THE ANALYSIS OF CATEGORICAL DATA FROM A COMPLEX SAMPLE SURVEY: CHI-SQUARED TESTS FOR HOMOGENEITY SUBJECT TO MISCLASSIFICATION ERROR

S. Heo and J. L. Eltinge, Texas A&M Unviersity
S. Heo, Department of Statistics, Texas A&M University
College Station, TX 77843-3143 (sunyeong@stat.tamu.edu)

Key Words: Heterogeneous misclassification probabilities; Contingency table; Logistic regression; Measurement error; Stratified multistage sample survey; Dual Frame National Health Interview Survey (NHIS)/ Random-Digit-Dialing (RDD) Methodology and Field Test.

1. Introduction

In the analysis of categorical data, if misclassification errors exist, then estimated cell probabilities may biased and standard Pearson chi-squared tests may have inflated true type I error rates. For some general background on the analysis of categorical data subject to misclassification, see, e.g., Mote and Anderson (1965), Tenenbein (1972), Hochberg and Tenenbein (1983)and Selén (1986). For specific work with misclassification problems in the analysis of stratified multistage sample survey data, see, e.g., Rao and Thomas (1991).

Rao and Thomas (1991) discussed methods to adjust chi-squared test statistics for goodness-of-fit with complex survey data subject to misclassification errors. They assumed that the misclassification probabilities are equal across all units in a given population.

This paper considers extensions of the Rao and Thomas (1991) method to tests of homogeneity, following Scott and Rao (1981). In addition, this paper examines cases in which misclassification probabilities may be heterogeneous within populations. For the latter case, we use estimated power curves to examine the extent to which heterogeneous misclassification probabilities may have a serious impact on inference. The proposed methods are applied to data from the Dual Frame National Health Interview Survey (NHIS) / Random-Digit-Dialing (RDD) Methodology and Field Test Project.

The authors thank Dr. Paul P. Biemer of the Research Triangle Institute and Dr. Van L. Parsons of the National Center for Health Statistics for providing the data from the Dual Frame NHIS/RDD Methodology and Field Test; and for helpful comments on this data set. This work was supported in part by the U.S. National Center for Health Statistics. The views expressed in this paper are those of the authors and do not necessarily represent the policies of the National Center for Health Statistics.

2. Notation

Suppose that there are two independent populations and that two independent samples of sizes n_1 and n_2, respectively are taken from these populations. In addition, suppose that there is a categorical variable with J mutually exclusive and exhaustive classes. Define $\pi_{i+} = (\pi_{i1}, \pi_{i2}, \ldots, \pi_{iJ})'$ and $p_{i+} = (p_{i1}, p_{i2}, \ldots, p_{iJ})'$ to be the vectors of J true and observed proportions, respectively, corresponding to the J classes for populations $i = 1, 2$. The hypothesis of homogeneity of the two populations is $H_0 : \pi_1 = \pi_2 = \pi_0$, against $H_a : \pi_1 \neq \pi_2$, where π_i are vectors with the first $(J-1)$ elements of π_{i+}, $i = 1, 2$; and π_0 is an unknown vector. In addition, define Z to be an observed class, Y the true class, X a predictor, S a population label. Let $P(Z = k|Y = j, X = x, S = i)$ equal the probability that a unit reports membership in class k conditional upon $Y = j$, $X = x$, and $S = i$. For convenience, we use the notation $P(Z = k|Y = j, X = x, S = i)$ and $P_i(Z = k|Y = j, X = x)$ interchangeably.

When misclassification errors exist, it can be important to determine the extent to which misclassification probabilities are homogeneous within specified groups. For this paper, we will say that misclassification probabilities are homogeneous within a population i if, for a given vector of explanatory variables x, $P_i(Z = k|Y = j, X = x)$ does not depend on x.

