Chapter 8
Insurance and Banking Habits Regulating Economic Stability of the People of Char Areas (Riverine) in Lower Assam

Kishore K. Das

Abstract Riverine (Char area) of the Brahmaputra flowing through Assam are most backward and neglected areas. There about 2251 Chars covering about 10% population of Assam. This chapter tries to examine the Insurance and Banking habits regulating economic stability of the people of char areas in Lower Assam. This study is based on the survey '*Dynamics of socio-economic development of Char areas of Lower Assam*' was conducted during the year 2003–2005 under the financial support of University Grants Commission, New Delhi on randomly selected char areas. Insurance and banking habits have been considered in the survey to study the economic stability of the people. A log-linear model and contingency table have been considered in this study. It is observed that the proportion of insured households in the low income group is nil; it rises to a small in the lower middle group and then goes up in the upper middle group and after that steadily increases to the high income group. The economic status reveals that only the upper middle income and high income group households have been able to avail the facility of banking and insurance. The households in the low income group and lower middle income group went unrepresented.

Keywords Insurance · Banking · Log-linear model

8.1 Introduction

A char is nothing but a landmass surrounded by flowing river water inhabited by people and animals. There are 2251 char areas in 14 districts out of 23 districts of Assam (Socio-economic Survey Report 2002–2003). Among these char areas, a large numbers are densely populated. These areas constitute those lands, which go under water, except a very few chars, during monsoon flow of the Brahmaputra River and remain dry thereafter. The majority inhabitants of the Chars are Muslim

K.K. Das (✉)
Department of Statistics, Gauhati University, 781 014 Guwahati, Assam, India
e-mail: daskkishore@gmail.com

U.K. De et al. (eds.), *Inequality, Poverty and Development in India*, India Studies in Business and Economics, DOI 10.1007/978-981-10-6274-2_8

(98.4%). The people living in these areas are very poor and they have to brave floods and erosion of their lands every year. The risks of health hazard are also high. The socio-economic conditions of the Chars are appalling.

Keeping in view these basic backgrounds of the people in char areas, the project was completed with the title Economic Development of Char Areas in Assam: Phase-I., which is a humble endeavour in a process to measure the level of development and study the socioeconomic status of the char areas and put forward necessary recommendations. The project was initiated under the banner of the Department of Statistics, Gauhati University, Guwahati, undertook a survey in the Char areas of Assam precisely, in the districts of Barpeta, Dhubri and Goalpara. This survey was sponsored by the University Grants Commission, New Delhi. These chars although not big in size, their socio-economic development plays a major role in the overall development of the state.

8.2 Data and Methodology

8.2.1 Sample Design

The sample design adopted for the survey is a multi-stage stratified sampling and systematic sampling. The main fieldwork for the survey in the three districts was carried out by the investigators. It was carried out in the year 2003–2004. There are 2251 char areas in 14 districts out of 23 districts of Assam. We have considered only three districts, which have been chosen using random number generated from Casio calculator by the Rnd function. The districts Barpeta, Dhubri and Goalpara were found for the study as first stage. In the second stage, the Chars of these districts were selected adopting the probability proportional to size (PPS) procedure. Households of the char areas were chosen by systematic sampling as third stage.

Here we investigated the economic planning of the households, like whether they are members of cooperative societies, have any bank account, insurance policy, etc. We have also asked the households if any member of that household has/had ever taken a loan from cooperative society or bank or against insurance policy.

From the survey, we have come to know that none of the sampled households is a member of any cooperative society. This speaks volumes of the fact that the culture of cooperative societies has not yet touched upon these backward areas of Assam. The authorities seriously need to do something about this.

As far as savings in private depository institutions are concerned, only negligible (0.3%) households have invested in these institutions.

A log-linear model and contingency table have been considered to find the study. It is observed that the proportion of insured households in the low income group is nil; it rises to a small in the lower middle group and then goes up in the upper middle group and after that steadily increases to the high income group. The economic status reveals that only the upper middle income and high income group households have been able to avail the facility of banking and insurance. The households in the low income group and lower middle income group went unrepresented.

