# Dealing with Distributions of Behavior Frequencies: An Example with Alcohol Use

Emilia Peytcheva, University of Nebraska-Lincoln
Andy Peytchev, University of Michigan

**KEY WORDS: Behavior frequencies, Alcohol, Left censoring, Poisson, Tobit, Mixture models.**

## Introduction

As a substantive interest in areas such as psychology, frequency of alcohol use is often an outcome (dependent) variable or an indicator (independent) variable in a wide range of statistical analyses. As in such fields the replication of results is the basis for future decisions, affecting human subjects and society as a whole, detection of consistent analytical mistakes across studies are of great importance. The purpose of this paper is not to criticize prior works, but to show evidence whether frequency of alcohol use is being used correctly and propose solutions to any problems. The importance is rather global in scope as any findings are likely to be applicable to other frequently used substance reports, such as tobacco and marijuana.

Frequency of alcohol use is a typical example of a ratio variable, with an absolute zero and no upper limit (theoretically). Since the underlying distribution is continuous, frequency scales are treated as continuous in statistical analyses in the social sciences despite the bracketed measure. However, these characteristics alone are insufficient for justifying the variable's use in common parametric analyses. The more apparent violation is the distribution of the variable – especially for frequency of alcohol use among adolescents in major studies like the National Longitudinal study of Youth (NLSY) and the National Longitudinal study of Adolescent Health (Add Health), the distribution is one approximating Poisson and not the occasionally assumed Gaussian distribution. This violation results in underestimation of associations, particularly with variables that are skewed to the right (opposite direction). Also because of the monotonic slope of the distribution, the associations are heavily influenced by cases with high frequency of alcohol use, as they act as outliers. These problems are difficult and sometimes impossible to overcome, as for example, an analysis that requires the fitting of Structural Equation Models is currently not possible without the assumption of normality[1] (e.g. Bui, Ellickson, and Bell, 2000). However, in other studies where frequency of alcohol use is the dependent variable in a multiple linear regression (e.g. Resnick et al., 1997, where Add Health data from the first wave was used), the absence of alternatives is no longer a justifiable explanation, as more appropriate methods like Poisson regression are readily available to handle such data.

**Table 1: Frequency of Alcohol Use in Wave 1 of the Add Health study**

During the past 12 months, on how many days did you drink alcohol?

| | |
|---|---|
| 0 | Never [skipped as it is asked in a previous question] |
| 1 | Not in the past 12 months |
| 2 | 1 or 2 days in the past 12 months |
| 3 | Once a month or less (3-12 times in the past 12 months) |
| 4 | 2 or 3 days a month |
| 5 | 1 or 2 days a week |
| 6 | 3 to 5 days a week |
| 7 | Every day or almost every day |
| | Refused |
| | Don't know |

There are other problems that arise from the violation of the normality assumption of various analytical methods, but it is the consideration of the substantive structure of alcohol use when selecting a statistical method that is the focus of discussion in this study. The alcohol variable found in the Add Health study, shown in Table 1 and it's distribution displayed in Figure 1d, is the one commonly used in analyses using the alcohol frequency question asked in Wave 1 of the Add Health study. Superficially, it seems rather convincing evidence to simply use it as a Poisson variable as supported both by the distribution (Figure 1a) and the meaning (self-reported average *count* of number of drinks consumed in the past 12 months).

[1] M*plus* version 2.12 allows the inclusion of categorical variables in the models, but not ordinal or non-normal continuous variables.

