

Research on Influencing Factors of Residents Watching Cross Talk based on Logistic Regression

Yu Pei^a, Xiang Zhang, Yu Ding

School of Statistics and Applied Mathematics, Anhui University of Finance and Economics,
Bengbu 233030, China

^azy35848@163.com

Abstract

Cross talk, a folk rap art in the form of speaking, learning, amusing and singing, was included in the first batch of intangible cultural heritage in Beijing in 2006. Starting from the dimensions of basic information, viewing cost, actor popularity, and performance theme, Logistic regression is used to analyze the influencing factors of residents' choice of cross talk. Finally, using text mining, qualitative comparative analysis and other methods to clarify the successful path of Deyun Club from the three dimensions of star effect, cultural effect and marketing effect.

Keywords

Cross Talk; Logistic Regression; Qualitative Comparative Analysis; Deyun Club.

1. Introduction

In the research process of this paper, a combination of macro social background and micro data analysis is adopted, starting from the current situation of Jiangsu residents' cognition of cross talk performances, comprehensively considering the development direction and form of cross talk, and understanding residents' willingness to inherit cross talk. The research contents are as follows:

- (1). To collect basic information on the cognition of cross talk among residents in Jiangsu Province, including gender, age, occupation, income, education, whether they have seen cross talk performances, and the level of understanding of cross talk terms. Construct resident portraits from four aspects, and analyze the understanding and cognition of cross talk performances by residents in Jiangsu Province.
- (2). Collect the influencing factors that affect residents' choice of watching cross talk performances, including the cost of watching performances, recommendations from relatives and friends, popularity of actors, performance themes, publicity efforts, performance appeal, etc., and find out the key factors. The residents' trust, satisfaction and influencing factors for cross talk therapy were studied, and some existing problems of cross talk were explored.

2. Literature Review

The word "cross talk" is a combination of the "phase" of appearance and the "sound" of sound. The earliest written records on the origin of cross talk can be traced back to the Han Dynasty, as stated in "Han Shu • Xu Le Biography": "The laughter of haiku and dwarfs is no less than before." However, it was not until the Tongzhi and Xianfeng years of the Qing Dynasty that cross talk became an independent art, and it had a mature system of its own[6]. The integration of mobile communication technology and Internet technology has brought mankind into a new era of mobile Internet, and the arrival of new media marketing has injected new vitality into the art of cross talk. Under the influence of the popularization of IP in the cultural industry and

the star effect, the cross talk industry has moved up in time and focused, and the cross talk group headed by Deyun Club has entered the public eye and won the love of the audience. To study cross talk, one cannot avoid the "Cross Talk Dictionary" edited by Xue Baokun (2012), which is the first cross talk dictionary in China. A comprehensive summary of the art of cross talk has been made through more than 5,000 entries in terms of trade rules, arts and customs, apprenticeship and art, famous names, practitioners groups, and documents, and it has become an encyclopedia about cross talk art[1].

Since the new era, in terms of cross talk theory, there have been many monographs focusing on the history of cross talk and the theory of cross talk art. The latter expounds the characteristics of "cross talk art" and the development history of cross talk. Again, there are many researches on related papers, and the perspectives are relatively new. For example: Liu Nan (2009), in his paper "Analysis of Guo Degang's Cross-talk Language Characteristics", expounds Guo Degang's language characteristics from three aspects: rhetorical characteristics, pragmatic characteristics, and burden characteristics[2]. In terms of the study of cross talk as a medium, the research is relatively closely related to this topic, so the author paid special attention to it and found that the research on cross talk as a medium at this stage is mainly about the change of the communication medium and the elaboration of the communication effect[3]. For example: Xiong Hanyu (2017), in his paper "Research on the Communication of Cross Talk Art in the New Media Environment", looks at the development stage of cross talk from the perspective of journalism and communication. His thesis mainly interprets the current situation of interpersonal communication, mass communication and new media communication[5].

Jiang Huiming (2016) explored the process of localization, marketization and idolization of contemporary cross talk art. Among them, Deyun Club is a well-deserved typical representative[4]. Wang Feng (2019) believes that the current predicament of cross talk art is in creation, and mining the heritage of traditional cross talk will never be a retrospective. Wang Rubing (2020) analyzed the current situation of commercialization of traditional cross talk, and believed that the cross talk industry, while returning to tradition, used the development of new media to expand the scope of audiences and break geographical restrictions[8]. However, with the rapid expansion, it also faces unavoidable challenges such as content-blind entertainment, impetuous actors' mentality, rigid business performance forms, deformed fan groups, and chaotic cultural markets[7].