8.3 Banking in the Char Areas of Lower Assam

The information on whether a household has any bank account throws some light on the economic condition of that particular household. Obviously those households, which earn their livelihood on a day-to-day basis, cannot imagine of having an account in bank. This is why we put a question on this aspect of economic planning. And what we have found matches with our expectation. Just 8.2% of the sampled households have any bank account. In other words, 91.8% of the sampled households cannot think of savings in a bank. So it can be said that, in general, the people residing in the char areas of lower Assam earn their livelihood on a day-to-day basis Fig. 8.1.

The picture is even dimmer in case of Barpeta district where only 5.3% of the households have bank account. The proportion of households having bank account in the other two districts of Dhubri and Goalpara are nearly equal (10.2 and 10.3% respectively) Fig. 8.2.

Almost half (47.3%) of the households that have any bank account, deposit their money in Pragjyotish Gaonlia Bank, a regional rural bank (RRB). The other two big

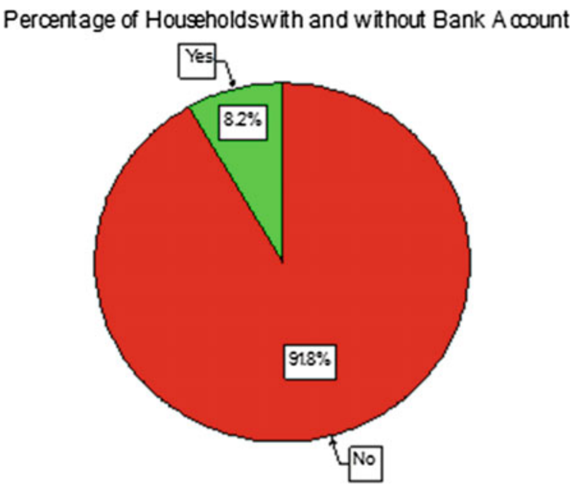


Fig. 8.1 Pie chart of percentage of households with and without bank accounts

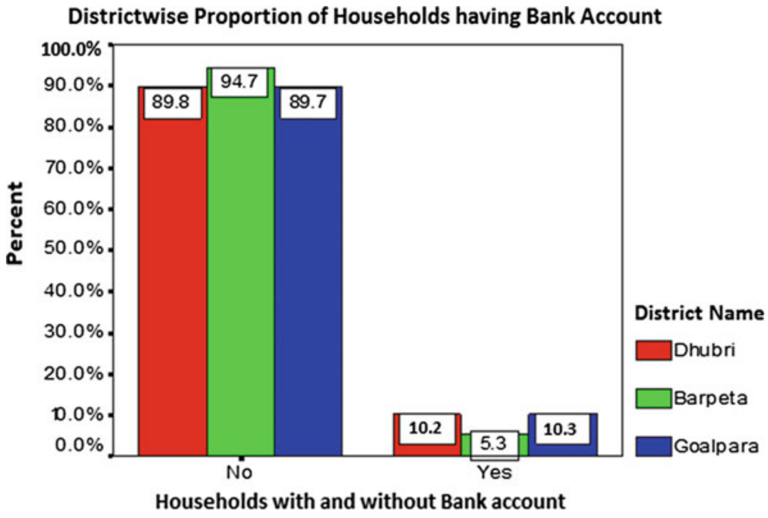


Fig. 8.2 Clustered bar chart showing district wise proportion of households having bank account

