \propto p(X \mid \theta) p(\theta)$

Method	Penalization term $A(\beta)$	Prior $p(\beta)$
Ridge	$-\lambda \sum \beta^2$	Normal
LASSO	$-\lambda \sum \beta $	Double exponential
Firth	$\frac{1}{2} \log \det(I(\beta))$	Jeffreys'
Greenland	$\sum \left(\frac{\beta}{2} - \log(1 + \exp \beta) \right)$	logF(1,1)

$I(\beta)$ = Fisher information matrix

Implementation in SAS®

- ▶ **PROC LOGISTIC**
 - ▶ Firth method is available
- ▶ **PROC NLMIXED**
 - ▶ Firth method is not feasible
 - ▶ logF(1,1) data priors are feasible
- ▶ **Why switch to PROC NLMIXED?**
 - ▶ Only PROC NLMIXED provides odds ratios, relative risks and absolute risk differences marginalized at-the-means out of one model.

Example used in presentation

▶ Analysis variable

- RESPONSE (0=“failure”, 1=“response”)

▶ Covariables

- Intercept
- TRT (categorical)
- SITE (categorical)
- SEX (categorical)
- BASEZ (continuous, standardized with Z transformation)

▶ Logistic regression model

- $\text{RESPONSE} = \text{TRT SITE SEX BASEZ}$

Example data

Obs	trt	site	sex	base	basez	response
1	1	1	1	45	-0.63	1
2	2	1	1	56	0.69	0
3	1	1	1	51	0.09	1
4	2	1	2	59	1.05	0
5	1	1	1	58	0.93	0
6	2	1	2	39	-1.35	0
7	1	2	1	48	-0.27	0
8	2	2	2	47	-0.39	0
9	1	2	2	36	-1.71	1
10	2	2	1	64	1.65	0
11	1	2	1	46	-0.51	0
12	2	2	2	54	0.45	0

Example data excerpt

Obs		trt	site	sex	basez	response	
1		1	1	1	-0.63	1	
2		2	1	1	0.69	0	
...		...	...	...	...	...	
11		1	2	1	-0.51	0	
12		2	2	2	0.45	0	

Design matrix of example data excerpt

Obs	intercept	trt2	site2	sex2	basez	response	
1	1	0	0	0	-0.63	1	
2	1	1	0	0	0.69	0	
...	...	...	...	...	...	...	
11	1	0	1	0	-0.51	0	
12	1	1	1	1	0.45	0	

Adding partial likelihood weights for original records

Obs	intercept	trt2	site2	sex2	basez	response	pweight
1	1	0	0	0	-0.63	1	1
2	1	1	0	0	0.69	0	1
...	...	...	...	...	...	...	...
11	1	0	1	0	-0.51	0	1
12	1	1	1	1	0.45	0	1

Adding data priors for trt2: itself

Obs	intercept	trt2	site2	sex2	basez	response	pweight
1	1	0	0	0	-0.63	1	1
2	1	1	0	0	0.69	0	1
...	...	...	...	...	...	...	...
11	1	0	1	0	-0.51	0	1
12	1	1	1	1	0.45	0	1
13		1					
14		1					

Adding data priors for trt2: all other covariables

Obs	intercept	trt2	site2	sex2	basez	response	pweight
1	1	0	0	0	-0.63	1	1
2	1	1	0	0	0.69	0	1
...	...	...	...	...	...	...	...
11	1	0	1	0	-0.51	0	1
12	1	1	1	1	0.45	0	1
13	0	1	0	0	0		
14	0	1	0	0	0		

Adding data priors for trt2: response variable

Obs	intercept	trt2	site2	sex2	basez	response	pweight
1	1	0	0	0	-0.63	1	1
2	1	1	0	0	0.69	0	1
...	...	...	...	...	...	...	...
11	1	0	1	0	-0.51	0	1
12	1	1	1	1	0.45	0	1
13	0	1	0	0	0	0	
14	0	1	0	0	0	1	

Adding data priors for trt2: partial likelihood weights

Obs	intercept	trt2	site2	sex2	basez	response	pweight
1	1	0	0	0	-0.63	1	1
2	1	1	0	0	0.69	0	1
...	...	...	...	...	...	...	...
11	1	0	1	0	-0.51	0	1
12	1	1	1	1	0.45	0	1
13	0	1	0	0	0	0	0.5
14	0	1	0	0	0	1	0.5