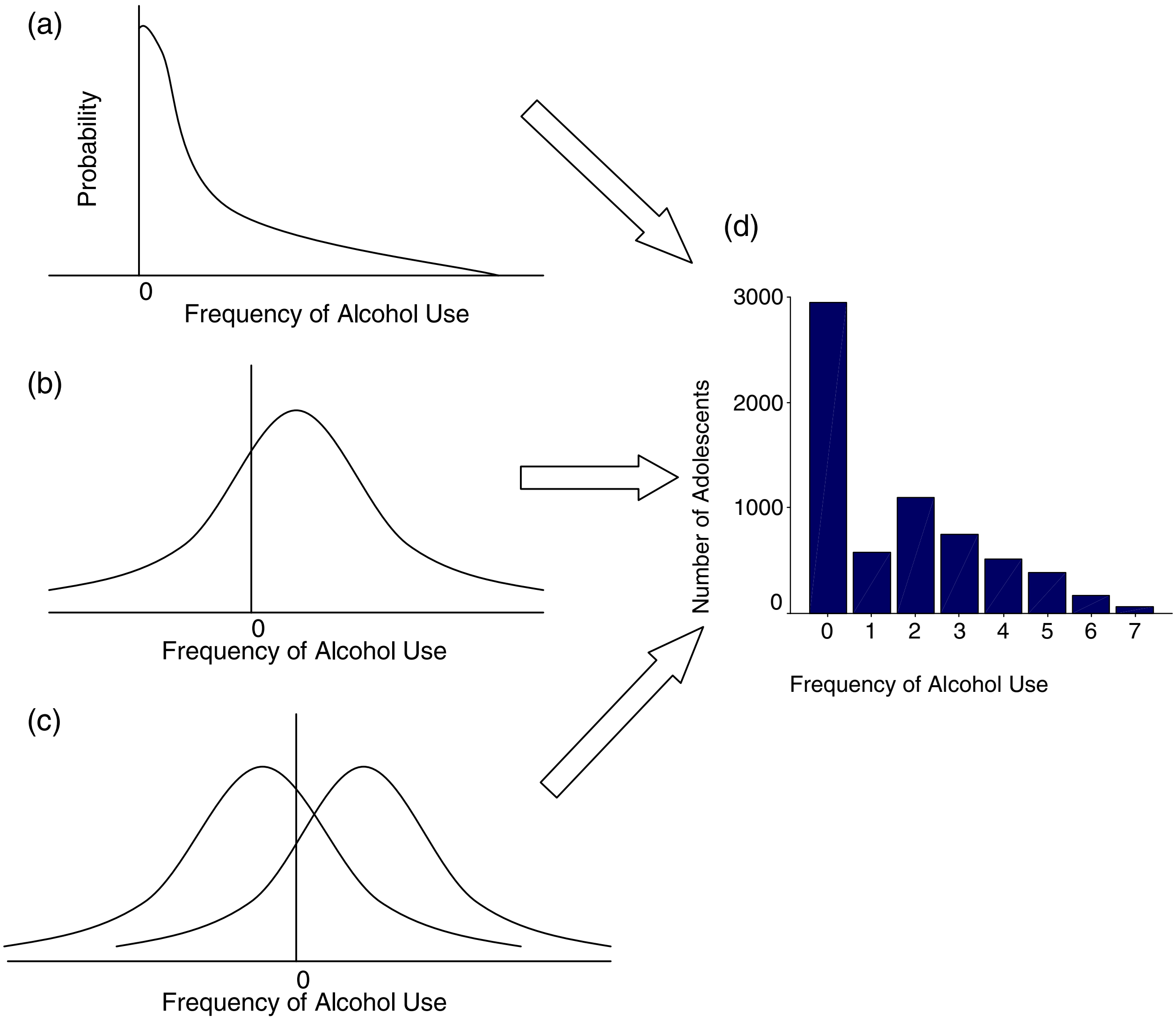

**Figure 1: Distribution of Frequency of Alcohol Use in Past 12 Months (d) and Possible Underlying Distributions: (a) Poisson; (b) Left-censored; and (c) Left-censored multiple populations.**

However, there are two other possible explanations for the shape of this distribution, which have been developed in the field of econometrics. The first is that the underlying distribution is indeed two-tailed but is truncated, as negative values for frequency of a behavior do not exist, otherwise known as left censoring. It is based on the premise that there are people who do not want to drink, but there are others who are even less likely to ever drink, all placed on an underlying continuum that cannot be measured by this indicator, which has a minimum of zero, as illustrated in Figure 1b. If this is the case, then selection models such as the Tobit model[2] (Tobin, 1958) would be appropriate as they model this truncation using the same covariates that are used in modeling values above zero. The other possibility is that there are two or more overlapping underlying population distributions, for example, one for inherent abstinents and another for those who do or would drink, as shown in Figure 1c. This is an example of a mixture model, a broad family of models that can be fitted to various types of data.

