International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2019): 7.583

Customer Relationship Management (CRM) Viewed by and Analysed by Probability Distribution

Kamlesh L Patel [1], Dr. Sanjay G Raval[2], Dr. S R Patel [3]

[1]Research Scholar, School of Sciences, Department of Statistics, Gujarat University, Ahmedabad-380009, India

[2]Head, Department of Statistics, Som-Lalit College of Commerce, Navrangpura Ahmedabad-380009, India

[3]Retired professor, Department of Statistics, Sardar Patel University, Vallabh Vidyanagar-388120, India

Abstract: ***In the modern age of competition banking sector forces on bank to deal with their customers more professionally, therefore most of the banks use Customer Relationship Management (CRM) practices. Bank maximizes the customer satisfaction level by trying to understand the behaviours of the customer. In this paper customer's behaviour can be understood by the associations of attributes like safety, stability, services and bank charges. We assume frailty multinomial model for associations. We simulated many observations for this model by Monte Carlo (MC) method with the help of R language. Further correlation, multiple correlations, Partial correlation between associations, mle of associations can be obtained. Also a particular case of stability and safety has been studied by binomial generation.***

Keywords: Customer Relationship Management (CRM), Multinomial distribution, Monte Carlo Simulation, Maximum Likelihood Estimate (mle)

1. Introduction

Fair size of literature is available in practice on Customer Relationship Management (CRM). It is studied by many authors and researchers such as (Peppers, 1997) have studied the relationships between customer and banks, (Petrusihn, 2000) have studied shopper model and simulation in his paper, (Wang, 2004) dealt with Framework for Customer Value and Customer Relationship Management, (Rootman, 2008) described variables influencing the CRM of bank. All the above have studied CRM with reference to banking system and (Mamoun, 2011) have used CRM practices as a business performance in a Development Country, (Zeynep, 2012) dealt the effect of CRM in business market in their paper, (Jha, 2013) studied CRM for AXIS bank, (Patil, 2014) defined management practices and marketing in modern age and (Patel, 2017) have develop customer preferential study of banks. All of them have used modern technique for studying CRM for various banks.

In this paper we further study the Customer Relationship Management (CRM) by using various factors like safety, stability, bank-services and bank-charges. We generate sample by simulation of assumed theoretical distribution like multinomial distribution for studying and verifying theoretically the usefulness of the different factors mentioned above and for that purpose section 2 is devoted for computation of partial and multiple correlation coefficients and testing them, section 3 deals with the simulation of frailty multinomial distribution used and estimation of parameters have been done by Monte Carlo (MC) method with the help of R language, section 4 binomial simulation is used to estimate number of respondents for association when its probability is given.

1) Partial and Multiple Correlation

There are many factors to measure satisfaction level of customer in CRM practices done by banks, among them trust is an important factor. Perception of customer about trust can be measured by sub- factors such as safety, stability of bank, services and bank charges. We consider some associations of sub-factors when customers give their perception about the trust. The formation of associations is done by selecting safety as a prime sub-factor and some other sub-factors are associate to safety and the following table 2.1shows the different associations obtained from respondents.

Table 2.1

Ways of answering(AID)	Sub factors(association)
1	Safety, stability
2	Safety, charges
3	Safety, stability, services
4	Safety, stability, services, charges
5	services
6	safety
7	charges
8	stability

With respect to the above associations we form an initial matrix in which 1 indicates that customer is selecting corresponding sub factor to measure the perception about trust and 0 represent that customer is not selecting corresponding sub factor to measure the perception about trust. The following table 2.2 represent above associations in binary form.

Table 2.2

(AID)	safety	stability	services	charges
1	1	1	0	0
2	1	0	0	1
3	1	1	1	0
4	1	1	1	1
5	0	0	1	0
6	1	0	0	0
7	0	0	0	1
8	0	1	0	0
probability	0.625	0.5	0.375	0.375

From the initial matrix (table 2.2) we consider only four associations and find their probabilities.