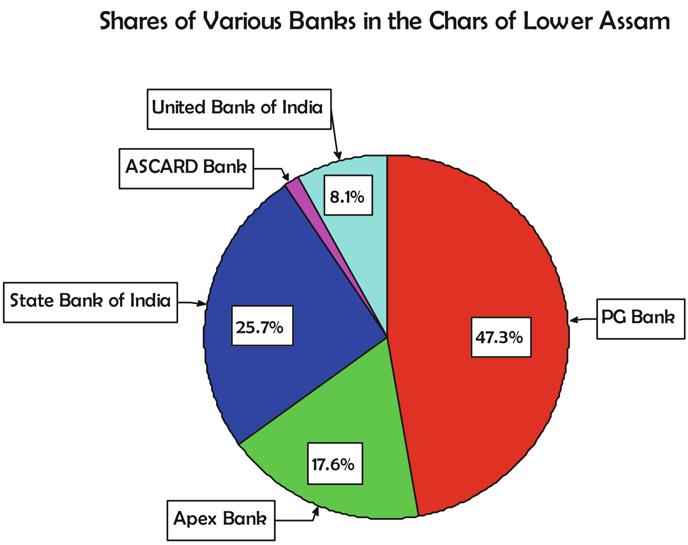


Fig. 8.3 Pie chart of various banks' shares in the char households

catchers of the money of these households are State Bank of India (SBI) (25.7%) and The Assam Co-operative Apex Bank Limited (17.6%). The remaining 9.5% of the households are shared by United Bank of India (UBI) (8.1%) and ASCARD Bank (1.4%) Fig. 8.3.

8.4 Insurance Sector in the Char Areas of Lower Assam

In case of insurance policy vis-à-vis the char households, almost a similar picture emerges. To be precise, only 5% of the sampled households have any insurance cover. The remaining 95% households do not have any security for the future. Dhubri district takes the lead with 9.1% insured households, followed by Barpeta (4%) and Goalpara (2.6%) Figs. 8.4 and 8.5.

Of all the households that are insured, 93.3% have got life insurance policies. One positive sign that have surfaced from this enquiry is that the people residing in the char areas have started insuring their crops, though in a very small proportion (only 2.2%). Crop insurance with respect to the char areas is very important when we take into account the fact that their crops are extremely vulnerable to natural calamities like flood.

8.5 Relation Between Banking and Insurance with Respect to Char Households

We have here made an attempt to examine if there exists a significant relationship between a char household having a bank account and it having an insurance policy. We construct a two-way contingency table on the nominal categorical variables—households with bank account and households with insurance policy; each with two categories—yes and no.

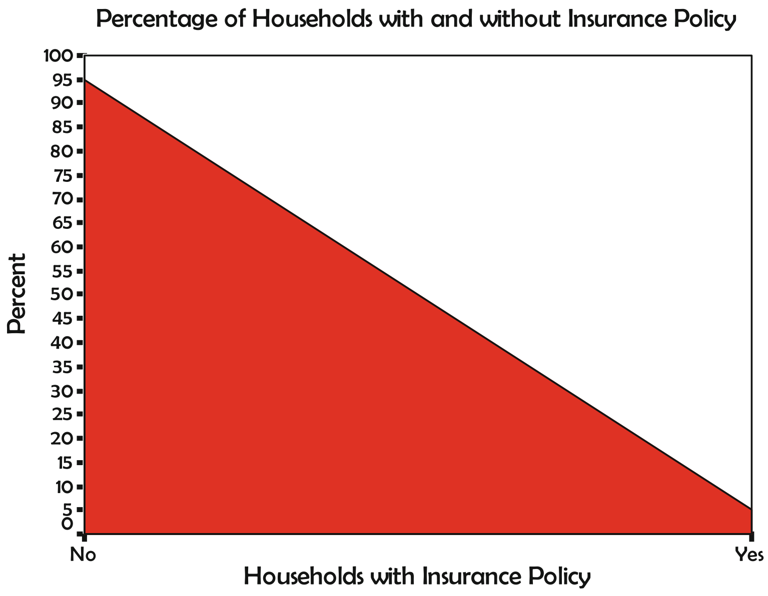


Fig. 8.4 Area diagram showing distribution of households with regard to insurance policy

We now proceed to perform the chi-quare Test of independence of attributes. The chi-square test compares the number of cases falling into each cell of the table with the frequency that would be expected if there were no association between the two variables that form the Table 8.2.

The contingency Table 8.1 demonstrates that 90.2% of the sampled households have neither a bank account nor an insurance policy and another 3.4% households have both a bank account and an insurance policy. Only a small 6.4% households have either a bank account or an insurance policy.