When misclassification probabilities are homogeneous, customary design based estimators of the proportions of reported classifications will converge to

$$p_{i+} = A_i' \pi_{i+} \tag{1}$$

where $A_i = [a_{i,jk}]$ is a $J \times J$ matrix with $(j,k)th$ element $a_{i,jk}$. The (j,k)th element of matrix A_i is the probability, denoted $P_i(Z = k|Y = j)$, of a unit being classified into the kth class when its true class is j.

Suppose now that there are categorical explanatory variables and that the intersection of all of the explanatory variable categories partitions the population i into C groups. Then, for group c and population i,

$$p_{ic+} = A'_{ic}\pi_{ic+} \tag{2}$$

where p_{ic+} is a vector of proportions of observed classification rates for group c in population i, π_{ic+} a vector of true proportions and A_{ic} is the associated misclassification matrix. More specifically, define $A_{ic} = [a_{ic,jk}]$ to be a $J \times J$ matrix with (j,k)th element $a_{ic,jk}$, where $a_{ic,jk} = P_{ic}(Z = k|Y = j)$ for group c and population i. The vector p_{ic+} is defined as

$$p_{ic+} = (M_{ic})^{-1}(\sum_{t\in U_{ic}} I_{t1}, \cdots, \sum_{t\in U_{ic}} I_{tJ})'$$

where U_{ic} is the subpopulation of persons in group c and population i, M_{ic} is the size of U_{ic} and I_{tj} is a dummy variable that equals one if a person gives answer j and zero otherwise. Similarly, the vector π_{ic+} is

$$\pi_{ic+} = (M_{ic})^{-1}(\sum_{t\in U_{ic}} \delta_{t1}, \cdots, \sum_{t\in U_{ic}} \delta_{tJ})'$$

where δ_{tj} equals one if a person's true category is j and zero otherwise. By this definition, the combined vector of observed proportions for population i is

$$p_{i+} = \sum_{c=1}^{C} R_{ic}A'_{ic}\pi_{ic+} \tag{3}$$

where $R_{ic} = M_i^{-1}M_{ic}$ and M_i is the number of units in population i. When $A_{i1} = \cdots = A_{iC} = A_i$, expression (3) is equal to $A'_i\pi_{i+}$ where $\pi_{i+} = M_i^{-1}(\sum_{t\in U_i} \delta_{t1}, \cdots, \sum_{t\in U_i} \delta_{tJ})'$ and U_i is the population i.

Assume now that all A_{ic} are all nonsingular matrices and that are not all equal. Let $B_{ic} = (A'_{ic})^{-1}$. Then from expressions (2) and (3)

$$\pi_{i+} = \sum_{c=1}^{C} R_{ic}B_{ic}p_{ic+} \tag{4}$$

and $B_{ic} = [b_{ic,jk}]$. When all A_{ic} are equal, expression (4) simplifies to $\pi_{i+} = (A'_i)^{-1}p_{i+}$ where $p_{i+} = M_i^{-1}(\sum_{t\in U_i} I_{t1}, \cdots, \sum_{t\in U_i} I_{tJ})'$.

3. Estimation of Cell Probabilities with Heterogeneous Misclassification Rates

3.1 Point Estimation

For population i, we assume the following design condition, quoted with minor modifications from Shao (1996, pp. 205-206).

(D.1) The sampling method follows a stratified multistage sampling design. The population has been stratified into L strata with N_h clusters in the hth stratum. For the hth stratum, $n_h \geq 2$ clusters are selected independently across the strata. These first-stage clusters are selected with unequal pre-draw probabilities, p_{hi}, and with replacement. Within the ith first-stage cluster in the hth stratum, $n_{hi} \geq 1$ ultimate units are sampled with selection probabilities p_{hij} from N_{hi} units, $j = 1, \cdots, n_{hi}$, $i = 1, \cdots, n_h$, $h = 1, \cdots, L$. The total number of ultimate units in the population is $N = \sum_{h=1}^{L}\sum_{i=1}^{N_h} N_{hi}$ and in the sample is $n = \sum_{h=1}^{L}\sum_{i=1}^{n_h} n_{hi}$.

