Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
May-June 2022 Page No. 1015-1022
Article History: Received: 24 February 2022, Revised: 05 April 2022, Accepted: 08 May 2022, Publication: 30 June 2022

Probability Model of Mobile Phone Tariff Selection Based on Poisson Regression Model

Yanhong Wu[1], Jianxia Bai[2]*, Jianfeng Liu[1], Juanjuan Song[1]
[1] Foundation Department, Shandong Huayu University of Technology, Shandong 253000, China
[2]Department of Mathematics, Tianjin Renai College, Tianjin 301636, China
*Corresponding Author.

Abstract:

With the deepening of the informatization degree of social life, mobile phones have become an indispensable tool in People's Daily life, study and work. In order to guide people to make more scientific choices, some reasonable models are needed as a reference. The choice of mobile phone tariff is a decision-making problem that people often encounter. For companies, how to set mobile phone rates and which mobile phone rates are more likely to be selected by whom are the decisions they need to make. Determine these problems more effectively in order to research, this article through early to collect data, the Poisson regression model was established, using the largest eigenvalue method and maximum likelihood estimation method to determine the parameters in the model, and finally determine the established model, on the basis of this model can calculate the probability of choosing a mobile phone charges.

***Keywords**: Poisson regression model; Probability; Maximum likelihood method; Mobile phone chargesestimation*

I. INTRODUCTION

Suppose $Y_1, Y_2, \cdots Y_m$ are independent of each other. $Y_k \sim P(\lambda_k)$.That is

$Var(Y_k) = E(Y_k) = \lambda_k$. Suppose

$$P(Y_k = y_k) = \frac{\exp(-\lambda_k)\lambda_k^{y_k}}{\Gamma(1+y_k)} \tag{1}$$

$\lambda_k = \exp(x_k^T \beta)$, $X_1, X_2, \cdots, X_m$ are independent of each other. $X_k \sim N(\mu_k, \Sigma_k)$ [1-6], $\mu_k, \Sigma_k > 0$ are unknown. x_k is the value of X_k . $x_k = (x_{k1}, x_{k2} \cdots, x_{kp})^{\mathrm{T}}$ is the $p \times 1$ dimensional column vector. $\beta = (\beta_1, \beta_2, \cdots, \beta_p)^{\mathrm{T}}$ is a vector of $p \times 1$ dimensional parameters.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
May-June 2022 Page No. 1015-1022
Article History: Received: 24 February 2022, Revised: 05 April 2022, Accepted: 08 May 2022, Publication: 30 June 2022

Poisson regression model[7-16] is widely used in real life. For example, there are n types of policies, and each type of policy has p rate factors.

II. POISSON REGRESSION MODEL

2.1. Model Establishment

The overall value of (X,Y) is shown in Table 1, where Y represents the sold quantity of various mobile phone charges, and X represents the monthly function fee (yuan), domestic data traffic (G), and call duration (min) of the tariff in sale:

TABLE I. Types of mobile phone charges and sales quantity

Y	7	9	7	11	9	9	4	8	4	8
X	$\begin{pmatrix}58\\0.15\\150\end{pmatrix}$	$\begin{pmatrix}88\\0.3\\350\end{pmatrix}$	$\begin{pmatrix}128\\0.6\\650\end{pmatrix}$	$\begin{pmatrix}158\\0.6\\900\end{pmatrix}$	$\begin{pmatrix}188\\0.6\\1200\end{pmatrix}$	$\begin{pmatrix}58\\0.5\\50\end{pmatrix}$	$\begin{pmatrix}88\\0.7\\200\end{pmatrix}$	$\begin{pmatrix}128\\1\\420\end{pmatrix}$	$\begin{pmatrix}158\\2\\510\end{pmatrix}$	$\begin{pmatrix}188\\2.5\\600\end{pmatrix}$

Poisson regression model is constructed as follows:

Suppose $Y_1,Y_2,\cdots Y_{10}$ are independent of each other. $Y_k \sim P(\lambda_k)$.That is $Var(Y_k)=E(Y_k)=\lambda_k$. Suppose

$$P(Y_k=y_k)=\frac{\exp(-\lambda_k)\lambda_k^{y_k}}{\Gamma(1+y_k)} \quad (k=1,2,\cdots,10) \tag{2}$$

$\lambda_k=\exp(x_k^T\beta)$, $X_1,X_2,\cdots,X_{10}$ are independent of each other. $X_k \sim N(\mu_k,\Sigma_k)$, μ_k , $\Sigma_k>0$ are unknown. x_k is the value of X_k . $x_k=(x_{k1},x_{k2},x_{k3})^{\mathrm{T}}$ is the eigenvariable column vector. $\beta=(\beta_1,\beta_2,\beta_3)^{\mathrm{T}}$ is a vector of 3×1 dimensional parameters.