If the underlying distributions are ignored, a model will be misspecified, and the combination of using an inappropriate method and ignoring the

[2] The name is a compilation of the author's name and *probit* analysis, aspects of which the model is based on.

complex survey design of such national studies (both misspecifications were found in major studies (e.g. Resnick et al., 1997) could invoke misleading results, finding nonsignificant relationships to be significant and vice versa. This is related to the issue with underlying distributions as sampling distributions in calculations assume simple random sampling, hence the omission of design effects such as stratification and clustering can severely bias the results and underestimate standard errors. Weights are also necessary for generalizations to a population and only the non-inclusion of post-stratification is not as detrimental, as it merely leads to more conservative tests.

## Methods

These analyses are based on the public use Add Health data, a nationally representative probability based sample of adolescents in grades 7 through 12 (from the large national studies, the National Longitudinal Survey of Youth would have necessitated the use of the second generation cohort, which lacks generalizability, while the Monitoring the Future study conducted by the University of Michigan's Institute for Social Research is a retrospective study that lacks the desired concurrent measures). The Add Health data was collected in three samples over time: in-school questionnaire (September 1994 – April 1995, n=90,118); in-home wave 1 (April 1995 – December 1995, n=20,745); and in-home wave 2 (April 1996 – August 1996, n=14,738). Only the latter two samples are nationally representative with weights, four strata based on U.S. geographic regions, and clustering (the sampling frames came from schools in the U.S.). Wave 1 data was preferred as it is not susceptible to attrition bias. In wave one 1,821 adolescents did not have weights as they were in a genetic sample, interviewed for a different analytical objective in the Add Health study. From those with available weights, 6,504 were randomly selected for the public use data. The region stratification identifier, however, was not among the public use variables. Its absence in analyses is not a large threat in terms of bias in estimates as differences in the variables of interest are much more likely to be affected by factors like the level of urbanization than by a four-region classification of the U.S. However, efficiency gains from stratification will not be possible, hence statistical significance tests may tend to be more conservative. Since race is also included in the current study, it is important to note that the public use sample includes data from the originally selected core sample, the high education black supplement sample, or both (probability of selection is accounted for in the weights).

The purpose of this study is not to build the best possible model explaining the frequency of alcohol use by adolescents by entering predictors that have not been used in prior research, but rather to build models that are plausible and substantiated as it is crucial to the generalizability of any methodological findings. If the former was the case, a split halves approach would have been more appropriate. Three demographic variables were entered: age, gender, and race. Gender has been found to have a strong association with alcohol at all age levels (e.g. White and Labouvie, 1989) to the extent that different quantifying criteria for binge drinking are used for girls and boys, while race has also been found to be related as studies have found that blacks are much less at risk than whites (e.g. Cook and Moore, 2001). In addition, seven psychological, social, and behavioral scales were created from existing scales and coherent groups of questions in the Add Health study: Depression (15 items from the "Feelings Scale," such as "You felt sad" and "You felt depressed," α=.899), Happiness (the remaining 4 items from the "Feelings Scale," which had moderate first-order correlations among items, but very low correlations with the rest of the items, such as "You were happy" and "You enjoyed life," α=.730), Anxiety (3 items from "General Health" – frequency of "Trouble relaxing," "Moodiness," and "Frequent crying," α=.594), Self-esteem (9 items from the "Personality and Family" section, such as "You have a lot to be proud of" and "You like yourself just the way you are," α=.874), Parental Control (7 items from the "Relations with Parents" section, such as "Do you let your parents make your decisions about the time you must be home on weekend nights," α=.968), Delinquency (all 15 items in the "Delinquency Scale," such as "Take something from a store without paying for it," administered by ACASI, α=.939), and Deviant Peers (3 items from the ACASI section on "Tobacco, Alcohol, and Drugs," asking for a self report of how many from the respondent's 3 best friends exhibit a specified frequency of smoking, drinking, and marijuana use, α=.800). Scales similar in meaning have often been used in analysis of alcohol use, e.g. Cahalan, 1970; Donovan and Jessor, 1978; Fillmore, 1974; Mayer and Filstead, 1980.