The exact p-value (<0.0005) under Fisher's Exact Test is significant at 5% level of significance, as it is less than our chosen significance level 0.05. So we reject the null hypothesis of independence between banking and insurance vis-à-vis char households in the population and conclude that there is a significant association between the two categorical variables.

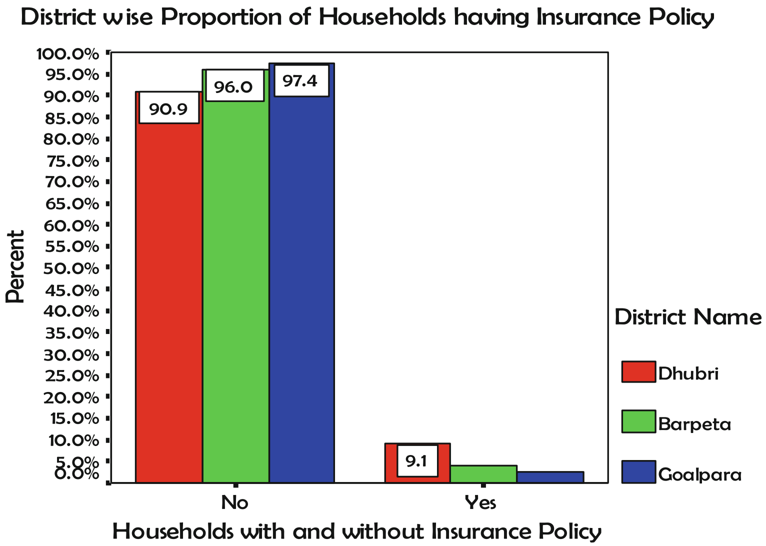


Fig. 8.5 Clustered bar chart showing district wise proportion of households having insurance policy

Table 8.1 Contingency table showing households with/without bank account and households with/without insurance policy

Banking * insured cross-tabulation					
Bank account	Yes	Count	31	43	74
		% to total	3.4	4.8	8.2
	No	Count	14	812	826
		% to total	1.6	90.2	91.8
Over all		Count	45	855	900
		% to total	5.0	95	100

Source Primary survey conducted during 2003–2004

Table 8.2 Chi-square test of independence between banking and insurance

Chi-square tests				
	Value	df	Exact sig. (2-sided)	Exact sig. (1-sided)
Pearson chi-square	231.027[b]	1	0.000	0.000
Continuity correction[a]	222.642	1		
Likelihood ratio	114.764	1	0.000	0.000
Fisher's exact test			0.000	0.000

Note [a]Computed only for a 2 × 2 table
[b]1 cells (25.0%) have expected count less than 5. The minimum expected count is 3.70
Source Primary survey conducted during 2003–2004

Table 8.3 Contingency table of economic class of households and bank account

Economic class of a household * bank account cross-tabulation					
			Bank account		Total
			No	Yes	
Economic class of a household	Low economic class	Count	179		179
		% within bank account	22.2		20.3
	Lower middle economic class	Count	258	14	272
		% within bank account	31.9	19.2	30.9
	Upper middle economic class	Count	249	17	266
		% within bank account	30.8	23.3	30.2
	High economic class	Count	122	42	164
		% within bank account	15.1	57.5	18.6
Total		Count	808	73	881
		% within bank account	100.0	100.0	100.0

Source Primary survey conducted during 2003–2004

8.5.1 Odds Ratio and Relative Risk for the "Insurance" × "Banking" Cross-tabulation

The sample odds ratio is defined as in Eq. (8.4.1).

$$\hat{\theta} = \frac{n_{11}n_{22}}{n_{12}n_{21}} \tag{8.4.1}$$

where $\hat{\theta}$ represents the maximum likelihood estimate (MLE) of the true odds ratio, and n is the count obtained in each cell.