2.2. Estimation of parameter unit vector

There are 100 samples $(X^{(i)},Y^{(i)})$ $(i=1,2,\cdots100)$ in the model population (X,Y). The distribution of sample quantity is shown in Table 2.

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
May-June 2022 Page No. 1015-1022
Article History: Received: 24 February 2022, Revised: 05 April 2022, Accepted: 08 May 2022, Publication: 30 June 2022

TABLE II. Types of mobile phone charges and the frequency

Y	**7**	**9**	**7**	**11**	**9**	**9**	**4**	**8**	**4**	**8**
X	$\begin{pmatrix}58\\0.15\\150\end{pmatrix}$	$\begin{pmatrix}88\\0.3\\350\end{pmatrix}$	$\begin{pmatrix}128\\0.6\\650\end{pmatrix}$	$\begin{pmatrix}158\\0.6\\900\end{pmatrix}$	$\begin{pmatrix}188\\0.6\\1200\end{pmatrix}$	$\begin{pmatrix}58\\0.5\\50\end{pmatrix}$	$\begin{pmatrix}88\\0.7\\200\end{pmatrix}$	$\begin{pmatrix}128\\1\\420\end{pmatrix}$	$\begin{pmatrix}158\\2\\510\end{pmatrix}$	$\begin{pmatrix}188\\2.5\\600\end{pmatrix}$
frequency	**6**	**12**	**16**	**10**	**7**	**5**	**10**	**11**	**16**	**7**

Let $\hat{\mu}_l = X^{(l)} - \bar{X}$, the values of $\hat{\mu}_l$ are shown in Table 3.

TABLE III. The values of $\hat{\mu}_l$

l	**1**	**2**	**3**	**4**	**5**
$\hat{\mu}_l$	$\begin{pmatrix}-69.7\\-0.725\\-371.3\end{pmatrix}$	$\begin{pmatrix}-39.7\\-0.575\\-171.3\end{pmatrix}$	$\begin{pmatrix}0.3\\-0.275\\128.7\end{pmatrix}$	$\begin{pmatrix}30.3\\-0.275\\378.7\end{pmatrix}$	$\begin{pmatrix}60.3\\-0.275\\678.7\end{pmatrix}$
l	**6**	**7**	**8**	**9**	**10**
$\hat{\mu}_l$	$\begin{pmatrix}-69.7\\-0.375\\-471.3\end{pmatrix}$	$\begin{pmatrix}-39.7\\0.125\\-321.3\end{pmatrix}$	$\begin{pmatrix}0.3\\1.125\\-101.3\end{pmatrix}$	$\begin{pmatrix}30.3\\-0.175\\-11.3\end{pmatrix}$	$\begin{pmatrix}60.3\\1.625\\78.7\end{pmatrix}$

$$\hat{V}_1 = \sum_{l=1}^{10} \frac{N_l}{100} \hat{\mu}_l \hat{\mu}_l^{\mathrm{T}} = \begin{pmatrix} 1628.91 & 10.6225 & 9579.39 \\ 10.6225 & 0.433725 & 0.3825 \\ 9579.39 & 0.3825 & 84041.31 \end{pmatrix} \quad (3)$$

$$\hat{\Sigma} = \frac{1}{10} \sum_{l=1}^{10} \hat{\mu}_l \hat{\mu}_l^{\mathrm{T}} = \begin{pmatrix} 2197.69 & 16.256 & 13509.71 \\ 16.256 & 0.517625 & 19.399 \\ 13509.71 & 19.399 & 112975.89 \end{pmatrix} \quad (4)$$

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
May-June 2022 Page No. 1015-1022
Article History: Received: 24 February 2022, Revised: 05 April 2022, Accepted: 08 May 2022, Publication: 30 June 2022

$$\hat{V} = \hat{\Sigma}^{-1}\hat{V}_1\hat{\Sigma}^{-1} = \begin{pmatrix} 0.0067 & -0.1938 & -0.0008 \\ -0.1938 & 7.1993 & 0.0218 \\ -0.0008 & 0.0218 & 0.0001 \end{pmatrix}, \quad (5)$$

The maximum characteristic root of $\hat{V}$ is λ_1=7.2046 , and its corresponding feature vector is $\hat{\eta} = \begin{pmatrix} -0.0269 \\ 0.9996 \\ 0.0030 \end{pmatrix}$.

