

Application of GLM Advancements to Non-Life Insurance Pricing

Leonardo Stincone

Università degli Studi di Trieste

11 Maggio 2021



1. Il Pricing nelle Assicurazioni Danni

2. Modelli Statistici per il Pricing nelle Assicurazioni Danni

- Modelli Lineari Generalizzati (GLM)

- Modelli Additivi Generalizzati (GAM)

- Stimatori Shrinkage per i GLM

- Stimatori Bayesiani per i GLM

- Algoritmi di Machine Learning

- Confronto tra i modelli

3. Applicazione Pratica



1. Il Pricing nelle Assicurazioni Danni

2. Modelli Statistici per il Pricing nelle Assicurazioni Danni

Modelli Lineari Generalizzati (GLM)

Modelli Additivi Generalizzati (GAM)

Stimatori Shrinkage per i GLM

Stimatori Bayesiani per i GLM

Algoritmi di Machine Learning

Confronto tra i modelli

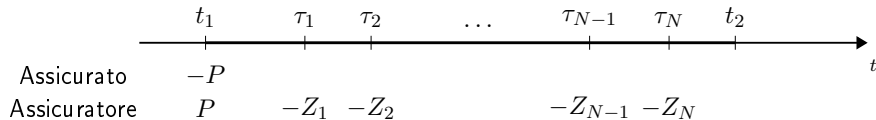
3. Applicazione Pratica



Contratto di Assicurazione, Art. 1882, Codice Civile Italiano

L'assicurazione è il contratto col quale l'**assicuratore**, verso il pagamento di un **premio**, si obbliga a rivalere l'**assicurato**, entro i limiti convenuti,

- ① del **danno** ad esso prodotto da un **sinistro**,
- ② ovvero a pagare un **capitale** o una **rendita** al verificarsi di un **evento** attinente alla **vita umana**.



Distribuzione composta

Assumiamo che:

- 1 $\forall n > 0, Z_1|N = n, Z_2|N = n, \dots, Z_n|N = n$ siano i.i.d.;
- 2 la distribuzione di $Z_i|N = n, i \leq n$ non dipenda da n .

Sotto queste ipotesi diciamo che:

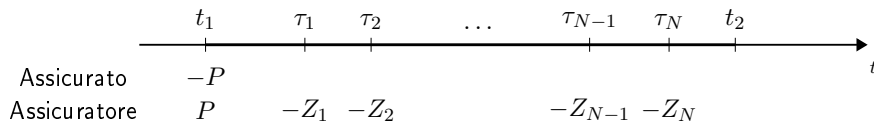
$$S = \begin{cases} 0 & \text{if } N = 0 \\ \sum_{i=1}^N Z_i & \text{if } N > 0 \end{cases}$$

ha distribuzione composta.

Proprietà



$$E(S) = E(N)E(Z)$$



Distribuzione composta

Assumiamo che:

- ① $\forall n > 0, Z_1|N = n, Z_2|N = n, \dots, Z_n|N = n$ siano i.i.d.;
- ② la distribuzione di $Z_i|N = n, i \leq n$ non dipenda da n .

Sotto queste ipotesi diciamo che:

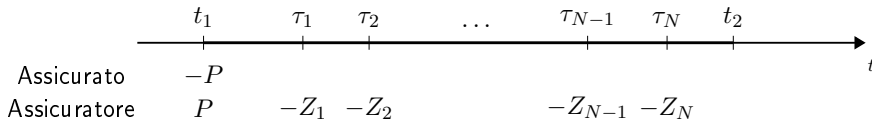
$$S = \begin{cases} 0 & \text{if } N = 0 \\ \sum_{i=1}^N Z_i & \text{if } N > 0 \end{cases}$$

ha distribuzione composta.

Proprietà



$$E(S) = E(N)E(Z)$$



Distribuzione composta

Assumiamo che:

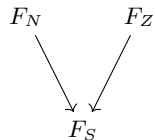
- ① $\forall n > 0, Z_1|N = n, Z_2|N = n, \dots, Z_n|N = n$ siano i.i.d.;
- ② la distribuzione di $Z_i|N = n, i \leq n$ non dipenda da n .

Sotto queste ipotesi diciamo che:

$$S = \begin{cases} 0 & \text{if } N = 0 \\ \sum_{i=1}^N Z_i & \text{if } N > 0 \end{cases}$$

ha distribuzione composta.

Proprietà



$$E(S) = E(N)E(Z)$$

Variabili esplicative

Possibili variabili esplicative per il pricing delle assicurazioni motor:

- Informazioni sul veicolo assicurato;
- Informazioni generiche sul contraente;
- Informazioni assicurative sul contraente;
- Opzioni sulla polizza assicurativa;
- Dati telematici.

Queste variabili possono essere codificate come un vettore di numeri reali:

$$\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip}) \in \mathcal{X} \subseteq \mathbb{R}^p$$

Regola di Pricing

Una *Regola di Pricing* è una funzione $f(\cdot)$ che da una $\mathbf{x}_i \in \mathcal{X}$ restituisce un prezzo P_i :

$$\begin{aligned} f : \mathcal{X} &\longrightarrow \mathbb{R}_+ \\ \mathbf{x}_i &\longmapsto P_i \end{aligned}$$

Modellare una variabile risposta

Modellare una variabile risposta Y_i significa stimare una funzione $r(\cdot)$ che da una $\mathbf{x}_i \in \mathcal{X}$ restituisce la distribuzione di Y_i o alcuni suoi momenti:

$$\begin{aligned} r : \mathcal{X} &\longrightarrow \mathcal{C} \\ \mathbf{x}_i &\longmapsto F_{Y_i}, E(Y_i), Var(Y_i) \end{aligned}$$



Variabili esplicative

Possibili variabili esplicative per il pricing delle assicurazioni motor:

- Informazioni sul veicolo assicurato;
- Informazioni generiche sul contraente;
- Informazioni assicurative sul contraente;
- Opzioni sulla polizza assicurativa;
- Dati telematici.