From the 6,504 respondents, 15 were missing frequency of alcohol use in the past 12 months but were not imputed as it would mean the assumption of one of the distributions that are being tested in the current study. The 4 respondents who were below 12 years of age were removed, as were the 144 adolescents over 18 years old (studies such as

Harford and Mills, 1978 have found that young people drink less often than adults, and when they drink, they tend to drink in larger quantities – one of many reasons for excluding the few adults in the data). As the 3 respondents missing age and the 51 missing race had not provided answers to the large majority of questions as well, they were removed from the analyses. Missing values for any of the 7 scales were imputed using sequential regression multiple imputation by means of a Bayesian algorithm in IVEWARE® (a SAS-callable macro library program created by Raghunathan, Solenberger, and Van Hoewyk, at the Institute for Social Research, University of Michigan), accounting for weighting and complex sample design. Bayesian and sequential regression methods of imputation have been found to be superior over complete-case analysis, simple deterministic imputation such as mean substitution, and other methods like the hot deck method (Heeringa, Little, and Raghunathan, 2002). In order to account for any statistical uncertainty in the results, all analyses were performed simultaneously on 5 versions of the data set with independently imputed values. Furthermore, all models were tested on the imputed data and on the original data using listwise deletion. The minimum sample size for the former was 6,288 and for the latter was 6,092, when all variables of interest are used in a single model.

In order to test whether all represented values for frequency of alcohol use lie on the same continuum and which of the distributions in Figure 1 is most plausible, 8 different models were tested. The focus was on examining the substantive differences between responses to each option, i.e. is it the same set of factors that would determine where on this "alcohol use" continuum an adolescent lies, or is alcohol use a set of categories with a continuum on the right side and categories to the left, with different factors accounting for the likelihood of an adolescent's location in each dimension or is there an underlying continuum on both sides, just unobservable in the left tail. The models differ in their treatment of "NEVER"s (never had alcohol), "NO"s (not in the past 12 months), and "YES"s (gave some non-zero frequency of alcohol use). They also differ in terms of their treatment of the "YES"s – as a category or as a count distribution (Poisson). The resulting 3 Logistic, 3 Poisson, and 2 Tobit regression models are described in Table 2. The same set of independent variables (age, gender, race, and 7 scales) were entered in all the models, hence initially, the models differed only in the left side of the equation. Statistically non-significant covariate contributions (P≥.05) were removed (starting with the highest P-value) and the model re-estimated in a sequential and iterative procedure. Any disparity in statistical significance and/or magnitude of parameter estimates would reflect on conclusions of dissimilarities or similarities between two models in terms of the differences in the dependent variables. All models were tested in IVEWARE® using 5 datasets with multiply-imputed missing values.

**Table 2: Model Descriptions Using Response Codes for the Dependent Variable from Table 1.**

Model 1: Logistic regression modeling probability of (2+3+4+5+6+7) over (0+1)
Model 2: Logistic regression modeling probability of (2+3+4+5+6+7) over (1)
Model 3: Logistic regression modeling probability of (1) over (0)
Model 4: Poisson regression on (2, 3, 4, 5, 6, 7)
Model 5: Poisson regression on (1, 2, 3, 4, 5, 6, 7)
Model 6: Poisson regression on (0, 1, 2, 3, 4, 5, 6, 7)
Model 7: Tobit regression on (1, 2, 3, 4, 5, 6, 7)
Model 8: Tobit regression on (0+1, 2, 3, 4, 5, 6, 7)

## Results

All three reduced logistic regression models differed in which predictors were statistically significant, and for those predictors that were common to at least two of the models, they differed in the magnitude of the parameter estimates, as can be seen in Table 3. Figure 2 shows that if all parameters are considered, they even differ in signs (direction), albeit not statistically significant. Model 1 groups the first two response categories together as is commonly done in practice when dichotomizing alcohol use, and therefore can act as a contrast group for the other models as it provides estimates that would result from typical analysis. When comparing the three models, the difference in the sample sizes should not be a concern – none of the regression coefficients in Model 1 are close to failing to reject the null hypothesis, hence the retention of more independent variables can be attributed more to the variance and covariance structure rather than sample size. Any parameter that remains in Model 1 also remains in at least one of the other two logistic regressions. A finding from the second and third analysis, which is of importance is that 6 of the parameter estimates are unique to only Model 2 or Model 3 and only 3 of the estimates are common in the two models. Happiness, Anxiety, Self-esteem, race, and Parental Control, have an effect on frequency of drinking, but have no association with whether alcohol has ever been consumed.