For our data, $\hat{\theta} = \frac{31 \times 812}{14 \times 43} = 41.81395$

Risk estimate			
	Value	95% confidence interval	
		Lower	Upper
Odds ratio for insured (Yes/No)	41.814	20.729	84.346
For cohort banking = Yes	13.698	9.640	19.463
For cohort banking = No	0.328	0.212	0.506

Source Primary survey conducted during 2003–2004

The relative risk estimate is a measure of association between the presence or absence of a factor and the occurrence of an event. Here we take the factor to be ‘Banking’, i.e. whether a household has a bank account or not, and having or not having an insurance policy as the event. For the contingency Table 8.1, the relative risk of being insured is nearly 42 (13.698/0.328 = 41.814) times among households which have a bank account than among the households which do not have a bank account. And the 95% confidence interval for the relative risk (which is equal to the *odds ratio* in this case as the probability of the outcome of interest is very small) does not include 1, indicating that there is a significant difference in the occurrence of insured households between households which have a bank account and households which do not have a bank account.

8.5.2 Measures of Association Between Banking and Insurance

Since the two nominal categorical variables-households with/without bank account and households with/without insurance policy, are not independent in the population, they are associated. We now give below some indices that measure the degree of association between these two variables.

Symmetric measures				
		Value	Approx. sig.	Exact sig.
Nominal by nominal	Phi	0.507	0.000	0.000
	Cramer’s V	0.507	0.000	0.000
	Contingency coefficient	0.452	0.000	0.000

Source Primary survey conducted during 2003–2004

The nominal symmetric measures under the above-referred chi-square test suggest that though the relationship between a household having a bank account and its having an insurance policy is significant, the strength of this relationship is not so strong. The relationship is moderate because value of each of the coefficients revolves around 0.5.

Directional measures				
			Value	Exact sig.
Nominal by nominal	Goodman and Kruskal tau	Bank account dependent	0.257	0.000
		Insurance policy dependent	0.257	0.000
	Uncertainty coefficient	Symmetric	0.264	0.000
		Bank account dependent	0.224	0.000
		Insurance policy dependent	0.321	0.000

Source Primary survey conducted during 2003–2004

The nominal directional measures are appropriate when both variables are nominal, categorical variables. The nominal directional measures indicate both the strength and significance of the relationship between the row and column variables of a contingency table. These coefficients are measures of association between categorical variables based on the idea of using one variable to predict the other. The value of each statistic can range from 0 to 1 and indicates the proportional reduction in error in predicting the value of one variable based on the value of the other variable.

In this case, the low significance values for both *Goodman and Kruskal tau* and the *uncertainty coefficient* indicate that there is a relationship between the two categorical variables. But the temperate values for both the test statistics indicate that the relationship between the two variables is a moderate one. Nevertheless, the test statistic value of 0.321 of *uncertainty coefficient* indicates that by trying to predict whether a char household will have an insurance policy when the household having or not having a bank account is known, one can reduce the error rate by 32.1% over what he/she could expect by random chance.

After all these analyses, we come to the conclusion that the households having a bank account are more likely to have an insurance policy than the households which do not have a bank account.

8.6 Relation Between Economic Class of a Household and its Having Bank Accounts

Common experience tells us that the more affluent a household is the more is the chance of its having a bank account. We would like to see whether this perception applies in case of char areas of lower Assam too. So we did a chi-square test of independence on these two categorical variables (Agresti 1984, 1990) Table 8.4.

The result was as one would expect. The test on the null hypothesis that there is no relationship between the two variables tabulated, produced a significant *p*-value at 5% level of significance. So we reject the null hypothesis of no association at 5% alpha-level and conclude that these two factors-economic statuses and having a bank account are associated in the population of char areas of lower Assam.

The 4 × 2 contingency Table 8.3, having four categories of economic class of a household (low, lower middle, upper middle and high) and two categories of bank account (yes or no), also illustrates this. From the table, we can see that in the households having a bank account, the proportion of low economic class is nil. The proportions of lower middle economic class and upper middle economic class in the households having a bank account are 19.2 and 23.3% respectively; the remaining 57.5% households fall in the high economic class.

Thus we are now in a position to infer that the households falling in the upper economic class are more likely to have a bank account than their counterparts in the lower economic class.