Final design matrix excerpt with data priors

Obs	intercept	trt2	site2	sex2	basez	response	pweight
1	1	0	0	0	-0.63	1	1
2	1	1	0	0	0.69	0	1
...	...	...	...	...	...	...	...
11	1	0	1	0	-0.51	0	1
12	1	1	1	1	0.45	0	1
13	0	1	0	0	0	0	0.5
14	0	1	0	0	0	1	0.5
15	0	0	1	0	0	0	0.5
16	0	0	1	0	0	1	0.5
17	0	0	0	1	0	0	0.5
18	0	0	0	1	0	1	0.5
19	0	0	0	0	1	0	0.5
20	0	0	0	0	1	1	0.5

PROC NLMIXED code for ordinary logistic regression

```
proc nlmixed data=example_design tech=newrap df=%eval(12 - 5);  
  eta = b_intercept * intercept  
        + b_trt2      * trt2  
        + b_site2     * site2  
        + b_sex2      * sex2  
        + b_basez     * basez  
;  
  prob = 1 / (1 + exp(-1 * eta));  
  
  model response ~ binary(prob);  
run; quit;
```

PROC NLMIXED code for logistic regression with penalized likelihood

```
proc nlmixed data=example_design_pen tech=newrap df=%eval(12 - 5);  
  eta = b_intercept * intercept  
        + b_trt2      * trt2  
        + b_site2     * site2  
        + b_sex2      * sex2  
        + b_basez     * basez  
;  
  if response eq 1 then prob = 1 / (1 + exp(-1 * eta));  
    else prob = 1 / (1 + exp(eta));  
  
  prob_adj = (prob>0 and prob<=1)*prob + (prob<=0)*1e-8 + (prob>1);  
  
  wloglike = pweight * log(prob_adj);  
  
  model response ~ general(wloglike);  
run; quit;
```

Estimates and standard errors of model fits

Parameters	Ordinary fit		Penalized fit	
	PROC LOGISTIC	PROC NLMIXED	PROC LOGISTIC with Firth method	PROC NLMIXED with logF(1,1) data priors
Intercept				
TRT				
Site				
Sex				
BaseZ				

Estimates and standard errors of model fits

Parameters	Ordinary fit		Penalized fit	
	PROC LOGISTIC	PROC NLMIXED	PROC LOGISTIC with Firth method	PROC NLMIXED with logF(1,1) data priors
Intercept	27.65 (142.8)	-0.75 (24.6)		
TRT	-17.58 (207.9)	-23.75 (1466.8)		
Site	-29.19 (169.6)	-42.87 (1138.1)		
Sex	18.61 (163.1)	24.87 (722.2)		
BaseZ	-12.72 (93.8)	-18.52 (609.3)		

Estimates and standard errors of model fits

Parameters	Ordinary fit		Penalized fit	
	PROC LOGISTIC	PROC NLMIXED	PROC LOGISTIC with Firth method	PROC NLMIXED with logF(1,1) data priors
Intercept	27.65 (142.8)	-0.75 (24.6)	2.74 (3.4)	-0.39 (1.1)
TRT	-17.58 (207.9)	-23.75 (1466.8)	-2.49 (3.3)	-1.16 (1.9)
Site	-29.19 (169.6)	-42.87 (1138.1)	-1.52 (2.0)	-1.62 (1.7)
Sex	18.61 (163.1)	24.87 (722.2)	1.47 (2.2)	0.56 (1.4)
BaseZ	-12.72 (93.8)	-18.52 (609.3)	-0.40 (1.6)	-1.14 (1.2)

- ▶ logF(1,1) data priors are a probate means to stabilize the fit of a logistic regression when the dataset consists of sparse data and PROC NLMIXED is necessary for deriving marginalized endpoints.

References

- ▶ Greenland S, Mansournia M A: Penalization, bias reduction, and default priors in logistic and related categorical and survival regressions. *Statist. Med.*, 34: 3133– 3143. doi: 10.1002/sim.6537. 2015.



Martin Dörr
Senior Biostatistician
mdoerr@clinipace.com