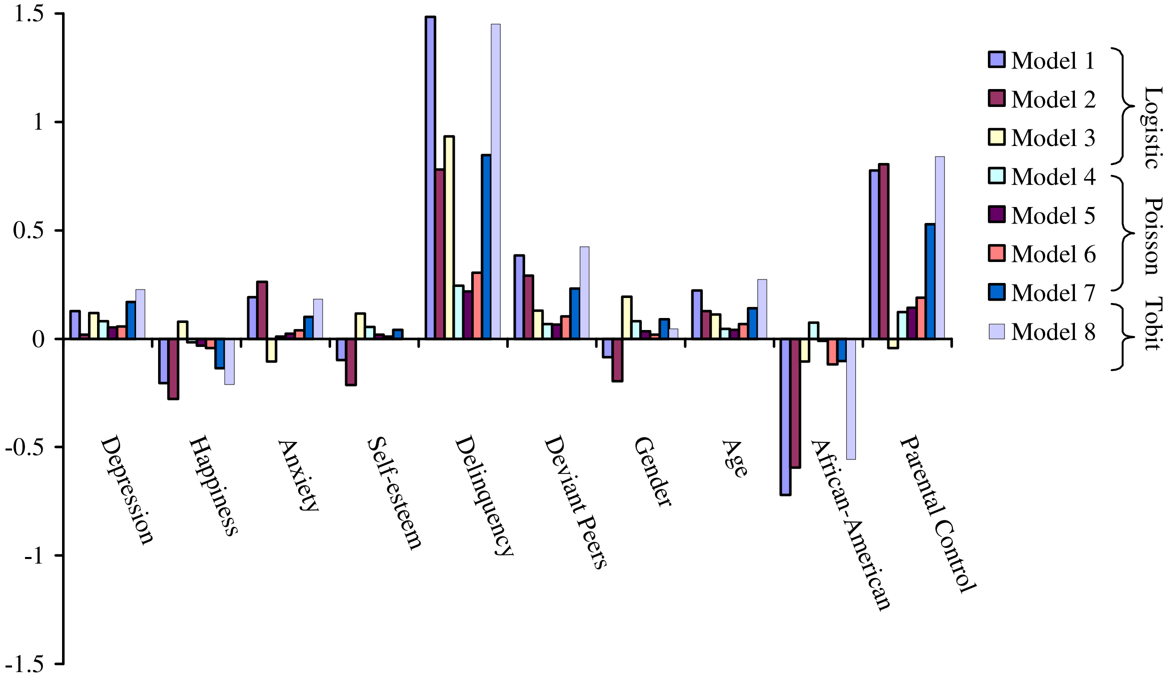

**Figure 2: Covariate Coefficients by Model, ignoring statistical significance.**

**Table 3: Statistically Significant Parameter Estimates**

| Regression Type | Logistic | | | | | | Poisson | | | | | | Tobit | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Model (Sample Size) | Model 1 | (6,288) | Model 2 | (3,434) | Model 3 | (3,447) | Model 4 | (2,891) | Model 5 | (3,461) | Model 6 | (6,288) | Model 7 | (3,434) | Model 8 | (6,288) |
| Independent Variable | $\hat{\beta}$ | $SE(\hat{\beta})$ | $\hat{\beta}$ | $SE(\hat{\beta})$ | $\hat{\beta}$ | $SE(\hat{\beta})$ | $\hat{\beta}$ | $SE(\hat{\beta})$ | $\hat{\beta}$ | $SE(\hat{\beta})$ | $\hat{\beta}$ | $SE(\hat{\beta})$ | $\hat{\beta}$ | $SE(\hat{\beta})$ | $\hat{\beta}$ | $SE(\hat{\beta})$ |
| Intercept | -5.494 | (0.384) | | | -3.728 | (0.558) | -4.279 | (0.252) | -2.120 | (0.056) | -2.721 | (0.040) | -1.953 | (0.376) | -6.353 | (0.380) |
| Depression | | | | | | | 0.147 | (0.050) | 0.051 | (0.014) | 0.052 | (0.014) | 0.214 | (0.096) | | |
| Happiness | -0.153 | (0.056) | -0.315 | (0.086) | | | | | -0.030 | (0.007) | -0.044 | (0.006) | -0.143 | (0.054) | -0.174 | (0.052) |
| Anxiety | 0.242 | (0.055) | 0.271 | (0.091) | | | | | 0.025 | (0.007) | 0.037 | (0.006) | | | 0.226 | (0.054) |
| Self-esteem | | | -0.344 | (0.090) | | | 0.113 | (0.039) | 0.021 | (0.008) | | | | | | |
| Delinquency | 1.484 | (0.147) | 0.695 | (0.231) | 0.919 | (0.224) | 0.386 | (0.034) | 0.219 | (0.008) | 0.303 | (0.010) | 0.887 | (0.072) | 1.487 | (0.102) |
| Deviant Peers | 0.386 | (0.022) | 0.296 | (0.031) | 0.130 | (0.024) | 0.117 | (0.007) | 0.065 | (0.001) | 0.101 | (0.002) | 0.232 | (0.012) | 0.425 | (0.017) |
| Gender | | | | | 0.233 | (0.118) | 0.131 | (0.037) | 0.033 | (0.007) | 0.021 | (0.008) | | | | |
| Age | 0.225 | (0.024) | 0.098 | (0.028) | 0.106 | (0.037) | 0.092 | (0.014) | 0.042 | (0.003) | 0.067 | (0.003) | 0.142 | (0.024) | 0.278 | (0.024) |
| African American | -0.712 | (0.120) | -0.561 | (0.142) | | | | | | | -0.117 | (0.016) | | | -0.533 | (0.146) |
| Parental Control | 0.764 | (0.183) | 0.816 | (0.289) | | | | | 0.143 | (0.026) | 0.189 | (0.020) | 0.541 | (0.195) | 0.810 | (0.188) |

