

INDIAN STATISTICAL INSTITUTE
CHENNAI CENTRE
M.Stat. : 2015-17
(Year I – Semester II)

Semester Examination – Categorical Data Analysis

Date: 6th May 2016

Duration : 3 Hours

ANSWER ANY FOUR QUESTIONS. ALL QUESTIONS CARRY EQUAL MARKS. TOTAL MARKS IS 40.

1. A general social survey has Y = belief in life after death, with categories (Yes, Undecided, No), and explanatory variables x_1 = education status and x_2 = gender. The survey data is shown in the following table. Develop an appropriate model. Test for goodness-of-fit of the model.

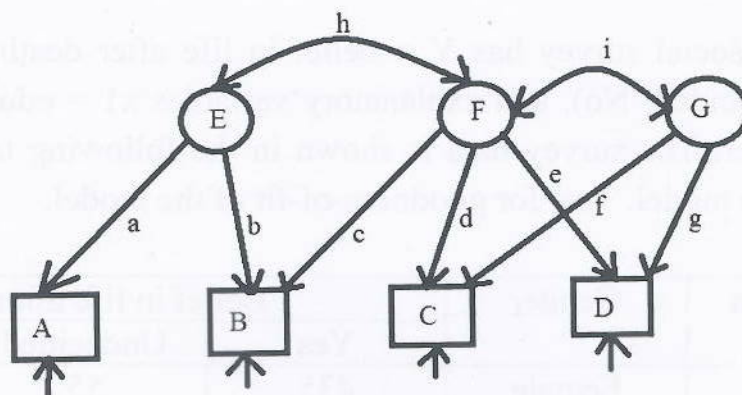
Education Level	Gender	Belief in life after death		
		Yes	Undecided	No
Not Educated	Female	435	55	85
	Male	286	44	70
Well Educated	Female	68	12	25
	Male	31	6	23

2. A longitudinal study was conducted for comparing a new drug with a standard drug for treating patients suffering mental depression. Patients were classified into two groups according to the initial severity of depression [mild (0), severe (1)]. In each group, patients were randomly assigned to one of the two drugs [standard (0), new (1)]. Following 1 month, 2 months, and 3 months of treatment, each patient's extent of suffering from depression was classified as Normal (1) or Abnormal (0). From the sample data, a model was estimated as

$$\log it[P(Y_t = 1)] = -0.51 - 1.316 * \text{severity} - 1.071 * \text{treatment} \\ + 0.483 * \text{time} + 1.012 * (\text{treatment} \times \text{time})$$

Compute expected proportions of Normal response for all categories at each time point.

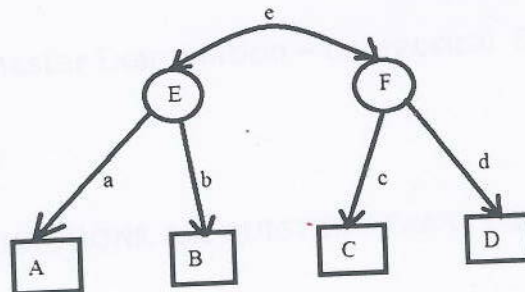
3. Consider the following three factor model:



- Find correlations between the manifest variables.
- Write the structural equations of this model.
- Give your comments about finding solutions of unknown constants a- i assuming that the sample correlation matrix is known.

4. Consider a three-way contingency table with each variable having only two levels. Assume that all observed cell frequencies are known. If we want to fit an additive log-linear model, how to estimate model parameters? How to test goodness-of-fit of such model? How to estimate the model parameters if someone opts for a saturated model?

5. Consider the following path diagram.



The sample covariance matrix of the four observed variables is:

	A	B	C	D
A	40.46			
B	12.20	14.27		
C	15.04	9.36	11.51	
D	18.89	15.67	11.89	25.11

Find values for the unknown constants a, b, c, d, and e.
