

Midterm Exam Winter 2006 ANSWER SHEET**Question I:**

- a) Do victims differ demographically from the general population? Using table 1, describe how the two subsamples seem to differ.**

Victims are less likely to be male, less likely to be very young or over 60 years, more likely to be African American, come from larger households, less likely to have incomes less than \$25,000 or more than \$75,000, less likely to have a college education or to have a parent with a college education. Victims were also less likely to have refused to answer the income question (4% versus 12% for the general sample).

These differences may be causal (e.g., having lower education may make you more vulnerable to predatory lending) or may be tied up with the fact that all victims did have loans (e.g., fewer lower income families could afford to finance a house so would be less likely to have a loan). In particular, having a larger family (perhaps having children), being at least middle income, and over 30 years old probably all lead someone to be more likely to buy a home and therefore to have a mortgage. So these groups are more likely to be in the victim pool because they have loans, though those characteristics might not make you more vulnerable to predatory lending GIVEN that you have a home loan.

- b) Does household size differ for victims and the general population? Calculate the mean and standard deviation of household size for the victim pool and general samples.**

See spread sheet for calculations. Victims averaged 3.2 people per household with a standard deviation of 1.48. The general population average 2.68 people per household with a standard deviation of 1.4 people. So households for victims are slightly larger. This could merely reflect the characteristics that make someone more likely to buy a home OR larger households could have more pressure to take on debt even among homeowners.

- c) How does income differ for victims and the general population? Calculate the chances that a person in the victim pool reported a household income less than \$55,000. Then calculate the chances for the sample from the general population. If all the people who refused to answer the question had incomes below \$55,000 what would the chances be for each sample? What would the chances be for each subsample if the people who refused to answer the question had incomes with the same distribution as those who answered?**

See spreadsheet for calculations. The probability that a victim reported income Less than \$55k is **.47** and for the general sample is **.49**. So, victims are only slightly less likely than the general population to be low income.

If all refusals were low income, the probabilities would be **51%** and **61%** respectively. This is a much bigger gap because more of the general population refused to answer the question.

To figure out what would happen If refusals had incomes distributed the same as those who reported, we need to find the probabilities of low incomes GIVEN that someone answered.

$$P(\text{Income} | \text{didn't refuse}) = P(\text{Income AND didn't refuse}) / P(\text{didn't refuse})$$

So we can rescale the probabilities by .96 for victims and .88 for general (the % who didn't refuse). If those who refused had the same distribution, then the % would stay the same when we added them in. So, the rescaled probabilities of low incomes are **49%** for victims and **56%** for the general population would have low incomes.

So, although about the same proportion of each group reported low incomes, once we account for the refusals, it's very likely that victims were LESS likely to be low income than were the general sample. Again this could be due to selection into the pool of home buyers.

d) *How is refinancing a mortgage related to being a victim?* Everyone was asked whether they had refinanced a home, but in order to refinance a mortgage, one needs to have bought a house. Write out this issue in probability terms and use the information in Table 2 to calculate the chances that someone refinanced their mortgage, given that they had bought a house (separately for both victims and the general sample).

$$P(\text{Refi} | \text{bought home}) = P(\text{refi and home}) / P(\text{home}) =$$

$$\text{Victims: } = .88 / .98 = \mathbf{.90}$$

$$\text{General: } = .37 / .77 = \mathbf{.48}$$

So you can see that even after accounting for the fact that many of the general sample had not bought homes so couldn't have refinanced, the rate of refinancing among victims who had bought homes (90%) was still much higher than the rate for those in the general sample who had bought homes (48%).

e) *How big is the group who are likely targets of predatory lending?* Suppose that, instead of two separate samples, both sets of data came from a random sample of the population. What would you estimate as

**the chances that someone would be a victim of predatory lending?
Explain why.**

If the two groups came from a random sample of the general population, then we could combine them into one group. If no one in the general sample was a victim (a good assumption if being a victim is rare), then our best guess at the chance of being a victim is :

$$P(V) = 862 / (499 + 862) = .62.$$

That seems very high (and would negate our initial assumption that being a victim was rare)! [Of course, they didn't come from one big random sample—it was a stratified sample with heavy oversampling of victims.]

Suppose instead that you were given the individual level data from this study. What strategy would you use to figure out who in the general sample might be especially likely to be victims of predatory lending. Given the information in the tables, what percentage of the general sample do you think would fall into that group?

We want to find factors that differ greatly for the victim sample and general sample—factors that are NOT statistically independent of victim status and use them to screen the general population.

As we already noted **homeownership** and **refinancing** are very different for the two groups. Also, note the big difference in the number of people who **paying off credit card balances each month** (25% of victims and 63% of the general sample).

If I had the individual data, I could screen the general sample for those who HAD refinanced AND didn't pay balances in full. Since for the general sample 37% had not paid balances in full, my sub sample would be smaller than that. If refinancing were independent of that, then my new subsample of those with both would likely be $p(\text{refi}) * p(\text{not paid in full}) = .37 * .37 = .14$. But it could be the same 37% people who do both so as large as .37. Most likely it is somewhere in between.

f) Oregon officials are trying to develop appropriate policies to prevent consumers from being victims of predatory lending. They are considering targeted education campaigns. Write a short description of your results (less than 1 page). Provide policy implications for their campaigns given your results and any other relevant information in the tables.