Let A represents association, (safety→ stability)
B represents association, (safety→ services)
C represents association, (safety→ stability, services)
D represents association, (safety→ stability, services and charges)

The probabilities of different association can be done by using the following formula.

$$\text{Prob. (A)} = \frac{Number\ \ of\ time\ \ (safety \rightarrow stability\ )association\ \ of\ answer\ \ occurs}{Total\ number\ \ of\ associations\ \ of\ answers}$$
$$= 3/8$$
$$= 0.375$$

Similarly we have obtained probabilities of other associations and represented in the following table 2.3

Table 2.3

Association No	Association	Probability
1	A	0.375
2	B	0.25
3	C	0.25
4	D	0.125

With the help of table 2.2 and table 2.3 we find correlation coefficient between the associations (Xinog, 2010) by following formula

$$r_{(\text{safety, stability})} = \frac{p(safety\ \ and\ \ stability\ ) - p(safety\ )\,p(stability\ )}{\sqrt{p(safety\ )\,p(stability\ )}\sqrt{(1-p(safety\ )\,(1-p(stability\ )}}$$
$$= \frac{0.375-(o.625)(0.5)}{\sqrt{(0.625)(0.5)}\ \sqrt{(0.375)(0.5)}}$$
$$= 0.2582$$

We represent various correlation coefficients in the following table 2.4 showing correlation matrix

Table 2.4

	Safety	Stability	Services	Charges
Safety	1			
Stability	0.2582	1		
Services	0.0666	0.2582	1	
Charges	0.0666	-0.2582	-0.0666	1

Among n respondents, let n_1 respondents select association A with probability p_1, n_2 respondents select association B with probability p_2, n_3 respondents select C with probability p_3 and n_4 respondents select association D with probability p_4 to measure the perception about trust on banks. The categorical variables n_1,n_2,n_3,n_4 follow multinomial distribution with probability mass function:

$$P\ (n,\ p_1, p_2\ , p_3, p_4) = \frac{n!}{n_1!\ n_2!\ n_3!\ n_4!}\ \ p_1^{n_1}\ p_2^{n_2}\ p_3^{n_3} p_4^{n_4} \quad (2.1)$$

where $\sum_{i=1}^{4} p_i = 1,\ \ 0 \le p_i\ \le 1$ and $\sum_{i=1}^{4} n_i = n$

For multinomial distribution P $(n,p_1,p_2,\ldots p_m,\ldots,p_k)$ the partial correlation coefficient between n_1 and n_2 when variables $n_3,\ldots\ldots,n_m$ (m < k) are held fixed , $r_{12\cdot3,\ldots,m}$, is given by (Kshirsagar, 1972).

$$r_{12\cdot3,\ldots,m} = \frac{\sqrt{p_1p_2}}{\sqrt{(1-p_1-p_3-\cdots-p_m)(1-p_2-p_3-\cdots-p_m)}}\ ,\ m < k \quad (2.2)$$

Using (2.2) different partial correlation coefficients for multinomial distribution (n, 0.375, 0.25, 0.25, 0.125) are:

$$r_{12.3} = \frac{\sqrt{p_1p_2}}{\sqrt{(1-p_1-p_3)(1-p_2-p_3)}} = 0.7071 \quad (2.3)$$
$$r_{23.1} = \frac{\sqrt{p_2p_3}}{\sqrt{(1-p_2-p_1)(1-p_3-p_1)}} = 0.6666 \quad (2.4)$$
$$r_{13.2} = \frac{\sqrt{p_1p_3}}{\sqrt{(1-p_1-p_2)(1-p_3-p_2)}} = 0.7071 \quad (2.5)$$
$$r_{13.4} = \frac{\sqrt{p_1p_3}}{\sqrt{(1-p_1-p_4)(1-p_3-p_4)}} = 0.5477 \quad (2.6)$$

The multiple correlation coefficient $\rho^2_{1(23\ldots m)}$ of n_1 on $n_2,\ldots,n_m$ is obtained for multinomial distribution $P(n,p_1,p_2,\ldots p_m,\ldots,p_k)$ as follows (Kshirsagar, 1972).

$$R^2_{1(23\ldots m)} = \frac{p_1(p_2+p_3+\ldots\ldots p_m)}{(1-p_1)(1-p_2-p_3+\cdots p_m)}\ \ , m < k \quad (2.7)$$