3. Model Establishment and Solution

3.1. Introduction to Logistic Regression

Logistic regression model is the most commonly used multivariate quantitative analysis method for regression analysis of binary dependent variables in sociology, psychology, demography and other fields[9]. And the weights of independent variables can be obtained through Logistic regression analysis, which prepares us to predict the possibility of events based on influencing factors. The formula is as follows:

$$P(Y_i = 1 | X_i) = \frac{e^{\alpha + \sum_{i=1}^k \beta_k x_{ki}}}{1 + e^{\alpha + \sum_{i=1}^k \beta_k x_{ki}}} \quad (1)$$

Among them, Y represents the dependent variable (0-1 variable, 0 means the event does not occur, 1 means the event occurs), and $X=[x_1, x_2, \dots, x_k]$ is a set of dependent variables corresponding to Y. $P(Y_i = 1 | X)$ represents the probability of an event occurring after a given set of values $X=[x_1, x_2, \dots, x_k]$. The optimal parameter $[\alpha, \beta_1, \beta_2, \dots, \beta_n]$ is determined by

establishing a likelihood function for a given set of dependent variables Y and a set of independent variables X corresponding to it[10].

Suppose the probability of an event occurring is

$$P(Y_i = 1 | X_i) = p_i \tag{2}$$

Then the ratio of the probability of an event occurring to the probability of not occurring is

$$\frac{p_i}{1 - p_i} = e^{\alpha + \sum_{k=1}^k \beta_k x_{ki}} \tag{3}$$

This ratio is called the occurrence ratio of events, abbreviated as odds. odds must be positive, because $0 < p_i < 1$, taking the natural logarithm of odds yields a linear function:

$$\ln\left(\frac{p_i}{1 - p_i}\right) = \alpha + \sum_{k=1}^k \beta_k x_{ki} \tag{4}$$

3.2. Data Processing

Quantify the indicators related to cross talk cognition and inheritance and development in this questionnaire, so as to use relevant software for statistical analysis and draw relevant conclusions[11]. Among them, gender and whether you have seen cross talk live performances are categorical variables, which are coded by 0-1, and age, education level, monthly income, etc. are ordinal variables, which are coded by integers with sizes[12].

3.3. Establishing Likelihood Functions

Estimate the parameters of the logistic regression model by maximum likelihood estimation. Assuming a population Y_1, Y_2, \dots, Y_n consisting of N cases, the observations are marked as y_1, y_2, \dots, y_n , and $P(y_i=1|x_i)$ is the conditional probability of obtaining the result $y_i = 1$ given the conditions of x_i ; and the conditional probability of obtaining the result $y_i = 0$ under the same conditions is $P(y_i = 0 | x_i) = 1 - p_i$. So, the probability of getting an observation is:

$$p(y_i) = p_i^{y_i} (1 - p_i)^{1 - y_i} \tag{5}$$

where $y_i = 0$ or $y_i = 1$. Because the observations are independent of each other, their joint distribution can be expressed as the product of the marginal distributions:

$$L(\theta) = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1 - y_i} \tag{6}$$

The above formula is the likelihood function of n observations. Next, find the parameter estimates that maximize the value of this likelihood function, taking the logarithm of the left and right sides of the above equation.

$$\ln[L(\theta)] = \ln\left[\prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1 - y_i}\right] = \sum_{i=1}^n [y_i(\alpha + \beta x_i) - \ln(1 + e^{\alpha + \beta x_i})] \tag{7}$$

The above formula is called the log-likelihood function. In order to estimate the values of the overall parameters α and β that can maximize $\ln[L(\theta)]$, first calculate the partial derivatives of α and β respectively, and then make them equal to 0, and the obtained parameters are the optimal parameters.