For convenience, we will replace the triple subscript (hij) with the single subscript t in the following expressions if it is not necessary to specify strata, clusters and ultimate units. Under the design (D.1), let w_t be a unit-level survey weight. Then we have standard estimators of R_{ic} and p_{ic},

$$\hat{R}_{ic} = \hat{M}_i^{-1}\hat{M}_{ic} \tag{5}$$

where $\hat{M}_i = \sum_{t\in s_i} w_t$ and s_i is the set of sample units in population i; $\hat{M}_{ic} = \sum_{t\in s_{ic}} w_t$ and s_{ic} the set of sample units in group c within population i; and

$$\hat{p}_{ic+} = \hat{M}_{ic}^{-1}(\sum_{t\in s_{ic}} w_tI_{t1}, \cdots, \sum_{t\in s_{ic}} w_tI_{tJ})'. \tag{6}$$

Thus from expressions (5) and (6),

$$\hat{R}_{ic}\hat{p}_{ic+} = \hat{M}_i^{-1}(\sum_{t\in s_{ic}} w_tI_{t1}, \cdots, \sum_{t\in s_{ic}} w_tI_{tJ})' = \hat{e}_{ic}, \tag{7}$$

say. In addition, from expression (4) we have

$$\hat{\pi}_{i+} = \sum_{c=1}^{C} B_{ic}\hat{e}_{ic} \tag{8}$$

and the jth element of $\hat{\pi}_{i+}$ equals $\hat{\pi}_{ij} = \sum_{c=1}^{C} B_{icj\cdot}\hat{e}_{ic}$ where $B_{icj\cdot} = (b_{ic,j1}, \cdots, b_{ic,jJ})$ is the jth row of $J \times J$ matrix B_{ic}.

3.2 Variance Estimation

Assume that the matrices A_{ic} and thus B_{ic} are known. Define a $CJ \times 1$ vector $\hat{e}_{i\cdot} = (\hat{e}'_{i1}, \cdots, \hat{e}'_{iC})'$. Note that $\hat{e}_{i\cdot}$ is a customary vector of sample ratios. Consequently, we can use standard methods (as in, e.g., Shao, 1996) to compute a design-based estimator of the variance of the approximate distribution of $\hat{e}_{i\cdot}$, $\hat{V}(\hat{e}_{i\cdot})$, say.

Also, note that expression (8) can be written as

$$\hat{\pi}_{i+} = B_{i\cdots}\hat{e}_{i\cdot} \tag{9}$$

where $B_{i\cdots}$ is a $J \times CJ$ matrix with jth row equal to a $1 \times CJ$ vector $B_{i\cdot j\cdot} = (B_{i1j\cdot}, \cdots, B_{iCj\cdot})$. Thus, an estimator of the variance of the approximate distribution of $\hat{\pi}_{i+}$ is

$$\hat{V}(\hat{\pi}_{i+}) = B_{i\cdots}\hat{V}(\hat{e}_{i\cdot})B'_{i\cdots} \tag{10}$$

with jth diagonal element $\hat{V}(\hat{\pi}_{ij}) = B_{i\cdot j\cdot}\hat{V}(\hat{e}_{i\cdot})B'_{i\cdot j\cdot}$.

4. Logistic Regression-Based Estimation of Misclassification Matrices

To estimate $a_{ic,jk}$, the logistic regression method can be considered. A simple logistic regression model is,

$$g_i(x, D_j) = \beta_0 + \beta_1 D_j + \beta x \tag{11}$$

where $(\beta_0, \beta_1, \beta)$ is a fixed vector of coefficients, x is a vector of demographic or other auxiliary variables and $g_i(x, D_j) = ln[P_i(Z=k|Y=j, X=x)/\{1-P_i(Z=k|Y=j, X=x)\}]$. In addition, D_j is an indicator variable indicating true category membership, and equals one when a unit's true category is j and 0 otherwise. For the following the true status D_j will be assumed known from response to a second interviews.