Queste variabili possono essere codificate come un vettore di numeri reali:

$$\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip}) \in \mathcal{X} \subseteq \mathbb{R}^p$$

Regola di Pricing

Una *Regola di Pricing* è una funzione $f(\cdot)$ che da una $\mathbf{x}_i \in \mathcal{X}$ restituisce un prezzo P_i :

$$\begin{array}{ccc} f : & \mathcal{X} & \longrightarrow R_+ \\ & \mathbf{x}_i & \longmapsto P_i \end{array}$$

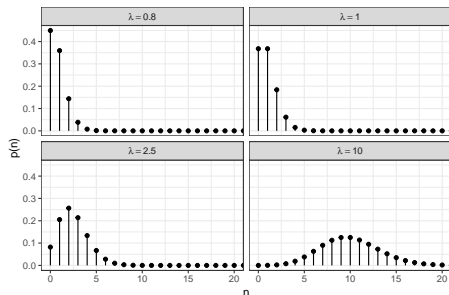
Modellare una variabile risposta

Modellare una variabile risposta Y_i significa stimare una funzione $r(\cdot)$ che da una $\mathbf{x}_i \in \mathcal{X}$ restituisce la distribuzione di Y_i o alcuni suoi momenti:

$$\begin{array}{ccc} r : & \mathcal{X} & \longrightarrow \mathcal{C} \\ & \mathbf{x}_i & \longmapsto F_{Y_i}, E(Y_i), Var(Y_i) \end{array}$$

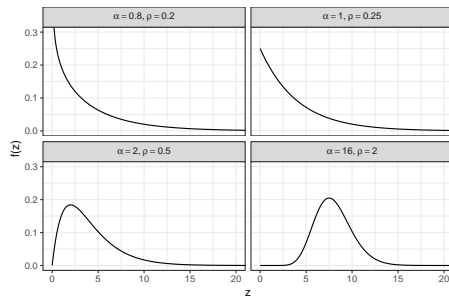
Distribuzione di Poisson

$$p_N(n) = P(N = n) = e^{-\lambda} \frac{\lambda^n}{n!}, \quad \lambda > 0$$



Distribuzione Gamma

$$f_Z(z) = \frac{\rho^\alpha}{\Gamma(\alpha)} z^{\alpha-1} e^{-\rho z}, \quad \alpha > 0, \rho > 0$$



Definizione di Premio

$$P_i^{(\text{risk})} = E(S_i)$$

$$P_i^{(\text{tech})} = E(S_i) + \text{Expenses}_i$$



Altri Caricamenti
Vincoli Normativi
Commercializzazioni

$$P_i^{(\text{tariff})}$$

$$P_i^{(\text{offer})} = P_i^{(\text{tariff})} - \text{Discount}_i$$

Ottimizzazione del Prezzo

Si basa su

- 1 Pricing Tecnico
- 2 Aspettativa del Cliente
- 3 Strategia di Business

Ulteriori modelli

- New Business: Probabilità di Conversion
- Rinnovi: Probabilità di Retention

Definizione di Premio

$$P_i^{(\text{risk})} = E(S_i)$$

$$P_i^{(\text{tech})} = E(S_i) + \text{Expenses}_i$$

↓
Altri Caricamenti
Vincoli Normativi
Commercializzazioni

$$P_i^{(\text{tariff})}$$

$$P_i^{(\text{offer})} = P_i^{(\text{tariff})} - \text{Discount}_i$$

Ottimizzazione del Prezzo

Si basa su

- 1 Pricing Tecnico
- 2 Aspettativa del Cliente
- 3 Strategia di Business

Ulteriori modelli

- New Business: Probabilità di Conversion
- Rinnovi: Probabilità di Retention

1. Il Pricing nelle Assicurazioni Danni

2. Modelli Statistici per il Pricing nelle Assicurazioni Danni

Modelli Lineari Generalizzati (GLM)

Modelli Additivi Generalizzati (GAM)

Stimatori Shrinkage per i GLM

Stimatori Bayesiani per i GLM

Algoritmi di Machine Learning

Confronto tra i modelli

3. Applicazione Pratica



Modelli Lineari Generalizzati (GLM)

Dato $\mathcal{D} = \{(\mathbf{x}_1, \omega_1, y_1), \dots, (\mathbf{x}_n, \omega_n, y_n)\}$

con $\mathbf{y} = (y_1, \dots, y_n)^t$ realizzazione di $\mathbf{Y} = (Y_1, \dots, Y_n)^t$.

