Car Insurance

Topic: Categorical variables, covariation and explanation

Context:

The insurance company CI specializes in car insurance, and wants to get some preliminary views from the market about added features to the comprehensive policy. At a break down within the country all expenses will be covered for home transport of passengers and car to a designated repair shop close to the owner. This will be offered for an additional charge less a deductible of NOK 2000.

The marketing manager in the company has interviewed a number of car owners with comprehensive insurance over telephone, about as many from own company as from the competitors. The first persons interviewed were sampled randomly from the telephone directory. This way information was obtained from 70 comprehensive policy holders, of these 10 were insured by CI. These were supplemented by 50 insured sampled from own customer file, thus having data from 120 respondents, 60 of each kind. The key question to be answered was:

"Will this supplement to the comprehensive insurance be of interest to you?"

Respondents who find the offer to be of "high" and "very high" interest will in the context of this case be labelled "the target customers".

Altogether the following information was gathered:

1. Insurance company (1=CI, 2=Other)
2. Car ownership (1=private, 2=firm)
3. Usage (1=single, 2=shared)
4. Age of major user
5. Car age (0=less than a year, 1=one year,...)
6. Driving length yearly
    (1=less than 16 000 km, 2=16 000-30 000km, 3=more than 30 000km)
7. Expected number of trips more than 500km away from home per year
8. Opinion about offer (on scale from 1 to 5 according to interest:
    1=no interest, 2=low interest, 3=moderate interest,
    4=high interest, 5=very high interest)

The data are stored as separate column in file Car_Insurance.XLS.

Task (version A):

What does this survey tell about the interest in the market for this kind of offer, and how is it related to the policy holder characteristics? What kind of supplementary information would be needed for decision support?

Task (version B):

Recode Opinion to a binary variable Target (1=target customer, 0=otherwise).

(a) Estimate the following population fractions, and give the error margins with an (approximate) confidence guarantee of 95%:

    (i) fraction of target CI customers
    (ii) fraction of target non-CI customers
    (iii) difference between the two fractions

    Only if the fraction of target customers among their own is above 20% they will go on developing the product (and do more market research). What is the conclusion? Are there reasons to believe that the two fractions are different?

(b) Explore whether and how the opinions on the new product differ between customer segments (defined by the other variables).

Hint: This may alternatively be done
(i) on recoded binary variables or the original ordinal opinion variables
(ii) in bivariate or multivariate manner (taking covariation into account)
(iii) by cross-tabulation or regression

(c) Are there interesting patterns in the data that may be of help when planning the development and/or the marketing the product?   Give critical comments on this "market study" and what you would have done differently.