When all x are categorical variables and they partition each population into C groups, model (11) indicates that the probability of misclassification $a_{ic,jk}$ of a unit that truly belongs to class j can be estimated by

$$\begin{aligned}&\hat{P}_{ic}(Z=k|Y=j, X=x)\\ &\quad= [1 + exp\{\hat{g}_i(x, D_j)\}]^{-1} exp\{\hat{g}_i(x, D_j)\},\end{aligned}$$

where $\hat{g}_i(x, D_j) = \hat{\beta}_0 + \hat{\beta}_1 D_j + \hat{\beta}x$ and $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta})$ is a consistent estimator of the vector $(\beta_0, \beta_1, \beta)$. Note that within group c and population i, all units in the sample have the same vector x. Thus, an estimator of A_{ic} is given by $\hat{A}_{ic} = [\hat{a}_{ic,jk}]$ where $\hat{a}_{ic,jk} = \hat{P}_{ic}(Z=k|Y=j, X=x)$.

5. Effect of Heterogeneous Misclassification Probabilities

When one considers heterogeneity of misclassification probabilities, the variances of adjusted estimators of cell proportions may be inflated due to the variability of A_{ic} within population i. If the bias of a biased test from incorrectly assuming their homogeneity is small relative to the amount of inflation in the point estimator variance that arise from accounting for heterogeneous misclassification probabilities, then one arguably might prefer the slightly biased test for some alternative hypothesis values. In this section, we will examine powers from unbiased and biased tests that do and do not account of heterogeneity of A_{ic}.

We will use the following condition which is quoted from Shao (1996, pp. 210-211).

(D.2) There is no survey weight that disproportionately large. That is

$$\max_{h,i,j}\{N^{-1}(n_{hi}w_{hij}n)\} = O(1).$$

Under design (D.1), condition (D.2) and additional regularity conditions, $\hat{e}_{i\cdot}$ in (7) is a consistent estimator of $R_{ic} = M_i^{-1}M_{ic}$; and $n_i^{1/2}(\hat{\pi}_i - \pi_i)$ converges in distribution to a $N_{J-1}(0, V_{\pi i})$ distribution where π_i and $\hat{\pi}_i$ are vectors with the first $(J-1)$ elements of π_{i+} and $\hat{\pi}_{i+}$ in expressions (4) and (9), respectively. For a Wald-type test, we will add the following condition.

(D.3) The matrix $n_i\{\hat{V}(\hat{\pi}_i)\}$ is a consistent estimator of $V_{\pi i}$ where $\hat{V}(\hat{\pi}_i)$ is the upper $(J-1) \times (J-1)$ submatrix of $\hat{V}(\hat{\pi}_{i+})$ in (10).

Then under design (D.1), conditions (D.2) and (D.3), and additional regularity conditions, the Wald test statistics for homogeneity, $H_0: \pi_1 = \pi_2 = \pi_0$,

$$X^2_{he} = (\hat{\pi}_1 - \hat{\pi}_2)'\hat{V}^{-1}(\hat{\pi}_1 - \hat{\pi}_2), \tag{12}$$

where $\hat{V} = \hat{V}(\hat{\pi}_1) + \hat{V}(\hat{\pi}_2)$, is asymptotically distributed as χ^2_{J-1}, a chi-square random variable on $(J-1)$ degrees of freedom under $H_0: \pi_1 = \pi_2 = \pi_0$ for sufficiently large n_i, $i = 1, 2$.