Estimation of unit vector $\beta' = \frac{\beta}{\|\beta\|}$ of parameter β is $\hat{\beta}' = \hat{\eta} = \begin{pmatrix} -0.0269 \\ 0.9996 \\ 0.0030 \end{pmatrix}$.

2.3. Estimation of parameter vector modulus

The maximum likelihood estimation method is used to estimate the modulus α of parameters β, and the Poisson regression model constructed in this paper can be rewritten as:

$$P(Y_k = y_k \mid X = x) = \frac{\exp\left[y_k x_k^T \beta'\alpha - \exp\left(x_k^T \beta'\alpha\right)\right]}{\Gamma(1+y_k)} \quad (6)$$

Replacing the estimator $\hat{\beta}'$ of the parameter unit vector β' in the above equation, Then we get:

$$P(Y_k = y_k \mid X = x) = \frac{\exp\left[y_k x_k^T \hat{\beta}'\alpha - \exp\left(x_k^T \hat{\beta}'\alpha\right)\right]}{\Gamma(1+y_k)} \quad (7)$$

Starting from the above formula, the joint density can be obtained as:

$$\begin{aligned} f(x,y) &= f_{Y|X}(y) f_X(x) \\ &= \frac{\exp\left[y_k x_k^T \hat{\beta}'\alpha - \exp\left(x_k^T \hat{\beta}'\alpha\right)\right]}{\Gamma(1+y_k)} \\ & \prod_{j=1}^{100} \frac{1}{\left(\sqrt{2\pi}\right)^3} \left|\Sigma_j\right|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(x_j-\mu_j\right)^T \Sigma_j^{-1}\left(x_j-\mu_j\right)\right] \end{aligned} \quad (8)$$

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
May-June 2022 Page No. 1015-1022
Article History: Received: 24 February 2022, Revised: 05 April 2022, Accepted: 08 May 2022, Publication: 30 June 2022

Omitting the irrelevant part, and getting the likelihood function:

$$L\left(\alpha; y_1,\cdots,y_{100}\right)=\prod_{k=1}^{100}\frac{\exp\left[y_k x_k^{\mathrm{T}}\hat{\beta}'\alpha-\exp\left(x_k^{\mathrm{T}}\hat{\beta}'\alpha\right)\right]}{\Gamma\left(1+y_k\right)} \tag{9}$$

The logarithmic likelihood function is

$$\begin{aligned} l\left(\alpha; y_1,\cdots,y_{100}\right)&=\sum_{k=1}^{100}\ln\frac{\exp\left[y_k x_k^{\mathrm{T}}\hat{\beta}'\alpha-\exp\left(x_k^{\mathrm{T}}\hat{\beta}'\alpha\right)\right]}{\Gamma\left(1+y_k\right)} \\ &=\sum_{k=1}^{100}\left[y_k x_k^{\mathrm{T}}\hat{\beta}'\alpha-\exp\left(x_k^{\mathrm{T}}\hat{\beta}'\alpha\right)\right]-\sum_{k=1}^{100}\ln\left[\Gamma\left(1+y_k\right)\right] \end{aligned} \tag{10}$$

The likelihood equation is

$$\sum_{k=1}^{100}\left[y_k x_k^{\mathrm{T}}\hat{\beta}'-x_k^{\mathrm{T}}\hat{\beta}'\cdot\exp\left(x_k^{\mathrm{T}}\hat{\beta}'\alpha\right)\right]=0 \tag{11}$$

By calculating the data in the paper, the results are shown in Table 4

TABLE IV. Product

Product	x_1^{T}	x_2^{T}	x_3^{T}	x_4^{T}	x_5^{T}
$\hat{\beta}'$	-0.9570	-1.0090	-0.8773	-0.9278	-0.8269
Product	x_6^{T}	x_7^{T}	x_8^{T}	x_9^{T}	x_{10}^{T}
$\hat{\beta}'$	-0.9100	-0.7635	-0.1743	-2.0089	-0.7445
Product	$6x_1^{\mathrm{T}}$	$12x_2^{\mathrm{T}}$	$16x_3^{\mathrm{T}}$	$10x_4^{\mathrm{T}}$	$7x_5^{\mathrm{T}}$
$\hat{\beta}'$	-5.7423	-12.1078	-14.0373	-9.2780	-5.7880
Product	$5x_6^{\mathrm{T}}$	$10x_7^{\mathrm{T}}$	$11x_8^{\mathrm{T}}$	$16x_9^{\mathrm{T}}$	$7x_{10}^{\mathrm{T}}$
$\hat{\beta}'$	-4.5500	-7.6348	-1.9178	-32.1417	-5.2116