8.7 Relation Between Income Level of a Household and its Having an Insurance Policy

Let us now test the null hypothesis that income level of a household does not influence whether it will have an insurance policy in the chars of lower Assam. For testing this hypothesis, we have constructed the following 4 × 2 contingency table and then performed a chi-square test of independence of attributes on the table, as shown in Table 8.6.

Table 8.4 Chi-square test of independence of economic class and having bank account for the households in the chars of lower Assam

Chi-square tests				
	Value	df	Asymp. sig. (2-sided)	Exact sig. (2-sided)
Pearson chi-square	85.721[a]	3	0.000	0.000
Likelihood ratio	80.059	3	0.000	0.000
Fisher's exact test	75.872			0.000

Note [a]0 cells (0.0%) have expected count less than 5. The minimum expected count is 13.59
Source Primary survey conducted during 2003–2004

Table 8.5 Contingency table of income level of households and insurance policy

Income level of a household * insurance policy cross-tabulation					
			Insurance policy		Total
			No	Yes	
Income level of a household	Low income group	Count	246		246
		% within insurance policy	29.4		27.9
	Lower middle income group	Count	204	1	205
		% within insurance policy	24.4	2.2	23.3
	Upper middle income group	Count	206	4	210
		% within insurance policy	24.6	8.9	23.8
	High income group	Count	180	40	220
		% within insurance policy	21.5	88.9	25.0
Total		Count	836	45	881
		% within Insurance policy	100.0	100.0	100.0

Source Primary survey conducted during 2003–2004

Table 8.6 Chi-square test of independence between income level of households and their having an insurance policy

Chi-square tests				
	Value	df	Asymp. sig. (2-sided)	Exact sig. (2-sided)
Pearson chi-square	104.298[a]	3	0.000	0.000
Likelihood ratio	94.485	3	0.000	0.000
Fisher's exact test	88.240			0.000

Note [a]0 cells (.0%) have expected count less than 5. The minimum expected count is 10.47
Source Primary survey conducted during 2003–2004

The chi-square test statistic with $(4-1) \times (2-1) = 3$ degrees of freedom is significant at 5% alpha-level; *p*-value (<0.0005) being less than the chosen level of significance (0.05). So the two categorical variables under consideration are not independent in the population of char areas of lower Assam. That is to say, these two variables are associated at 5% risk level. An inspection of the 4×2 contingency in Table 8.5 shows that the proportion of insured households in the low income group is nil; it rises to a small 2.2% in the lower middle group and then goes up to 8.9% in the upper middle group and after that steadily increases to 88.9% in the high income group. To summarise, the chance of possessing an insurance policy for a household falling in the upper income group is more than that of a household falling in the middle or lower income group.

8.8 Contingency Table of Economic Level, Banking and Insurance

We will consider here examining the associations between the three factors-economic class of a household, its having bank account and its having an insurance policy, simultaneously. For this, the contingency table is presented below Table 8.7.

8.9 A Log-linear Model to Describe the Three-way Contingency Table 8.7

Generalised linear models (GLM) with a log link are known as *log-linear* models (Christensen 1997). Log-linear analysis involves the generation of tables containing the possible combinations of the factors and a model of the expected frequencies of each combination, followed by the generation of statistical values that allow one to examine the significance of each factor and the goodness-of-fit of the model. The analysis typically begins with all the associations and interactions possible between the factors and then proceeds to eliminate those which can be, while retaining a reasonable fit between the expected cell frequencies generated by the model and the observed frequencies.

Data collected for categorical variables can be described in terms of a count of the number of instances which occurred for each combination of variables. Generalised linear models for count data (Log-linear Models) assume that the random component will follow a *Poisson* distribution. This distribution provides us with an estimate of the probability of the response variable—the frequency count of each cell of a contingency table and can therefore be used to make predictions about the likelihood of obtaining a value in each cell of the contingency table.

Table 8.7 Way contingency table of economic level, banking and insurance

			Insurance policy	
Economic class of a household			No	Yes
Low economic class	Bank account	No	179	0
		Yes	0	0
Lower middle economic class	Bank account	No	257	1
		Yes	14	0
Upper middle economic class	Bank account	No	246	3
		Yes	14	3
High economic class	Bank account	No	112	10
		Yes	14	28

Source Primary survey conducted during 2003–2004

Here we are interested to build a parsimonious model for prediction purposes, i.e. to predict the frequency count of each cell of the contingency table.