Assumiamo che:

- 1 $\mathbf{Y} = (Y_1, \dots, Y_n)^t$ siano indipendenti con distribuzione appartenente a una stessa famiglia esponenziale lineare:

$$f(y_i; \theta_i, \phi, \omega_i) = \exp \left\{ \frac{\omega_i}{\phi} [y_i \theta_i - b(\theta_i)] \right\} c(y_i, \phi, \omega_i), \quad y_i \in \mathcal{Y} \subseteq \mathbb{R}$$

- 2 $\mathbf{x}_i = (1, x_{i1}, \dots, x_{ip})^t$ agisca su Y_i tramite il predittore lineare η_i

$$\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip}$$

- 3 η_i sia legato a $\mu_i = E(Y_i)$ tramite la funzione legame $g(\cdot)$

$$g(\mu_i) = \eta_i = \mathbf{x}_i^t \boldsymbol{\beta}$$



Stima di massima verosimiglianza

Data la funzione di verosimiglianza

$$\begin{aligned} L : \mathbb{R}^{p+1} \times \Lambda &\longrightarrow [0, +\infty[\\ (\boldsymbol{\beta}, \phi) &\longmapsto f_{\mathbf{Y}}(\mathbf{y}; \boldsymbol{\theta}, \phi) \end{aligned}$$

La stima di massima verosimiglianza è:

$$\hat{\boldsymbol{\beta}} = \arg \max_{\boldsymbol{\beta} \in \mathbb{R}^{p+1}} L(\boldsymbol{\beta}, \phi; \mathbf{y})$$

Devianza

La devianza è

$$D(\hat{\boldsymbol{\beta}}, \mathbf{y}) = -2\phi \left(\ell(\hat{\boldsymbol{\beta}}, \phi; \mathbf{y}) - \ell_s(\boldsymbol{\beta}^*, \phi; \mathbf{y}) \right)$$

dove $\ell(\hat{\boldsymbol{\beta}}, \phi; \mathbf{y}) = \log L(\hat{\boldsymbol{\beta}}, \phi; \mathbf{y})$
e $\boldsymbol{\beta}^*$ sono i parametri del modello saturo.

La stima di massima verosimiglianza può essere ottenuta come:

$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta} \in \mathbb{R}^{p+1}} D(\boldsymbol{\beta}, \mathbf{y})$$



Stima di massima verosimiglianza

Data la funzione di verosimiglianza

$$\begin{aligned} L : \mathbb{R}^{p+1} \times \Lambda &\longrightarrow [0, +\infty[\\ (\boldsymbol{\beta}, \phi) &\longmapsto f_{\mathbf{Y}}(\mathbf{y}; \boldsymbol{\theta}, \phi) \end{aligned}$$

La stima di massima verosimiglianza è:

$$\hat{\boldsymbol{\beta}} = \arg \max_{\boldsymbol{\beta} \in \mathbb{R}^{p+1}} L(\boldsymbol{\beta}, \phi; \mathbf{y})$$

Devianza

La devianza è

$$D(\hat{\boldsymbol{\beta}}, \mathbf{y}) = -2\phi \left(\ell(\hat{\boldsymbol{\beta}}, \phi; \mathbf{y}) - \ell_S(\boldsymbol{\beta}^*, \phi; \mathbf{y}) \right)$$

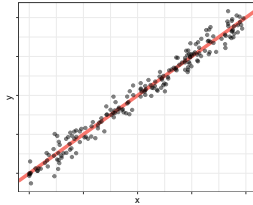
dove $\ell(\hat{\boldsymbol{\beta}}, \phi; \mathbf{y}) = \log L(\hat{\boldsymbol{\beta}}, \phi; \mathbf{y})$
e $\boldsymbol{\beta}^*$ sono i parametri del modello saturo.

La stima di massima verosimiglianza può essere ottenuta come:

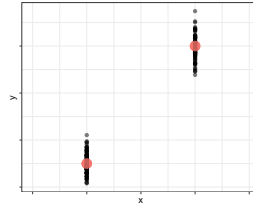
$$\hat{\boldsymbol{\beta}} = \arg \min_{\boldsymbol{\beta} \in \mathbb{R}^{p+1}} D(\boldsymbol{\beta}, \mathbf{y})$$