Accordingly, the likelihood equation can be written as

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
May-June 2022 Page No. 1015-1022
Article History: Received: 24 February 2022, Revised: 05 April 2022, Accepted: 08 May 2022, Publication: 30 June 2022

$$\begin{aligned}&7\times(-5.7423)+9\times(-12.1078)+7\times(-14.0373)+11\times(-9.2780)\\&+9\times(-5.7880)+9\times(-4.5500)+4\times(-7.6348)+8\times(-1.9178)\\&+4\times(-32.1417)+8\times(-5.2116)+100\times0.9570\times\exp(-0.9570\alpha)\\&+100\times1.0090\times\exp(-1.0090\alpha)+100\times0.8773\times\exp(-0.8773\alpha)\\&+100\times0.9278\times\exp(-0.9278\alpha)+100\times0.8269\times\exp(-0.8269\alpha)\\&+100\times0.9100\times\exp(-0.9100\alpha)+100\times0.7635\times\exp(-0.7635\alpha)\\&+100\times0.1743\times\exp(-0.1743\alpha)+100\times2.0089\times\exp(-2.0089\alpha)\\&+100\times0.7445\times\exp(-0.7445\alpha)=0\end{aligned}\quad(12)$$

After calculation, the result is

$$\begin{aligned}&658.6686-95.70\times\exp(-0.9570\alpha)\\&-100.90\times\exp(-1.0090\alpha)-87.73\times\exp(-0.8773\alpha)\\&-92.78\times\exp(-0.9278\alpha)-82.69\times\exp(-0.8269\alpha)\\&-91\times\exp(-0.9100\alpha)-76.35\times\exp(-0.7635\alpha)\\&-17.43\times\exp(-0.1743\alpha)-200.89\times\exp(-2.0089\alpha)\\&-74.45\times\exp(-0.7445\alpha)=0\end{aligned}\quad(13)$$

Let $A=e^{\alpha}$,the above formula can be transformed into

$$\begin{aligned}&658.6686-95.70\times A^{-0.9570}-100.90\times A^{-1.0090}-87.73\times A^{-0.8773}\\&-92.78\times A^{-0.9278}-82.69\times A^{-0.8269}-91\times A^{-0.9100}-76.35\times A^{-0.7635}\\&-17.43\times A^{-0.1743}-200.89\times A^{-2.0089}-74.45\times A^{-0.7445}=0\end{aligned}\quad(14)$$

Multiply both ends of the equation by $A^{3.0089}$,Then get the equation

$$\begin{aligned}&658.6686\times A^{3.0089}-95.70\times A^{2.0519}-100.90\times A^{1.9999}-87.73\times A^{2.1316}\\&-92.78\times A^{2.0811}-82.69\times A^{2.1820}-91\times A^{2.0989}-76.35\times A^{2.2454}\\&-17.43\times A^{2.8346}-74.45\times A^{2.2644}=200.89\times A\end{aligned}\quad(15)$$

Divide both ends of the equation by 200.89, we get

$$\begin{aligned}&3.2788\times A^{3.0089}-0.4764\times A^{2.0519}-0.5023\times A^{1.9999}-0.4367\times A^{2.1316}\\&-0.4618\times A^{2.0811}-0.4116\times A^{2.1820}-0.4529\times A^{2.0989}-0.3801\times A^{2.2454}\\&-0.0868\times A^{2.8346}-0.3706\times A^{2.2644}=A\end{aligned}\quad(16)$$

The iterative method is used to obtain A=1.3644, that is e^{α}=1.3644, then $\hat{\alpha}$=0.3107.