Model 4 benefits from the additional information from having 6 categories and correct distribution for the analysis, which is reflected in the much smaller standard errors. Although only two of the parameter estimates in Model 4 are not present in Model 3, some of the common parameters are different in magnitude. None of those that are similar are reflective of the respondents' cognition and behavior (age, deviant peers, and to some extent gender). Furthermore, there were two additional independent variables in Model 4 that were very close to the specified alpha level, namely Parental Control (P=.056) and Asian Race (P=0.059), which were found statistically significant when the model was estimated without the use of multiple imputation. Three more coefficients become significant when the "NO" category is added in Model 5, yet all the common coefficients decrease in size. When the adolescents who never had alcohol are included in Model 6, the magnitude of the coefficients changes in both directions, race gains a significant effect, yet Self-esteem proves to be significant only for non-zero frequency reports.

Model 7 is in a way a mix between models 2 and 4 – it estimates the logististic relationship between the "NO"s and the "YES"s and it also regresses the non-zero values as in the Poisson regression model. Model 8 does the same between models 1 and 4, hence it includes the "NEVER"s in the analysis. All the estimated coefficients are larger in absolute value in Model 8 with the exception of Depression.

## Discussion

The vast difference between models 2 and 3 is strong evidence that adolescents who have never had a drink should not be analyzed together with those who have, as the likelihood of never having had alcohol over not having had alcohol in the past 12 months is explained by a different set of independent variables than the likelihood of not having had alcohol in the past 12 months over having had alcohol more recently. The impact on results from typical analysis (Model 1) is apparent – too many factors would be found to have an effect on frequency of alcohol use when in fact some of them should be attributed to explaining the degree of abstinence from alcohol. Another implication stems from this and is also supported by the above results – the magnitude of the effect of the variables explaining frequency of alcohol use will be underestimated as the model is "contaminated" by a group that has a different set of explanatory variables (or equivalently, predictors in this case).