Effetto delle variabili in un GLM

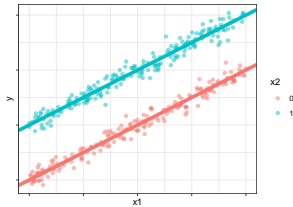
Quantitativa



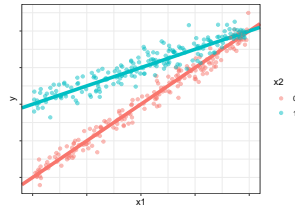
Qualitativa



Quantitativa e qualitativa
senza interazione

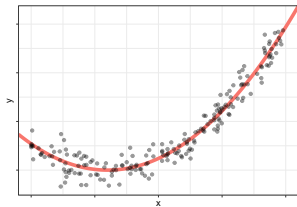


Quantitativa e qualitativa
con interazione

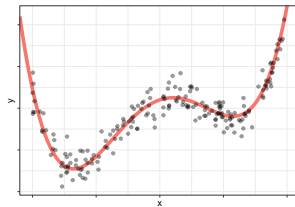


Variabili quantitative ed effetti non lineari

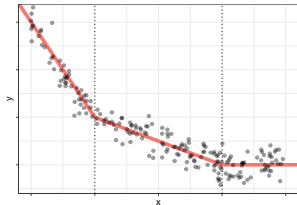
Polinomiale di grado 2



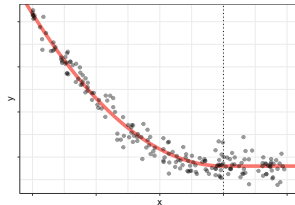
Polinomiale di grado 4



Piece-wise lineare

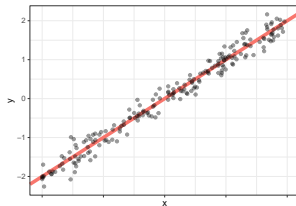


Piece-wise polinomiale di grado 2

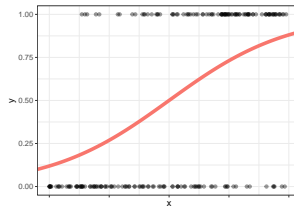


Funzione Legame e risposta

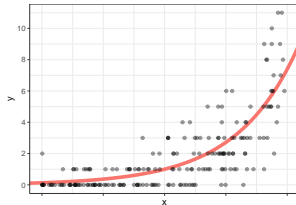
Normale - identità



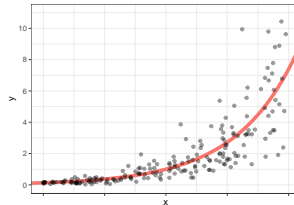
Binomiale - logit



Poisson - log

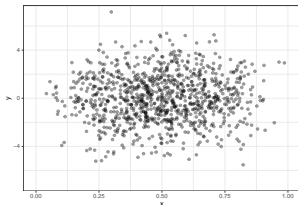


Gamma - log

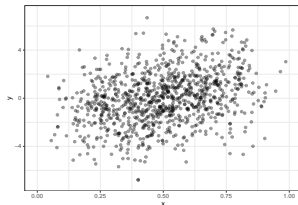


Grafici per visualizzare l'effetto delle variabili

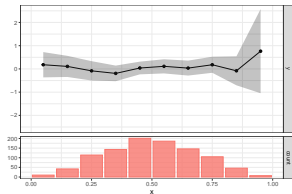
Nessun effetto - non raggruppati



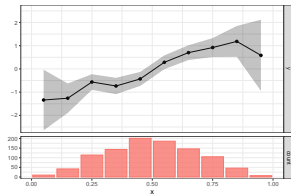
Effetto positivo - non raggruppati



Nessun effetto - raggruppati



Effetto positivo - raggruppati



Criteri per la selezione delle variabili

- Visualizzazione
- Test di verifica di ipotesi

$$\begin{cases} H_0 : \beta_{j_k} = 0 \quad \forall k \in \{1, 2, \dots, s\} \\ H_1 : \exists k : \beta_{j_k} \neq 0 \end{cases}$$

- Criteri di informazione

$$AIC = -2\ell(\boldsymbol{\beta}) + 2(p + 1)$$

$$BIC = -2\ell(\boldsymbol{\beta}) + \log(n)(p + 1)$$

- Divisione del dataset tra training set e test set
- Cross validation

\implies Algoritmi stepwise



1. Il Pricing nelle Assicurazioni Danni

2. Modelli Statistici per il Pricing nelle Assicurazioni Danni

Modelli Lineari Generalizzati (GLM)

Modelli Additivi Generalizzati (GAM)

Stimatori Shrinkage per i GLM

Stimatori Bayesiani per i GLM

Algoritmi di Machine Learning

Confronto tra i modelli

3. Applicazione Pratica



Modello Additivo Generalizzato (GAM)

- ❶ Variabile risposta \mathbf{Y} come GLM;
- ❷ Predittore lineare

$$\eta_i = \mathbf{x}_i^t \boldsymbol{\beta} + \sum_{l=1}^q f_l(z_{i,l}), \quad i \in \{1, 2, \dots, n\}$$

con $f_l(\cdot)$ spline cubica;

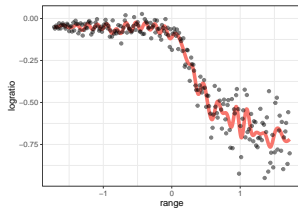
- ❸ Funzione legame $g(\cdot)$ come GLM.

Stima di Massima Verosimiglianza con Penalizzazione

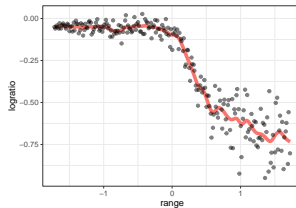
$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \left\{ D(\mathbf{f}, \mathbf{y}) + \sum_{l=1}^q \lambda_l \int_{a_l}^{b_l} (f_l''(x_l))^2 dx \right\}$$

con $\lambda_1, \lambda_2, \dots, \lambda_q$ iperparametri di smoothing.

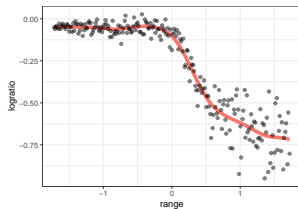
$\lambda = 0$



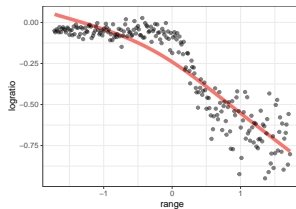
$\lambda = 10$



$\lambda = 10^3$



$\lambda = 10^6$



1. Il Pricing nelle Assicurazioni Danni

2. Modelli Statistici per il Pricing nelle Assicurazioni Danni

Modelli Lineari Generalizzati (GLM)

Modelli Additivi Generalizzati (GAM)

Stimatori Shrinkage per i GLM

Stimatori Bayesiani per i GLM

Algoritmi di Machine Learning

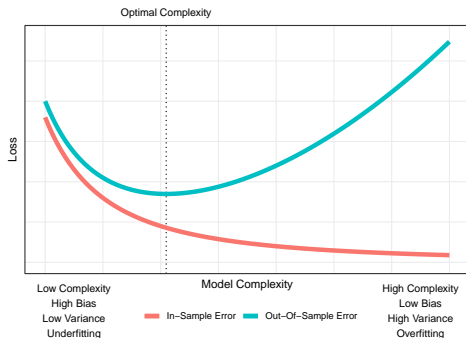
Confronto tra i modelli

3. Applicazione Pratica



Scomposizione dello scarto quadratico medio (MSE)

$$MSE(\tilde{\beta}_j) \stackrel{\text{def}}{=} E\left((\tilde{\beta}_j - \beta_j)^2\right) = \underbrace{\left(E(\tilde{\beta}_j) - \beta_j\right)^2}_{\text{Bias}^2} + \underbrace{\text{Var}(\tilde{\beta}_j)}_{\text{Variance}}$$