The Oregon Department of Consumer and Business Services recently sponsored a telephone survey to aid us in understanding the characteristics

victims of predatory lending and to craft appropriate outreach and education campaigns. We surveyed a group of households with loans with a company recently sanctioned for its lending practices and a comparison group from the general state population.

On the whole, the demographics of the sample from the predatory loan group (“victims”) and those from the general sample were very similar. Those in the victim sample were slightly more likely to be female, had slightly larger households (average of 3.2 people compared to 2.7 in the general sample), were more likely to be in the middle age groups, and were more likely to be in the middle income groups rather than low or high. Many of those characteristics may have been related to the fact that almost all the victims had bought a house and most home owners are middle-aged, have at least a middle income, and have larger than average households. Other demographics may be indicative of direct impacts on the chances of being a victim. For example, victims were less likely to have a college education than the general sample, and less likely to have had a parent with a college education. This may suggest that having a college education makes being a victim less likely.

Differences in some household financial practices are even more marked between victims and the general sample. For example, those in the victim sample were more likely to have bought a home (98 percent vs. 85 percent for the whole sample) and, among home owners, more likely to have refinanced their home (90 percent versus 48 percent), and much less likely to pay credit cards in full (25 percent versus 63 percent). Clearly, being a homeowner makes one more likely to take a loan and therefore is likely to increase the chances of being a victim. However, refinancing a home and not paying credit card bills in full, could either lead to needing extended credit or, alternatively, could be the result of having very high interest loan that makes it difficult to pay other obligations.

Given the information from this study, we suggest that educational campaigns be targeted to those likely to become homeowners (especially families with at least middle incomes) and perhaps be packaged with other financial literacy education (e.g., good credit practices and information on managing debt and on appropriate debt levels).

Our targeting of campaigns would benefit from having better information about the number of victims in our state so future studies should ask questions that help us assess whether or not a family has been victims of predatory lending (or been targeted by the companies). Information on victims for this analysis was limited in that it came from customers of only one company. Longitudinal information on family financial behaviors (e.g., credit card payments) could also help us distinguish between outcomes of predatory lending and the behaviors that make families vulnerable.

Question II:

An energy study in Gainesville Florida found that household energy use of electricity has a mean of 1200 kilowatt-hours (kWh) and a standard deviation of 500 kWh in a typical summer month. In June 2002, to encourage conservation, Gainesville Regional Utilities announced that the base cost would be \$.05 per kWh up to 1500 kWh and then \$.06 per kWh beyond that.

- a) If the distribution is approximately normal, about what proportion of households would have their bill computed entirely on the base amount?

$$\begin{aligned} P(X < 1500) &= P(Z < (1500 - 1200)/500) \\ &= P(Z < .6) \\ &= .5 + P(0 < Z < .6) \\ &= .5 + .2257 = .7257 \end{aligned}$$

So 73.7 percent of households would have their bills computed with the regular rate.

- b) How much would a household at the median and 75 percentiles of kWh pay in a month?

The median for a normal distribution is also the mean, so here it is 1200. So for the first part we know that a household at the median would pay:
1200*.05 = \$60.00

For the 75th percentile, we need to find the Z score that gives us 75 percent of the distribution to the left and convert it to the X scale.

$$P(Z < Z_0) = .75$$

But we can use the table to get the Z with .25 above the mean (since .5 is below)

$$\rightarrow P(0 < Z < Z_0) = .25$$

Look up .25 in the body of the Z table to get:

$$P(0 < Z < .67) = .2486 \quad \text{so our } Z_0 = .67$$

Now use the formula to convert to X_0 :

$$X_0 = \mu + ZS = 1200 + (.67)(500) = 1535$$

So, these people will pay .05 up until 1500 then .06:
1500*(.05) + 35(.06)=77.10

- c) If the utility chooses to put a note on conservation in the bill of any household with more usage more than 1.5 standard deviations above

the mean, what percent of households would get a note? What would be the cut-off in kilowatt hours?

We need the X_0 score and probability associated with households with consumption more than 1.5 standard deviations above the mean.

$$P(Z > 1.5) = P(X > X_0) \rightarrow X_0 = \mu + ZS = (1200) + (1.5)(500) = 1950$$

$$P(Z > 1.5) = .5 - P(0 < Z < 1.5) = .5 - .4332 = .0668$$

So, with a cut-off of 1950 kWh for households getting the note, about 7 percent of households would get the note.

d) If winter usage is lower by 200 kWh, on average, what percentage of clients would receive a note if the utility used the same kilowatt cut-off? What assumptions do you have to make in getting your estimate?

If only the mean of the distribution changed and we still used 1950 as our cut-off, then our new problem is:

$$P(X > 1950) = P(Z > (1950 - 1200)/500) = P(Z > 1.9) = .5 - P(0 < Z < 1.9) = .5 - .4713 = \mathbf{.0287}$$

So, now about 3 percent of households would get the notes. We had to assume that only the mean changed (not the standard deviation) which seems a little difficult to believe. It's likely that the variation would also decrease in the winter would leave the percentage receiving notes even lower in the winter. We also assumed that the distribution was still approximately normal—a little easier to imagine.