Высш. шк., 2001. –
[6]. P. Krause, O.Wasynczuk. Analysis of electric machinery and drive systems. *IEEE* Press, 2002.
[7]. С. Г. Герман-Галкин, Г. А.Кардонов. *Электрические машины: Лабораторные работы на ГТК.* СПб.: КОРОНА принт, 2003.
[8]. А. О. Лозинський, В. І.Мороз, Я. С. Паранчук. Розв'язування задач електромеханіки в середовищах пакетів MathCAD і MATLAB. Львів: Видавництво НУ” ЛП, 2000.

METHODS OF ESTIMATING PARAMETERS OF GENERALIZED LINEAR MODELS IN THE ANALYSIS OF ACTUARIAL RISKS

Roman PANIBRATOV[1]

National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine, ORCID 0000-0002-8604-4420, e-mail: roman.panibratov@gmail.com

Actuarial risk is defined as the possibility that the actuaries' assumptions used to develop the model for calculating the costs of insurance policies are unreliable or erroneous. The phrase "insurance risk" is another name for it. The accuracy of the assumptions employed in the pricing models that insurance firms use to determine premium prices directly relates to the amount of actuarial risk. The cost of insurance policies is determined by probability estimations, which enables insurers to execute payments subject to routine business activities. If the offered assumptions are incorrect, occurrences that were overlooked cause payments to be made more frequently, which has major financial ramifications for the insurer.

Making explicit assumptions about the nature of the insurance data and how it relates to the expected variables is possible with generalized linear models (GLM). In addition, GLM offers statistical diagnostics that support the process of identifying the important variables and validating model assumptions. This method is widely acknowledged as the norm for pricing various insurance products in various markets and countries.

The GLM includes a wide range of models, with the linear regression model being one of many particular cases. The latter's assumptions, which often include normal distribution, constant variance, and additive effects, are disproved. The target variable, for instance, might be chosen from the exponential family of distributions[1].

The general form of the exponential family of distributions is as follows:

$$f_i(y_i, \theta_i, \varphi) = \exp\{\frac{y_i\theta_i - b(\theta_i)}{a_i(\varphi)} + c(y_i, \varphi)\}$$

where $a_i(\varphi)$, $b(\theta_i)$, та $c(y_i, \varphi)$ are functions, that are defined at the beginning;

θ_i is a parameter related to the mean value;

φ is a scale parameter related to variance.

The variance and the distribution's mean are both subject to change. It is expected that explanatory factors have an additive effect on a different scale. For GLM, the following presumptions are made:

1. Stochastic component: each independent component of Y comes from the same exponential family distribution.

2. Systematic component: a linear predictor η is formed using p covariates (explanatory variables)

$$\eta = X\beta$$

3. Link function: a differentiable and monotonic link function is used to establish the connection between the random and systematic components.

$$E[Y] = \mu = g^{-1}(\eta)$$

The estimation of GLM parameters is a significant problem that requires sufficient attention. For the aim of comparing techniques, the following algorithms were employed to assess the parameters: Weighted least squares iterative-recursive approach (IRWLS), Adaptive moment estimation (Adam) optimization algorithm, and Markov chain Monte Carlo method (MCMC).

A full description of these algorithms is given in the works [2][3][4].

It was chosen to create insurance indicators and target variables at random since insurance data is not always readily accessible to the general public. The variables in the data were as follows:

— Age (range from 18 to 64 years);
— Sex;
— Body mass index (normal distribution was used);
— Number of children (range from 0 to 5);
— Smoker status;

- Region (generated from sample view ['north', 'south', 'east', 'west', 'center']);
- Charges.

The target is the final variable, and the following distribution laws and associated link functions were applied to it:

- normal distribution with a known variance σ and a logarithmic link function;
- exponential distribution with the identity link function;
- Pareto distribution with a known scale parameter x_m and a link function of the form

$$f(x) = -1 - x$$

To estimate the model quality, the following metrics were used: Mean squared error (MSE), Root mean squared error (RMSE) and Mean absolute error(MAE).

Tables 1, 2, and 3 demonstrates the results of the estimate of GLM parameters using the three approaches.

Table 1 – Results of GLM construction for a target variable with a known variance Gaussian distribution and a logarithmic link function

Metric	MCMC	ADAM	IRWLS
MSE	4410.78	686.65	2458.21
RMSE	66.41	26.20	49.58
MAE	51.88	25.81	49.58

Table 2 – Results of GLM construction for a target variable with a known scale parameter of Pareto distribution and a negative linear link function

Metric	MCMC	ADAM	IRWLS
MSE	52725.56	82188.07	638121.57
RMSE	229.62	286.69	798.83
MAE	154.02	205.31	773.33

Table 3 – Results of GLM construction for exponential distribution and an identity link function

Metric	MCMC	ADAM	IRWLS
MSE	213.52	148.95	288.53
RMSE	14.61	12.20	16.98
MAE	11.17	3.64	15.66

The Adam technique produced generally fairly good results, as can be shown from the GLM building results for three cases. The Pareto distribution case was another one where the MCMC approach performed well. After using models in real-world situations to address the issue of foreseeing potential losses, the final parameter estimate approach is decided upon.

References

[1]. D. Anderson et al, A Practitioner's Guide to Generalized Linear Models – a foundation for theory, interpretation and application, 3rd edition.Towers Watson, 2007, 122 p.

[2]. P. McCullagh and J. Nelder, Generalized Linear Models, 2nd edition. Chapman & Hall, 1989, 532 p.

[3]. D. Kingma, and J. Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014. DOI: 10.48550/arXiv.1412.6980.

[4]. V. Roy. MCMC for GLMMs. arXiv preprint arXiv:2204.01866, 2022, DOI: 10.48550/arXiv.2204.01866