$$\frac{\partial \ln[L(\theta)]}{\partial \alpha} = \sum_{i=1}^n [y_i - \frac{e^{\alpha+\beta x_i}}{1 + e^{\alpha+\beta x_i}}] \tag{8}$$

$$\frac{\partial \ln[L(\theta)]}{\partial \beta} = \sum_{i=1}^n [y_i - \frac{e^{\alpha+\beta x_i}}{1 + e^{\alpha+\beta x_i}}] x_i \tag{9}$$

3.4. Selection Factors

In order to study the influencing factors of whether residents choose to watch cross talk live performances, this section selects 12 factors such as gender, age, and educational level as independent variables, and selects Saw a live performance (Yes/No) as the dependent variable for Logistic regression. The specific factors are selected The situation is shown in Table 1

Table 1. Independent and Dependent Variables for Logistic Regression

	Selection factors	Marks
Independent variable	Gender	M1
	Age	M2
	Education level	M3
	Income	M4
	Knowledge of cross talk	M5
	Cost of viewing	M6
	Recommended by relatives and friends	M7
	Actor popularity	M8
	Performance theme	M9
	Propaganda power	M10
	Infectant power of performance	M11
	Cross talk understanding	M12
Dependent variable	Saw a live performance (Yes/No)	S

3.5. Model Results

Table 2. Comprehensive test of model coefficients

	chi-square	degrees of freedom	salience
Step 1	43.24	12	0.000
	43.24	12	0.000
	43.24	12	0.000
Step 2	2.00	1	0.112
	36.89	5	0.001
	36.89	5	0.000

Omnibus refers to the global likelihood ratio test of the model, step statistic is the result of the likelihood ratio test at each step compared to the previous step, and block refers to the likelihood ratio test comparing the nth block to the n-1th block As a result, the model is the

result of the likelihood ratio test of the previous model and the model after the variables in the current model have changed[13]. The significance level is less than 0.05, and the model passes the test.

According to the Hosmer and Lemeshow test contingency table, comparing the observed value and the expected value, we can know that the observed value is roughly the same as the expected value, so it can be considered that the model fit is good, with 0.5 as the threshold, according to the observed value and the predicted value, it can be seen that the model The prediction accuracy can reach 68.77%.

Table 3. Analyze data

	B	S.E	Wals	df	Sig.	Exp(B)	EXP (B) 95%C.I.	
Education level (x)	0.343	0.211	5.903	1	0.002	0.343	0.200	0.876
Income (y)	0.461	0.443	2.110	1	0.000	1.587	1.333	4.018
Actor popularity (z)	0.278	0.190	3.000	1	0.005	2.009	1.002	1.332
Infecant power of performance (m)	0.521	0.542	3.004	1	0.001	0.111	0.223	0.866
Cross talk understanding (n)	0.877	0.226	6.988	1	0.026	0.200	1.098	3.887
constant	-0.192	1.633	0.233	1	0.005	0.549		

Through stepwise regression, the final fitting results are obtained. As shown in Table 3, the significance level of the five variables of education level, monthly income, actor popularity, performance appeal, and understanding of cross talk are all less than 0.05. Whether residents choose to watch The live crosstalk performance has a significant impact, and the P value of the remaining variables is greater than 0.05, which has no obvious correlation with whether residents choose to watch the live crosstalk performance. The fitted equation is:

$$\ln\left(\frac{pi}{1-pi}\right) = -0.192 + 0.343x + 0.461y + 0.278z + 0.521m + 0.877n \quad (10)$$

4. Conclusion

4.1. The Educational Level Variable has a Significant Positive Effect on Whether Residents Choose to Watch Cross Talk Live Performances

The higher the education level of a person, the wider the knowledge. Under the same financial resources, the higher the education level, the more interest in the traditional cross talk culture. When the education level is primary school and below, junior high school, high school/secondary school, the proportion of people who choose to watch live cross talk performances is roughly the same. The proportion of people who choose to watch cross talk performances is 48.8%. The survey team further analyzed this. According to the pie chart in the figure below, it can be concluded that 53.3% of the people whose education level is college/undergraduate do not choose to watch cross talk live performances[14]. The proportion is more than those who choose to watch cross talk live performances.

4.2. The Monthly Income Variable has a Significant Positive Effect on Whether Residents Choose to Watch Cross Talk Performances

A person's income level affects his quality of life. When the income is low, it is difficult for people to find suitable entertainment methods within the economic range they can afford. A higher income level means that they can pursue a higher quality of life. Focusing on mental health, the

wider the range of entertainment options, the greater the possibility of choosing to watch cross talk performances. As the income level increases, the proportion of people who choose to watch cross talk performances continues to increase, and the proportion of people who do not choose to watch cross talk performances shows a downward trend.