Using (2.7) different multiple correlation coefficients for multinomial distribution (n, 0.375, 0.25, 0.25, 0.125) are:

$$R^2_{1(23)} = \frac{p_1(p_2+p_3)}{(1-p_1)(1-p_2-p_3)} = 0.6000 \text{ i.e } R_{1(23)} = 0.7746 \quad (2.8)$$
$$R^2_{2(13)} = \frac{p_2(p_1+p_3)}{(1-p_2)(1-p_1-p_3)} = 0.5555 \text{ i.e } R_{2(13)} = 0.7453 \quad (2.9)$$
$$R^2_{3(12)} = \frac{p_3(p_1+p_2)}{(1-p_3)(1-p_1-p_2)} = 0.5555 \text{ i.e } R_{3(12)} = 0.7453 \quad (2.10)$$

2) Test for partial correlation coefficients

(i) We use the result for testing H: $\rho_{12.34\ldots m} = \rho_0$ vs K: $\rho_{12.34\ldots m} \neq \rho_0$ at level α for a random sample from multinomial distribution P $(n,p_1,p_2,\ldots p_m,\ldots,p_k)$ is to reject H whenever

$$t_{n-(p-q)} = \sqrt{n-(p-q)}\ \frac{|r_{12.34\ldots m}-\rho_0|}{\sqrt{1-r^2_{12.34\ldots m}}} \ge t_{n-(p-q)}(\tfrac{\alpha}{2}) \quad (2.11)$$

Where n is the number of observation in a sample
p is total number of variables
q is number of variables held fixed.
$t_{n-(p-q)}(\frac{\alpha}{2})$ is upper $\frac{\alpha}{2}$% value of Student's t distribution on n-(p-q) degree of freedom.

So for testing H: $\rho_{12.3} = 0.72$ say vs K: $\rho_{12.3} \neq 0.72$ at 5% level, for 100 respondents the partial correlation coefficient between A and B when C is held fixed is $r_{12.3} = 0.7071$ and $t_{100-(3-1)} = 0.18578 \le 1.98$

The upper 0.025% value of t distribution at 98 d.f which shows that the null hypothesis is not rejected at 5% level.

(ii) Also for testing H:$\rho_{23.1} = 0.64$ vs K:$\rho_{23.1} \neq 0.64$, the calculated value t statistic is 0.3532 which is lesser than 1.98 at 5% level and hence the partial correlation coefficient between association B and C when A is fixed is not rejected.

3) Test for multiple correlation coefficient:

We want to test the effect of association A on B and C, that is to test H: $\rho_{1(23)}=0.8$ vs K: $\rho_{1(23)} \neq 0.8$ we reject H whenever

$$F_{(p-1,n-p)} = \frac{|R^2_{1(23)}-\rho^2_{1(23)}|}{1-R^2_{1(23)}} \times \frac{n-p}{p-1} \ge F_{(p-1,n-p)}(\alpha) \quad (2.12)$$

Where $F_{(p-1,n-p)}(\alpha)$ is the α% value of F on (p-1, n-p) degree of freedom.

For the sample of 50 the multiple correlation coefficients of A on B and C is
$R_{1(23)} = 0.7746$ and
$F_{(2,47)} = 2.35 \le F_{(2,47)}\ (0.05) = 3.18$

The table value at 5% level and on (2, 47) degree of freedom is $F_{(2,\ 47)}(0.05) = 3.18$. which shows that hypothesis is not rejected that is effect of association B and C on A is more than 77.46%.