The visual similarity between models 3 and 4 could lead one to think that the factors that determine the degree of alcohol abstinence are the same as those indicating frequency of alcohol consumption. However, from the common independent variables with similar magnitudes of parameter estimates, only age and deviant peers (based on substance use by closest friends) are comparable, hence it cannot be claimed that the models are very similar. Models 4, 5, and 6 show that even when the frequency categories are preserved, the covariate structures of both the "NEVER"s and the "NO"s are different from the structure of the non-zero frequency reports. However, this difference is not as large as when the non-zero frequencies are collapsed into a single "YES" category (Models 1 and 2).

The Tobit models (7 and 8) model the zero category separately from the non-zero values, hence either of the two parts of the Tobit model could have an influence. A visual comparison between estimates in models 7 and 8 could lead one to believe that the differences are not distinguishing, but keeping in mind that the part of the models evaluating the non-zero values is the same in both cases. Furthermore, Model 8 is more similar to Model 1 as it models the same logit relationship between the probability of "NEVER"s and "NO"s to that of the "YES"s. This relationship is so influential (an additional 2944 cases) that it increases some estimates, brings others to statistical significance, and eliminates the effect of Depression at α=.05 level. The great similarity between the parameter estimates in Model 1 and Model 6 indicates that it is the propensity to drink at all (in the past 12 months) rather than how frequently alcohol is consumed that the models are explaining.

So why not use this model in future analyses? To answer this question, all the conclusions from the 8 models have to be evaluated simultaneously. The differences in both magnitude and significance of parameter estimates support a mixture population distribution, as the one shown in Figure 3. When frequency of alcohol use is an independent variable in an analysis, it is much more justifiable based on these results to create two separate alcohol variables in order to estimate the effect of the "NEVER"s separately and yet still benefit from the ordinal/continuous non-zero part of the distribution ("YES"s), without muffling the results by forcing a line through a nonlinear and incoherent relationship. As the differences between the drinking frequency categories in terms of associations with other variables indicates, simply including interaction terms with the original variable in an analysis is not going to resolve the problem. If the analysis does not permit Poisson variables, using two indicator variables to denote "NEVER," "NO," and "YES" will lose some information in the categorization, but should allow more accurate estimation of relationships. If alcohol consumption is the dependent variable of interest, moving from a multiple linear regression to a more distributionally appropriate technique like Poisson regression is only the first step. The complex meaning of the zero category requires for it to be included in the model, hence a two-stage Tobit and mixture models allowing different types of predictors as implemented in Econometrics would be more appropriate. With large data sets, such as the Add Health study, Categorical Latent Class analysis is also possible, but will impose limitations on the number and data type of indicators, due to sparseness effects from the number of variables and the number of their categories.

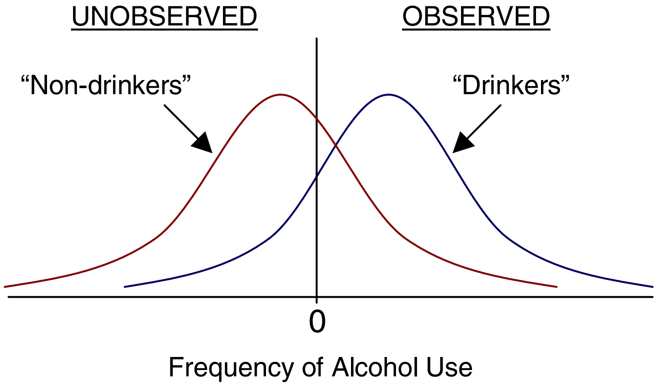

**Figure 3: Left-censored Mixed Population Distribution Model for Frequency of Alcohol Use.**

## Conclusion

There seem to be different factors affecting adolescent propensity to have never tried drinking alcohol and the frequency of alcohol consumption, hence it is erroneous to include all in a single continuous variable. Furthermore, simply reporting mean and standard deviation for such a variable does not confirm normality, but rather assumes it, as is the case of providing a Gaussian standard deviation on a Poisson distributed variable.