Stima di Massima Verosimiglianza con Penalizzazione

$$\hat{\beta} = \arg \min_{\beta \in \mathbb{R}^{p+1}} \left\{ D(\beta, \mathbf{y}) + \lambda \sum_{j=1}^p \beta_j^2 \right\}$$

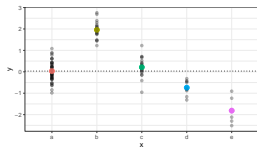
con

- $\lambda \geq 0$ iperparametro di penalizzazione

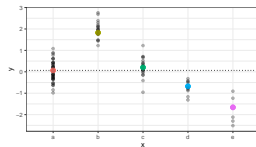
Modello sottostante: GLM

Regressione Ridge: esempio

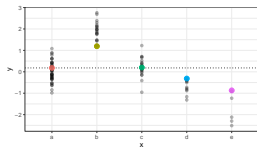
$\lambda = 0$



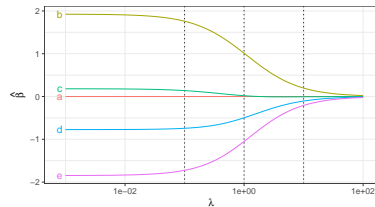
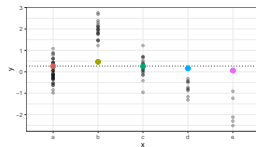
$\lambda = 0.1$



$\lambda = 1$



$\lambda = 10$



Stima di Massima Verosimiglianza con Penalizzazione

$$\hat{\beta} = \arg \min_{\beta \in \mathbb{R}^{p+1}} \left\{ D(\beta, \mathbf{y}) + \lambda \sum_{j=1}^p |\beta_j| \right\}$$

con

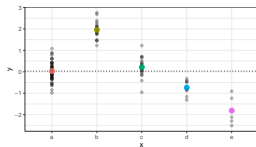
- $\lambda \geq 0$ iperparametro di penalizzazione

Modello sottostante: GLM

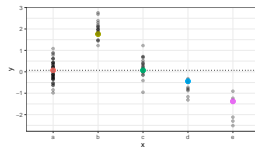


Regressione LASSO: esempio

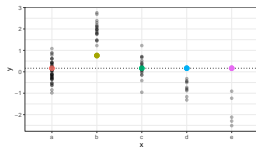
$\lambda = 0$



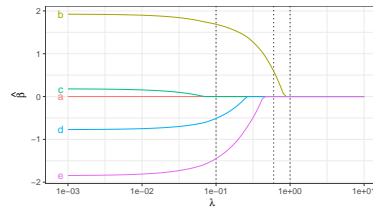
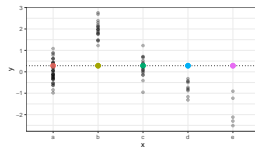
$\lambda = 0.1$



$\lambda = 1$



$\lambda = 10$



Stima di Massima Verosimiglianza con Penalizzazione

$$\hat{\beta} = \arg \min_{\beta \in \mathbb{R}^{p+1}} \left\{ D(\beta, \mathbf{y}) + \lambda \sum_{j=1}^p (\alpha |\beta_j| + (1 - \alpha) |\beta_j|^2) \right\}$$

dove

- $\lambda \geq 0$ iperparametro di penalizzazione
- $\alpha \in [0, 1]$ iperparametro che determina il peso della penalizzazione LASSO
 - ▶ $\alpha = 0 \implies$ Regressione Ridge
 - ▶ $\alpha = 1 \implies$ Regressione LASSO

Modello sottostante: GLM



1. Il Pricing nelle Assicurazioni Danni

2. Modelli Statistici per il Pricing nelle Assicurazioni Danni

Modelli Lineari Generalizzati (GLM)

Modelli Additivi Generalizzati (GAM)

Stimatori Shrinkage per i GLM

Stimatori Bayesiani per i GLM

Algoritmi di Machine Learning

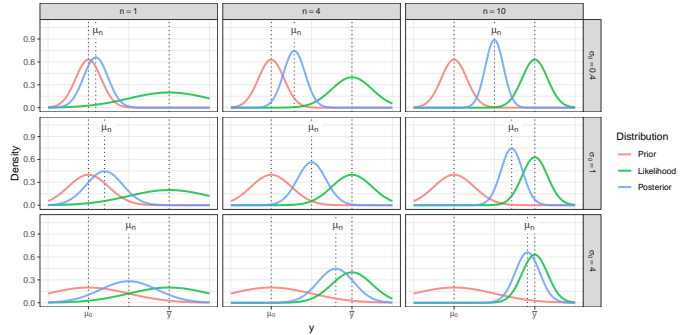
Confronto tra i modelli

3. Applicazione Pratica



Teorema di Bayes

$$\pi(\theta|\mathbf{y}) = \frac{p(\mathbf{y}|\theta)\pi(\theta)}{p(\mathbf{y})}$$