4.3. Actor Popularity Variable has a Significant Positive Effect on Whether Residents Choose to Watch Cross Talk Performances

The higher the influence of the actor popularity variable on the choice of watching cross talk performances, the more inclined they are to choose to watch cross talk performances; It reflects that in order to improve the acceptance of cross talk among residents, it is indispensable to enhance the popularity of actors. At present, cross talk still has a major problem, and there are still problems of slow development of cross talk teams of more types or factions.

4.4. The Understanding of Cross Talk has a Significant Positive Effect on Whether Residents Choose to Watch Cross Talk Live Performances

As people's understanding of cross talk increases, the proportion of people who choose to watch cross talk live performances is higher. When people don't know about cross talk at all (the level of understanding is 1), most of them don't choose to watch live cross talk performances. When people are very familiar with cross talk (the level of understanding is 5), the proportion of people who choose and don't choose live cross talk performances is close to 6:1, it can be inferred that the reason why some people do not choose to watch cross talk live performances is that they do not know enough about cross talk.

Acknowledgments

This work was supported by the Innovative Training Program for College Students in Anhui Province, Research on the Development Trend of Deyun Cross Talk Based on Text Mining and QCA (no. S202110378011).

References

- [1] Xue Baokun. The Big Cross Talk Dictionary [M]. Tianjin: Baihua Literature and Art Publishing House, 2012.10.
- [2] Jiang Huiming. Cross talk [M]. Beijing: China Federation of Literary and Art Circles Press, 2009.
- [3] Chen Jianhua. On Chinese Cross Talk Art [M]. Jinan: Qilu Book Club, 2013.
- [4] Liu Nan. Analysis of the language characteristics of Guo Degang's cross talk [D]. Jinan: Shandong University, 2009.
- [5] Xiong Hanyu. Research on the dissemination of cross talk art in the new media environment [D]. Nanchang: Nanchang University, 2017.
- [6] Gao Yucong. The return of traditional cross talk and the development of cross talk art [J]. Literature and Art Research, 2003(02): 100-105.
- [7] Li Lidan. "Localization" and "Presence": Intangible Cultural Heritage in the Game - Taking Tianjin Cross Talk's Inheritance Context Analysis as an Example [J]. National Art, 2010(04):31-35 +121. DOI: 10.16564/j.cnki.1003-2568.2010.04.015.
- [8] Wang Zhenyu. On the communication of cross talk art in the new media environment [J]. Modern Communication (Journal of Communication University of China), 2016,38(08):167-168.
- [9] Wu Shanghua. Research on the current situation of Deyun cross talk in the new media environment [D]. Zhengzhou University, 2018.
- [10] Han Qi. How to Create a Traditional Cultural Brand in the New Media Era: A Study on the Survival and Development of Cross Talk Performing Arts [J]. Sichuan Opera, 2018(06): 29-32.

- [11] Zhang Yun. A brief analysis of the modern dissemination of cross talk—taking Deyun Club as an example [J]. *Audiovisual*, 2019(06):211-212.DOI:10.19395/j.cnki.1674-246x.2019.06.116 .
- [12] Hong Diep Nguyen Thi,Nguyen Can Trong,Diem Phan Kieu,Hoang Nguyen Xuan,Kafy Abdulla Al. Assessment on controlling factors of urbanization possibility in a newly developing city of the Vietnamese Mekong delta using logistic regression analysis[J]. *Physics and Chemistry of the Earth, Parts A/B/C*,2021(prepublish).
- [13] Xu Xin Hong,Ye Lei,Pei Yu,Zhao Lin,Wang Jing Jing. Research on the Comprehensive Evaluation of the Higher Education System Based on FCE and ARMA Models[J]. *Complexity*,2022,2022.
- [14] Wang, Quan,Wang, Haijun. Correction to: An integrated approach of logistic-MCE-CA-Markov to predict the land use structure and their micro-spatial characteristics analysis in Wuhan metropolitan area, Central China[J]. *Environmental Science and Pollution Research*, 2022 (prepublish).