Although the issue with adolescent frequency of alcohol consumption is common to Psychology, possible solutions were borrowed from Econometrics (Tobit and Selection models for left-censored data) and even the very practice of comparing parameter estimates of the same predictors in models with different dependent variables has been used by Survey Research methodologists in contrasting location and cooperation factors in longitudinal surveys (Lepkowski and Couper, 2002). Furthermore, the estimation and comparison of such models is currently possible without ignoring the complex survey design – programs such as IVEWARE, SUDAAN, STATA, WesVar, and M*plus* estimate approximately unbiased standard errors when stratification and/or clustering is part of the design[3].

Failure to either divide such frequency distributions into separate variables, using a model that allows for differential modeling of the categories, or at least testing for the possibility of substantive differences between the categories undermines the validity of a study’s findings. Other substantive distributional issues not directly related to the focus of this study should also be considered, such as measurement equivalence[4] – is frequency of alcohol use an equivalent measure of the behavior across different subgroups in the target population. Other behavior frequencies should also be investigated.

## References

Bui KVT, Ellickson PL, and Bell R.M. Cross-lagged Relationships among Adolescent Problem Drug Use, Delinquent Behavior, and Emotional Distress. J Drug Issues 2000;30:283-304.

Cahalan D. Problem Drinkers. San Francisco: Jossey-Bass, Inc., Pubs., 1970.

Cook PJ, Moore MJ. Environment and Persistence in Youthful Drinking Patterns. Gruber J (ed.). Risky Behavior Among Youths: An Economic Analysis (Conference Report, National Bureau of Economic Research). U of Chicago Press, 2001.

Donovan JE and Jessor R. Adolescent Problem Drinking: Psychosocial Correlates in a National Sample study. J Stud. Alcohol 1978;39:1506-1524.

Fillmore KM. Drinking and Problem Drinking in Early Adulthood and Middle Age: An Exploratory 20-year Follow-up study. Q J Stud. Alcohol 1974;35:819-840.

Harford TC and Mills GS. Age-related Trends in Alcohol Consumption. J Stud. Alcohol 1978;39:207-210.

[3] These programs use variance replication techniques, such as Jackknife Repeated Replication (JRR) and Taylor Series Linearization methods.

[4] Methods of asserting it, like Item Response Theory, are provided by Knight and Hill, 1998

Heeringa SG, Little RJA, and Raghunathan TE. Multivariate Imputation of Coarsened Survey Data on Household Wealth. In: Groves RM, Dillman DA, Eltinge JL, and Little RJA (Eds.) Survey Nonresponse. New York, NY: John Wiley & Sons, Inc., 2002:357-371.

Knight GP and Hill NE. Measurement Equivalence in Research Involving Minority Adolescents. In: McLoyd VC and Steinberg L (Eds.) studying Minority Adolescents: Conceptual, Methodological, and Theoretical Issues. Mahwah, NJ: Lawrence Erlbaum Associates, Inc., 1998.

Lepkowski JM and Couper MP. Nonresponse in the Second Wave of Longitudinal Household Surveys. In: Groves RM, Dillman DA, Eltinge JL, and Little RJA (Eds.) Survey Nonresponse. New York, NY: John Wiley & Sons, Inc., 2002:259-273.

Mayer JE and Filstead WJ. Empirical Procedures for Defining Adolescent Alcohol Misuse. In: Mayer JE and Filstead WJ (Eds.) Adolescence and Alcohol. Cambridge, Mass.: Ballinger Publishing Co., 1980;51-68.

Resnick MD, Bearman PS, Blum RW, Bauman KE, Harris KM, Jones J, Tabor J, Beuhring T, Sieving RE, Shew M, Ireland M, Bearinger LH, and Udry JR. Protecting Adolescents from Harm: Findings from the National Longitudinal study of Adolescent Health. JAMA 1997;278:823-832.

Tobin J. Estimation of Relationships for Limited Dependent Variables. Econometrica 1958;26:24-36.

White HR and Labouvie EW. Towards the Assessment of Adolescent Problem Drinking. J Studies on Alcohol 1989;50:30-37.