Stima di Massima Verosimiglianza

$$\begin{aligned}\hat{\beta}^{ML} &= \arg \max_{\beta \in \mathbb{R}^{p+1}} L(\beta, \phi \mid \mathbf{y}) \\ &= \arg \max_{\beta \in \mathbb{R}^{p+1}} \ell(\beta, \phi \mid \mathbf{y}) \\ &= \arg \min_{\beta \in \mathbb{R}^{p+1}} D(\beta, \mathbf{y})\end{aligned}$$

Stima di Massimo a Posteriori

$$\begin{aligned}\hat{\beta}^{MAP} &= \arg \max_{\beta \in \mathbb{R}^{p+1}} \pi(\beta, \phi \mid \mathbf{y}) \\ &= \arg \max_{\beta \in \mathbb{R}^{p+1}} \{L(\beta, \phi \mid \mathbf{y}) \pi(\beta)\} \\ &= \arg \max_{\beta \in \mathbb{R}^{p+1}} \{\ell(\beta, \phi \mid \mathbf{y}) + \log(\pi(\beta))\} \\ &= \arg \min_{\beta \in \mathbb{R}^{p+1}} \{D(\beta, \mathbf{y}) - 2\phi \log(\pi(\beta))\}\end{aligned}$$

Stima di Massima Verosimiglianza

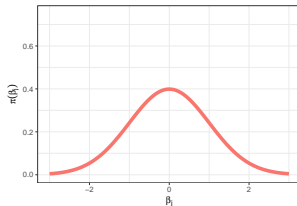
$$\begin{aligned}\hat{\beta}^{ML} &= \arg \max_{\beta \in \mathbb{R}^{p+1}} L(\beta, \phi \mid \mathbf{y}) \\ &= \arg \max_{\beta \in \mathbb{R}^{p+1}} \ell(\beta, \phi \mid \mathbf{y}) \\ &= \arg \min_{\beta \in \mathbb{R}^{p+1}} D(\beta, \mathbf{y})\end{aligned}$$

Stima di Massimo a Posteriori

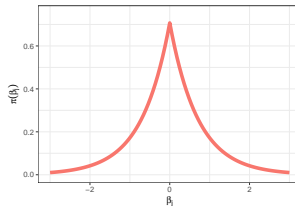
$$\begin{aligned}\hat{\beta}^{MAP} &= \arg \max_{\beta \in \mathbb{R}^{p+1}} \pi(\beta, \phi \mid \mathbf{y}) \\ &= \arg \max_{\beta \in \mathbb{R}^{p+1}} \{L(\beta, \phi \mid \mathbf{y}) \pi(\beta)\} \\ &= \arg \max_{\beta \in \mathbb{R}^{p+1}} \{\ell(\beta, \phi \mid \mathbf{y}) + \log(\pi(\beta))\} \\ &= \arg \min_{\beta \in \mathbb{R}^{p+1}} \{D(\beta, \mathbf{y}) - 2\phi \log(\pi(\beta))\}\end{aligned}$$

Regressione Ridge e LASSO come Stimatori Bayesiani

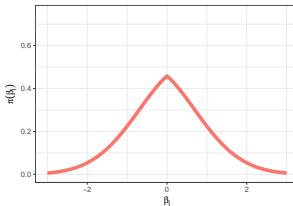
Distribuzione Normale
⇒ Regressione Ridge



Distribuzione di Laplace
⇒ Regressione LASSO



Distribuzione intermedia
⇒ Elastic Net



Altri distribuzioni a priori

- Diverse varianze a priori

$$\beta_j \sim \mathcal{N}(0, \sigma_j^2)$$

- Diverse medie a priori

$$\beta_j \sim \mathcal{N}(\beta_{j0}, \sigma_j^2)$$

- Altre distribuzioni a priori

$$\pi(\beta_j) = \begin{cases} \frac{\sqrt{2}}{\sqrt{\pi}\sigma} e^{-\frac{1}{2\sigma^2}\beta_j^2} & \text{if } \beta_j \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

Vantaggi stimatori Bayesiani

- Introduzione informazione esterna ai dati con una robusta metodologia statistica
- Rimpiazzamento degli offset
Scelgo σ_j^2 tale che $\hat{\beta}^{MAP} = \hat{\beta}^{\text{offset}}$
 - ▶ Ho accortezza di quanto è forte la correzione applicata
 - ▶ Se cambio qualche altro parametro e rifitto il modello, in automatico $\hat{\beta}^{MAP}$ viene ristimato

Altri distribuzioni a priori

- Diverse varianze a priori

$$\beta_j \sim \mathcal{N}(0, \sigma_j^2)$$

- Diverse medie a priori

$$\beta_j \sim \mathcal{N}(\beta_{j0}, \sigma_j^2)$$

- Altre distribuzioni a priori

$$\pi(\beta_j) = \begin{cases} \frac{\sqrt{2}}{\sqrt{\pi}\sigma} e^{-\frac{1}{2\sigma^2}\beta_j^2} & \text{if } \beta_j \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

Vantaggi stimatori Bayesiani

- Introduzione informazione esterna ai dati con una robusta metodologia statistica

- Rimpiazzamento degli offset

Scelgo σ_j^2 tale che $\hat{\beta}^{MAP} = \hat{\beta}^{\text{offset}}$

- ▶ Ho accortezza di quanto è forte la correzione applicata
- ▶ Se cambio qualche altro parametro e rifitto il modello, in automatico $\hat{\beta}^{MAP}$ viene ristimato

1. Il Pricing nelle Assicurazioni Danni

2. Modelli Statistici per il Pricing nelle Assicurazioni Danni

Modelli Lineari Generalizzati (GLM)

Modelli Additivi Generalizzati (GAM)

Stimatori Shrinkage per i GLM

Stimatori Bayesiani per i GLM

Algoritmi di Machine Learning

Confronto tra i modelli

3. Applicazione Pratica



Modelli di Machine Learning

- Gradient Boosting Machine (GBM)
- Random Forest (RF)
- Neural Network (NN)
- Altri ...