When all possible combinations of factors, i.e. all possible effects, are included in a log-linear analysis, then the model is said to be *saturated*. The saturated model in our case will be

$$\text{Log}\,\mu_{ijk} = \lambda + \lambda_i^E + \lambda_j^B + \lambda_k^I + \lambda_{ij}^{E,B} + \lambda_{ik}^{E,I} + \lambda_{jk}^{B,I} + \lambda_{ijk}^{E,B,I} \tag{8.4.2}$$

(i = 1, 2, 3, 4; j = 1, 2; k = 1, 2)

where μ_{ijk} are the cell frequencies; the letters **E, B** and **I** are simply variable labels and stand for 'Economic Class', 'Banking' and 'Insurance' respectively and do not raise λ to the *power* of E or B or I and i (= 1, 2, 3, 4), j (= 1, 2) & k (= 1, 2) are the levels of these three categorical variables in that order.

A saturated model provides a perfect fit for the data, i.e. the expected frequencies equal the observed frequencies. Saturated models do not offer any simplification or smoothing of the data and since the observed values are the same as the expected ones, we cannot obtain any goodness-of-fit statistics. In model building, one aims to build a simpler description of the data through the inclusion of only some of the factors and their combinations. To put it another way, the rationale behind generating a model is to obtain an *unsaturated* model with the fewest effects, but which still approximates the observed frequencies sufficiently well.

We had tried to fit many unsaturated models to our data and found that the following model is a good fit of the observed frequencies.

$$\text{Log}\,\mu_{ijk} = \lambda + \lambda_i^E + \lambda_j^B + \lambda_k^I + \lambda_{ij}^{E,B} + \lambda_{ik}^{E,I} \tag{8.4.3}$$

(i = 1, 2, 3, 4; j = 1, 2; k = 1, 2)

The Model 4.3 [hereinafter we shall refer to it as **Model (EB, EI, BI)]** is nested within Model 4.2 [Model (EBI)] as all the terms which appear in this model also appear in Model 4.2. Simply speaking, Model (EB, EI, BI) is a subset of the Model (EBI). Both the models are *hierarchical*, which means that each term includes all lower order terms.

The *likelihood ratio statistic* (G^2), which has a chi-square distribution for large samples, for the Model (EB, EI, BI) is 1.5335 with 3 degrees of freedom and the p-value is equal to 0.6746. So the Model (EB, EI, BI) is a good approximation of the data since the G^2 is *non-significant*, because we are looking for models in which the expected frequencies generated by the model are not significantly different from the observed cell frequencies. This is an important point, which is in contrast to the common practice of looking for p-values of 0.05 or less, and significant differences. In a log-linear model, we are testing the difference between the expected counts from the model and the actual data, rather than comparing a model of independence with the observed counts as one would do in a traditional chi-square test.

Table 8.8 Three-way contingency table of economic level, banking and insurance with the *predicted counts*

			Insurance policy	
Economic class of a household			No	Yes
Low economic class	Bank account	No	179	0
		Yes	0	0
Lower middle economic class	Bank account	No	257.51	0.49
		Yes	13.49	0.51
Upper middle economic class	Bank account	No	246.14	2.86
		Yes	13.86	3.14
High economic class	Bank account	No	111.35	10.65
		Yes	14.65	27.35