4) Estimation using simulation

In section 2 we defined multinomial distribution with parameters (n, p_1, p_2, p_3, p_4) which is singular distribution hence we cannot use this singular distribution for estimation , we use following non singular multinomial distribution:

$$P = \frac{n!}{n_1!\,n_2!\,n_3!\,n_4!}\; p_1^{n_1}\, p_2^{n_2}\, p_3^{n_3}\,(1-p_1-p_2-p_3)^{(n-n_1-n_2-n_3)} \qquad (3.1)$$

Where $\sum_{i=1}^{4} p_i \leq 1$, $\sum_{i=1}^{4} n_i \leq$ n and $0 \leq p_i \leq 1$, i=1,2,3,4

Let association A be one and half time probable than association C also association B be equally probable to association C and association D be half time probable than association C. Thus $p_1= 1.5\, p_3$, $p_2= p_3$, and $p_4= 0.5p_3$. For generating data from the above multinomial distribution, let $0.55 \leq r \leq 0.89$ be an arbitrary number and taking $\sum_{i=1}^{3} p_i \leq$ r and using the above relation p_3=r/3.5, p_1, p_2 and p_4 can be generated for different values of r using R-Programming, table 3.1 shows the mle obtained by the use of Monte Carlo method.

Table 3.1

n	$\widehat{p_1}$	$\widehat{p_2}$	$\widehat{p_3}$	$\widehat{p_4}$
100	0.3307	0.2205	0.2205	0.2283
200	0.3284	0.2189	0.2189	0.2338
300	0.3288	0.2192	0.2192	0.2328
400	0.3322	0.2215	0.2215	0.2248
500	0.3307	0.2205	0.2205	0.2283
600	0.3280	0.2187	0.2187	0.2346
700	0.3285	0.2190	0.2190	0.2335
800	0.3290	0.2194	0.2194	0.2322
900	0.3299	0.2199	0.2199	0.2303
1000	0.3201	0.2134	0.2134	0.2531
2000	0.3186	0.2124	0.2124	0.2566
3000	0.3198	0.2132	0.2132	0.2538
4000	0.3195	0.2130	0.2130	0.2545
5000	0.3184	0.2122	0.2122	0.2572
6000	0.3189	0.2126	0.2126	0.2559
7000	0.3192	0.2128	0.2128	0.2552
8000	0.3191	0.2127	0.2127	0.2555
9000	0.3190	0.2127	0.2127	0.2556
10000	0.3191	0.2127	0.2127	0.2555

The following table 3.2 shows the mle of ($\widehat{n_1}$, $\widehat{n_2}$,$\widehat{n_3}$, $\widehat{n_4}$) by the use of probabilities obtained in table 3.1.

Table 3.2

n	$\widehat{n_1}$	$\widehat{n_2}$	$\widehat{n_3}$	$\widehat{n_4}$
100	33	22	22	23
200	66	44	44	47
300	99	66	66	70
400	133	89	89	90
500	165	110	110	114
600	197	131	131	141
700	230	153	153	163
800	263	176	176	207
900	297	198	198	207
1000	320	213	213	253
2000	637	425	425	513
3000	959	640	640	761
4000	1278	852	852	1018
5000	1592	1061	1061	1286
6000	1913	1276	1276	1535
7000	2234	1490	1490	1786
8000	2553	1702	1702	2044
9000	2871	1914	1914	2300
10000	3191	2127	2127	2555

5) Real Data

We have collected data of 530 respondents as bank customers and among 530 respondents, 151customers have selected association A,101 customers have selected association B ,101 customers have selected association C and 177 customers have selected association D. Thus

No of respondents n = 530
No of respondents selecting association A = 151 that is n_1 =151
No of respondents selecting association B = 101 that is n_2 =101
No of respondents selecting association C =101 that is n_3 = 101
No of respondents selecting association D =177 that is n_4 =177

6) Simulated data:

By Monte Carlo Simulation technique using underline distribution as a Frailty multinomial distribution and taking total number of respondents as 530 the simulated values of four variables are estimated as under
No of respondents selecting association A = 154 that is $\widehat{n_1}$ = 154
No of respondents selecting association B = 102 that is $\widehat{n_2}$ = 102
No of respondents selecting association C = 102 that is $\widehat{n_3}$ =102
No of respondents selecting association D = 172 that is $\widehat{n_4}$ = 172

Thus for the collected real data of 530 respondents of various banks has been compared with data generated by frailty multinomial distribution in a table 3.3

Table 3.3

Categorical variables	Collected Data	Simulated Data
n_1	151	154
n_2	101	102
n_3	101	102
n_4	177	172
Total	530	530

Table 3.3 shows that the real data and generated data using frailty multinomial distribution are very close to each other. So the assumed frailty multinomial distribution is appropriate.