Caratteristiche

- Funzione di regressione con minime assunzioni

$$E(Y_i) = f(x_{i1}, \dots, x_{ip})$$

- Sofisticati algoritmi per prevenire l'overfitting



Modelli di Machine Learning

- Gradient Boosting Machine (GBM)
- Random Forest (RF)
- Neural Network (NN)
- Altri ...

Caratteristiche

- Funzione di regressione con minime assunzioni

$$E(Y_i) = \mathbf{f}(x_{i1}, \dots, x_{ip})$$

- Sofisticati algoritmi per prevenire l'overfitting

1. Il Pricing nelle Assicurazioni Danni

2. Modelli Statistici per il Pricing nelle Assicurazioni Danni

Modelli Lineari Generalizzati (GLM)

Modelli Additivi Generalizzati (GAM)

Stimatori Shrinkage per i GLM

Stimatori Bayesiani per i GLM

Algoritmi di Machine Learning

Confronto tra i modelli

3. Applicazione Pratica



GLM Classici vs GBM/RF/NN

- 👍 Maggior interpretabilità
- 👍 Maggior controllo delle variabili
 - ▶ Selezione variabili
 - ▶ Effetti variabili quantitative
 - ▶ Offset
- 👍 Facile utilizzo di informazioni esterne
- 👎 Minore automazione e scalabilità
- 👎 Minore flessibilità

GLM Advancements vs GLM Classici

- 👍 Mantenimento interpretabilità
- 👍 Mantenimento controllo delle variabili
- 👍 Miglior utilizzo delle informazioni esterne (stimatori bayesiani)
- 👍 Maggior automazione e scalabilità
 - ▶ Selezione variabili
 - ▶ Effetti variabili quantitative

GLM Classici vs GBM/RF/NN

- 👍 Maggior interpretabilità
- 👍 Maggior controllo delle variabili
 - ▶ Selezione variabili
 - ▶ Effetti variabili quantitative
 - ▶ Offset
- 👍 Facile utilizzo di informazioni esterne
- 👎 Minore automazione e scalabilità
- 👎 Minore flessibilità

GLM Advancements vs GLM Classici

- 👍 Mantenimento interpretabilità
- 👍 Mantenimento controllo delle variabili
- 👍 Miglior utilizzo delle informazioni esterne (stimatori bayesiani)
- 👍 Maggior automazione e scalabilità
 - ▶ Selezione variabili
 - ▶ Effetti variabili quantitative



1. Il Pricing nelle Assicurazioni Danni

2. Modelli Statistici per il Pricing nelle Assicurazioni Danni

Modelli Lineari Generalizzati (GLM)

Modelli Additivi Generalizzati (GAM)

Stimatori Shrinkage per i GLM

Stimatori Bayesiani per i GLM

Algoritmi di Machine Learning

Confronto tra i modelli

3. Applicazione Pratica



Set	Osservazioni	Esposizione (rischi anno)	Assicurati	Esposizione per Assicurato	Numero Sinistri	Frequenza Sinistri
Train	227 226	107 998.4	27 346	3.95	4 823	0.045
Test	56 603	26 806.3	6 824	3.93	1 131	0.042
Tot	283 829	134 804.7	34 170	3.95	5 954	0.044

Descrizione	Numero di variabili per categoria
Informazioni sul veicolo assicurato	12
Informazioni generiche sul contraente	14
Informazioni assicurative sul contraente	9
Opzioni sulla polizza assicurativa	11
Customer information on the policyholder	2
Dati telematici	4
Totale	52

Modelli considerati

Id	Model
Mod1	GLM Tot
Mod2	Elastic Net Tot
Mod3	Ridge Tot
Mod4	GLM AIC
Mod5	Elastic Net AIC
Mod6	GAM AIC
Mod7	GBM Tot

Valutazione

Metrica di confronto:

Devianza della distribuzione di Poisson
calcolata sul test set

$$D(\hat{\beta}, \mathbf{y}) = 2 \sum_{i=1}^n \left\{ y_i \log \left(\frac{y_i}{\hat{\mu}_i} \right) - (y_i - \hat{\mu}_i) \right\}$$

Modelli considerati

Id	Model
Mod1	GLM Tot
Mod2	Elastic Net Tot
Mod3	Ridge Tot
Mod4	GLM AIC
Mod5	Elastic Net AIC
Mod6	GAM AIC
Mod7	GBM Tot

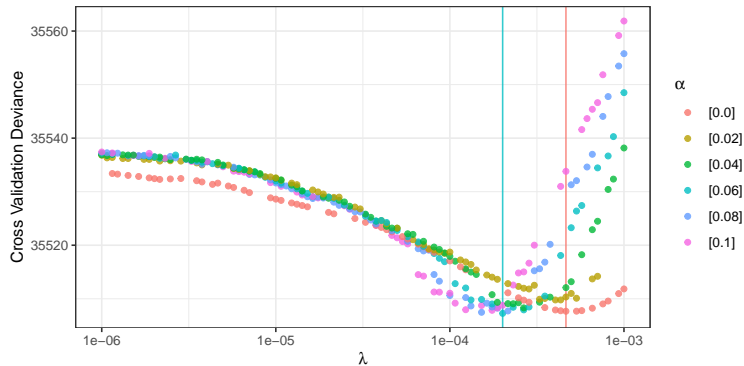
Valutazione

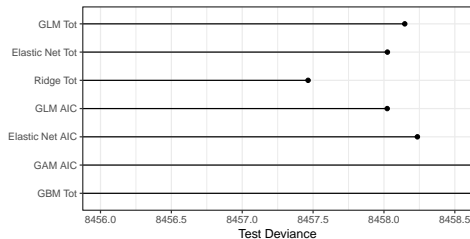
Metrica di confronto:

Devianza della distribuzione di Poisson
calcolata sul test set

$$D(\hat{\boldsymbol{\beta}}, \mathbf{y}) = 2 \sum_{i=1}^n \left\{ y_i \log \left(\frac{y_i}{\hat{\mu}_i} \right) - (y_i - \hat{\mu}_i) \right\}$$

Elastic Net: Tuning degli Iperparametri





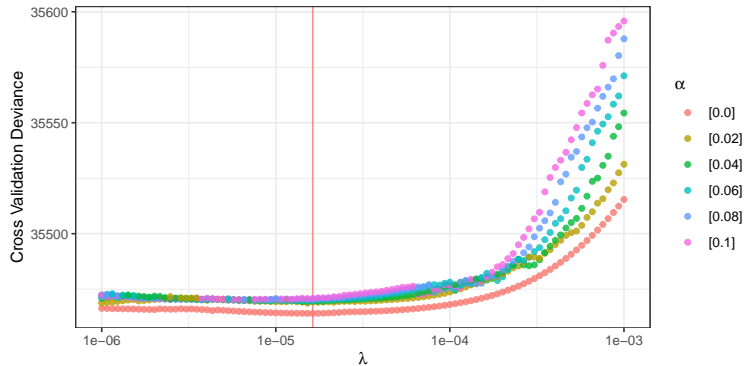
Id	Model	Test Deviance	Time	α	λ
Mod1	GLM Tot	8 458.147	2.7s	0	0
Mod2	Elastic Net Tot	8 458.024	1h 30m	0.06	2.01e-04
Mod3	Ridge Tot	8 457.465	1h 30m	0	4.64e-04
Mod4	GLM AIC	8 458.023	7h 27m	0	0
Mod5	Elastic Net AIC	8 458.236	8h 54m	0	1.63e-05
Mod6	GAM AIC	9 728.570	7h 45m	0	0
Mod7	GBM Tot	8 504.178	2h 30m		



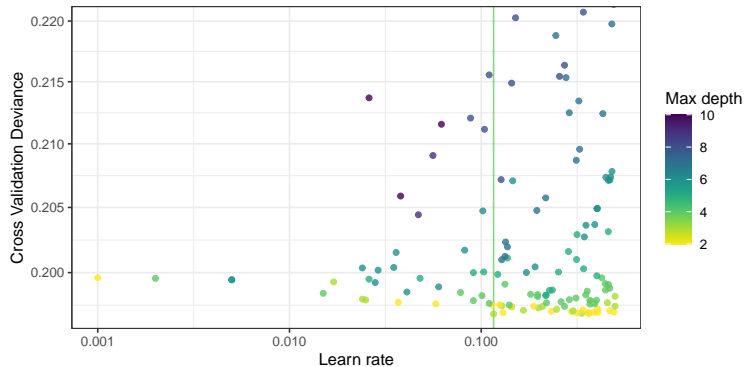
Backup



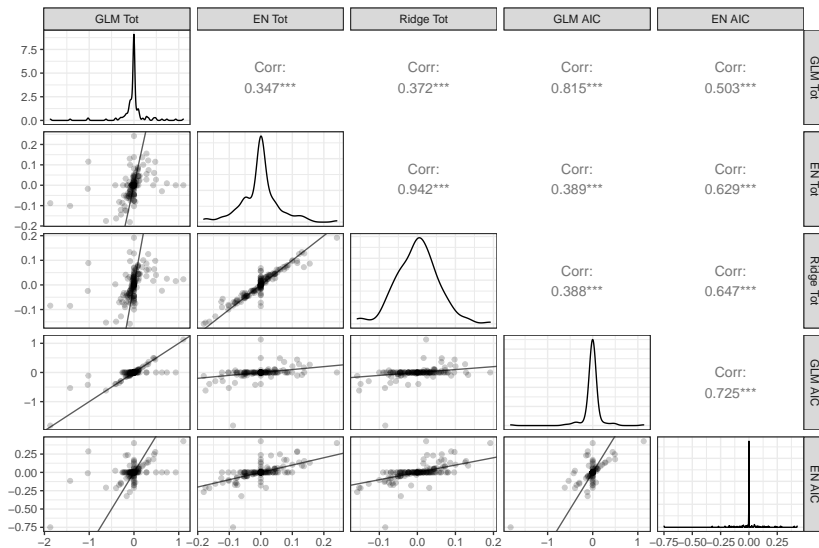
Elastic Net AIC: Tuning degli Iperparametri



Elastic Net AIC: Tuning degli Iperparametri



Confronto tra i coefficienti



model	value
GLM Tot	-6.447
EN Tot	0.000
Ridge Tot	-0.003
GLM AIC	-6.574
EN AIC	-0.095

GLM Tot	Elastic Net Tot	GLM AIC	n	
$\neq 0$	$\neq 0$	$\neq 0$	51	121
$\neq 0$	$\neq 0$	0	48	
$\neq 0$	0	$\neq 0$	6	
$\neq 0$	0	0	16	
0	$\neq 0$	0	23	38
0	0	0	15	