Source Primary survey conducted during 2003–2004

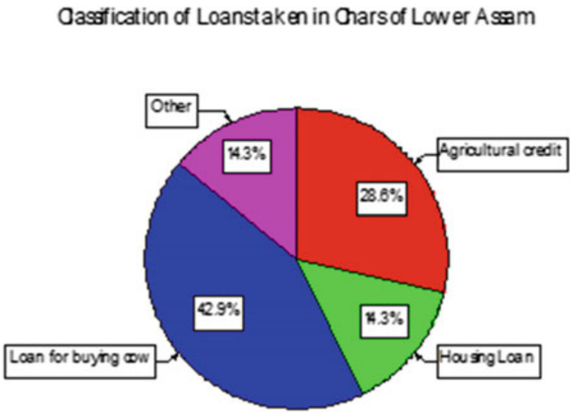
There are in total 29 parameters, including the constant λ, in the Model (EB, EI, BI). Of these 29 parameters, 16 are aliased (or redundant) parameters and they are set to zero. Most of the statistical software packages give the estimates of the parameters along with their 95% confidence intervals. In this way, the log μ_{ijk} s (i = 1, 2, 3, 4; j = 1, 2; k = 1, 2) can be obtained. By taking exponentiation of the log μ_{ijk} s (i = 1, 2, 3, 4; j = 1, 2; k = 1, 2), we would get the required 16 expected (predicted/fitted) frequencies μ_{111}, μ_{112}, μ_{121}, μ_{122}, μ_{211}, μ_{212}, μ_{221}, μ_{222}, μ_{311}, μ_{312}, μ_{321}, μ_{322}, μ_{411}, μ_{412}, μ_{421} and μ_{422}.

In log-linear analysis, adjusted residuals larger than 1.96 in absoluter magnitudes indicate a poor model fit for that cell of the table, while the sign indicates the direction of the discrepancy. Negative values indicate expected values lower than observed, and positive values indicate expected values that are larger than the observed count for that cell. In our case, there is no such adjusted residual which is greater than 1.96 in absolute magnitude Table 8.8.

8.10 Institutional Loan in the Char Areas of Lower Assam

Next we investigated whether the char households have availed the facility of taking loan either from cooperative societies or from banks or against their insurance policies (if any) during the last 5 years from the reference date of the survey (i.e. in between 01-01-1998 and 31-12-2002). What we have found out is only a meagre 1.6% sampled households have been able to avail this facility. On making further enquiry on this issue, we came to know that 42.9% of these households have taken loan for buying cow. Another 28.6% households told that the loan taken was an agricultural credit. This is another indicator of the agrarian economy of the char areas of lower Assam Fig. 8.6.

Fig. 8.6 Pie chart showing classification of loans taken in chars of lower Assam



We feel that there is an urgent need to provide cheap loan to the people of char areas who depend mainly on agriculture for their livelihood. The banks can provide loans for buying seeds, manures, pesticides, chemicals and other vital inputs like thresher, tractor, etc. which are needed in agricultural activities. This way the agricultural output would increase and the farmers will also be able to repay the loan within the stipulated period. Hence the nationalised banks should devise some special schemes for the char areas.

8.11 Economic Status of Households that Have Taken Institutional Loan

The cross-tabulation between the households that have taken loan from a cooperative society or bank or against insurance policy during the last 5 years (from the reference date of the survey) and its economic status reveals that only the upper middle income group and high income group households have been able to avail this facility. The households in the low income group and lower middle income group went unrepresented as far as taking loan is concerned. The proportions of upper middle income group and high income group households among the households that have taken loan are 28.6 and 71.4% respectively Table 8.9.

Table 8.9 Cross-tabulation of income level and institutional loan for households

Cross-tabulation of income level and institutional loan							
		Loan from co-op/Bank/Insurer				Total	
		No		Yes		Count	Col %
		Count	Col %	Count	Col %		
Income level of a household	Low income group	246	28.4	0	0.0	246	27.9
	Lower middle income group	205	23.6	0	0.0	205	23.3
	Upper middle income group	206	23.8	4	28.6	210	23.8
	High income group	210	24.2	10	71.4	220	25.0
Total		867	100	14	100	881	100

Source Primary survey conducted during 2003–2004

References

Agresti, Alan. (1984). *Analysis of ordinal categorical data*. New York: John Wiley and Sons.

Agresti, Alan. (1990). *Categorical data analysis*. New York: John Wiley and Sons.

Christensen, Ronald. (1997). *Log-linear models and logistic regression* (2nd ed.). New York: Springer.

Socio-economic Survey Report (2002–2003) of char areas of Assam (2004). Directorate of char areas development Assam.