7) Special Case (Binomial Distribution)

Dorman at el (1968) have estimated the parameter n in the Binomial distribution when the value of the parameter p is given. They have obtained asymptotic estimate of n by using maximum likelihood and minimum variance method. Here we generate the samples and use Monte Carlo method to get mle of n for given p in Binomial (n, p) distribution. In our case we obtain sample size, estimates of the association A (safety → stability). Using technique of section 3 we obtain the sample size of association and give it in the following table 5.1

For generating the data based on Binomial distribution let probability of success (occurrence of association A) be 1.5 time probability of failure.

Let $0.7 \leq r \leq 0.9$ be an arbitrary number for generating data set by using $p \leq \frac{r}{2.5}$ in Binomial distribution. For given p (given r) the number of trials n can be generated by Monte Carlo simulation and mles are obtained for different values of p, estimated value of n are given in the following table 5.1

Table 5.1

p	0.6532	0.6526	0.6506	0.6524	0.6514	0.6512	06537	0.6521	0.6504	0.6508
$\hat{n}$	196	196	195	196	195	195	196	196	195	195

2. Conclusion

The main focus of this paper was on modelling the frailty multinomial distribution model. Correlation coefficients, partial correlation coefficient and multiple correlation coefficients between associations have been obtained. Further the tests for multiple and partial correlations have been carried out. We have estimated the probabilities as well as mles of number of occurrence of associations. Further for given probability, we have estimated the number of trails of Binomial distribution as a special case of multinomial distribution. Although significant improvements are possible, remaining factors associated with CRM may be studied by same manner.

References

[1] Cardez, & el, a. (2001). Probabilistic Modeling of Transaction Data with application to Profiling Visualization and Prediction. *KDD 2001 conference*, (pp. 37-46).

[2] Feldman, D., & Fox, M. (1968). Estimation of Parameter n in Binomial Distribution. *The American Statistical Association , 63* (321), 150-158.

[3] Jha, N., & Bhome, S. M. (2013). "A Study of CRM-As a strategic Tool Adopted by Banks With Reference to AXIS Bank in Thane City". *Indian Journal of Applied Research , 3* (5), 454-457.

[4] Kshirsagar. (1972). *Multivariate Analysis.* Marcel Dekker INC.

[5] Mamoun, N., & at.el. (2011). Customer Relationship Management Implimentation: An Investigation of a Scale's Generalizabitity and its Relationship with Business Performance in a Developing Country Context". *International Journal of Commerce and Management , 21* (2), 158-190.

[6] Patel, R. C., & Trivedi, C. J. (2017). "CRM Customer Preferential Study on Financial Services For Public and Private Sector Banks With Reference to Ahmadabad District". *International Journal of Engineering and Technology , 4* (6), 386-396.

[7] Patil, D. B., & Bhakkad, D. D. (2014). *"Redefining Management Practices and Marketing in Modern Age".* Athrar Publications.

[8] Peppers, & Rogers. (1997). "The one to one Future: Building Relationships one Cuatomer at a Time". Double Day Publications.

[9] Petrusihn, V., & at.el. (2000). "e-shoper Modeling and Simulation". *SPIE conference on Data Mining.*

[10] Rootman, C., Tait, M., & Bosch, J. (2008). Variables infiuencing the customer relationship management of bank. *Journal of Financial Services Marketing , 13* (1), 52-62.

[11] Wang, Y., Lo, H. P., Chi, R., & Yang, Y. (2004). "An Integrated Framwark for Customer Value and Customer Relationship Management Performance: A Customer base Perspective from China". *Managing Services Quality* , 169-182.

[12] Xinog, H., & Shahshi, S. (2010). A two steps approach for all strong pairs correlation query in large Database. *Transactions on Knoledge and Data Engineering , x.*

[13] Zeynep, A. U., & Toker, A. (2012). "The effect of Customer Relationship Management Adoption in Business to Business Markets". *Journal of Business and Industrial Marketing , 27* (6), 497-507.