III. CONSISTENCY ESTIMATION OF PARAMETER VECTORS

According to Slutsky's theorem, the estimation of parameter vector β is consistent estimation:

Forest Chemicals Review
www.forestchemicalsreview.com
ISSN: 1520-0191
May-June 2022 Page No. 1015-1022
Article History: Received: 24 February 2022, Revised: 05 April 2022, Accepted: 08 May 2022, Publication: 30 June 2022

$$\hat{\beta} = \hat{\alpha} \cdot \hat{\beta} = 0.3107 \begin{pmatrix} -0.02699 \\ 0.99966 \\ 0.00300 \end{pmatrix} = \begin{pmatrix} -0.008 \\ 0.31 \\ 0.0009 \end{pmatrix} \quad (17)$$

VI. CONCLUSION

Based on the analysis of the sales data of the tariff package launched by a mobile company, this paper chooses to use the Poisson model to calculate the probability that customers would choose a certain package. After building the model, the unit vector estimation of parameters is obtained by using the method of maximum eigenvalue and eigenvector, and the modulus estimation of parameters is obtained by using the maximum likelihood estimation method, so as to obtain the consistency estimation of parameters. Through this Poisson regression model, it is easy to determine the probability of each mobile phone tariff package. This model also has a very wide range of applications in all aspects of society.

ACKNOWLEDGEMENTS

This research was supported by the project grant from the second first-class undergraduate course of Shandong Huayu University of Technology “complex function and integral transform”(Grand No.2022YL-18), the second first-class undergraduate course of Shandong Huayu University of Technology “Probability theory and Mathematical Statistics”(Grand No.2022YL-19) , National Natural Science Foundation of China key Project: Experimental Study on Lagrangian Turbulence Structure and its Influence on Transport diffusion (No.11732010). Scientific research project of Tianjin Education Commission "The Evolution of Community in Text-rich Social Network" (Project No. 2021KJ083). Research project of Tianjin Renai College “Research on the Active Control of Turbulent Boundary Layer drag Reduction by Piezoelectric Actuator" (Project No. XX18002).

REFERENCES

[1]Lv Ping. Application of weighting in variance estimation of complex surveys. Beijing: Statistical Research, 2011.
[2]MAO Shisong, Wang Jinglong, Pu Xiaolong. Advanced mathematical statistics. 3rd edition. Beijing: Higher Education Press, 2006.
[3]. Song Zhaoji, Xu Liumei. MATLAB 6.5 in the application of scientific calculation. Beijing: Tsinghua University Press, 2005.
[4]. Wang Songgui, Shi Jianhong, Yin Suju, Wu Mixia. Introduction to linear Models. Beijing: Science Press, 2004.
[5]. Jia Lanxiang, Zhang Jianhua, Linear Algebra. Tianjin: Nankai University Press, 2004.
[6]. Yan Qingjin. Numerical analysis. Beijing: Beihang University Press, 2006.
[7]. R D Gupta, D Kundu. Generalized exponential distributions: Existing results and some recent developments. Journal of Statistical Planning and Inference, vol. 137, 2007, pp. 3525-3536.
[8]. Sarhan, N Balakrishnan. A new class of bivariate distribution and its mixture. Journal of the Multivariate Analysis, vol. 98, 2007, pp. 1508-1527.
[9]. Ming-Tien Tsai. Maximum likelihood estimation of covariance matrices under simple tree ordering. Journal of Multivariate Analysis, vol. 89, 2004, pp. 292-303.

[10]. Ming-Tien Tsai. Maximum likelihood estimation of Wishart mean matrices under lowner order restrictions. Journal of Multivariate Analysis, vol. 98, 2007, pp. 932-944.

[11]. NADWA K R, ALGAMAL Z Y. A new ridge estimator for the Poisson regression model. Iranian Journal of Science and Technology, vol. 6, no. 43, 2019, pp. 2921-2 928

[12].MERVE K C, SELAHATTIN K. Improved twoparameter estimators for the negative binomial and Poisson regression model. Journal of Statistical Computation and Simulation, vol. 14, no. 89, 2019, pp. 2 645-2660.

[13].QASIM M, KIBRIA B M G, MANSSON K, et al. A new Poisson Liu regression estimator: Method and application. Journal of Applied Statistics, vol. 4, no. 6, 2019, pp. 1-14.

[14]. NOORIASL M, BEVRANI H, BELAGHI R A. Penalized and ridge type shrinkage estimators in Poisson regression model. Communications in Statistics-Simulation and Computation, vol. 3, no. 3, Sept. 2020, pp. 1-18.

[15]. NOORIASL M, BEVRANI H, BELAGHI R A. Penalized and ridge type shrinkage estimators in Poisson regression model. Communications inStatistics-Simulation and Computation, vol. 3, no. 3, Sept. 2020, pp. 1-18.

[16]Amin M, Qasim M, Amanullah M. Performance of Asar and Genç and Huang and Yang's Two Parameter Estimation Methods for the Gamma Regression Model. Iranian Journal of Science and Technology, Transactions A: Science, vol. 6, no. 43, 2019, pp. 2951-2963.