For any nonzero $D_\pi = \pi_1 - \pi_2$, the test statistic X^2_{he} is distributed asymptotically as $\chi^2_{J-1}(\lambda)$, a chi-square random variable on $(J-1)$ degrees of freedom with noncentrality parameter λ, where $\lambda = D'_\pi V^{-1} D_\pi/2$ and $V = V(\hat{\pi}_1) + V(\hat{\pi}_2)$. Thus the power of the Wald test in (12) is

$$\begin{aligned}1 - \beta_{he} &= Pr(X^2_{he} > \chi^2_{J-1,\alpha}|D_\pi)\\ &\doteq Pr(W_{J-1} > \chi^2_{J-1,\alpha}|D_\pi)\end{aligned}$$

where W_{J-1} is distributed as $\chi^2_{J-1}(\lambda)$ and $\chi^2_{J-1,\alpha}$ is the upper αth quantile of χ^2_{J-1}. When $H_0 : \pi_1 = \pi_2 = \pi_0$ is true, the Wald test in (12) achieves the nominal type I error rate α.

Now assume $A_{i1} = \cdots = A_{iC} = A_i$ and assume that A_i are known. Then for known A_i, the estimator of π_{i+} is

$$\hat{\pi}^*_{i+} = (A'_i)^{-1}\hat{p}_{i+} \quad (13)$$

where $\hat{p}_{i+}$ are observed probabilities. Its variance is estimated by

$$\hat{V}(\hat{\pi}^*_{i+}) = (A'_i)^{-1}\hat{V}(\hat{p}_{i+})A_i^{-1}$$

From a sample obtained by design (D.1), $\hat{p}_{i+} = \hat{M}_i^{-1}(\sum_{t\in s_i} w_t I_{t1}, \cdots, \sum_{t\in s_i} w_t I_{tJ})'$ for ith population, $i = 1, 2$. As with $\hat{e}_{i\cdot}$, $\hat{p}_{i+}$ is a vector of sample ratios and $\hat{V}(\hat{p}_{i+})$ is obtained by the same methods as $\hat{V}(\hat{e}_{i\cdot})$. Then the Wald test statistic for homogeneity, $H_0 : \pi_1 = \pi_2 = \pi_0$, is

$$X^2_{ho} = (\hat{\pi}^*_1 - \hat{\pi}^*_2)'(\hat{V}^*)^{-1}(\hat{\pi}^*_1 - \hat{\pi}^*_2), \quad (14)$$

where $\hat{V}^* = \hat{V}(\hat{\pi}^*_1) + \hat{V}(\hat{\pi}^*_2)$; $\hat{\pi}^*_i$ is a vector with the first $(J-1)$ elements of $\hat{\pi}^*_{i+}$; and $\hat{V}(\hat{\pi}^*_1)$ is a upper $(J-1) \times (J-1)$ submatrix of $\hat{V}(\hat{\pi}^*_{i+})$. The power of the Wald test in (14) is

$$\begin{aligned} 1 - \beta_{ho} &= Pr(X^2_{ho} > \chi^2_{J-1,\alpha}|D_\pi) \\ &\doteq Pr(W^*_{J-1} > \chi^2_{J-1,\alpha}|D_\pi) \end{aligned}$$

where W^*_{J-1} is distributed as $\chi^2_{J-1}(\lambda^*)$; $\lambda^* = (D_\pi - B)'(V^*)^{-1}(D_\pi - B)/2$; $V^* = V(\hat{\pi}^*_1) + V(\hat{\pi}^*_2)$; $B = b_1 - b_2$; and $b_i = \mathrm{E}(\hat{\pi}^*_i) - \pi_i$.

When misclassification probabilities are heterogeneous, b_i is not zero. Due to this bias in $\hat{\pi}^*_i$, the power $1 - \beta_{ho}$ under $H_0 : \pi_1 = \pi_2 = \pi_0$ may be different from the nominal type I error rate α and the Wald test statistic in (14) gives a biased test.

6. Application to Health Survey Data

6.1 Dual Frame NHIS/RDD Data

The National Health Interview Survey (NHIS) is a national level face-to-face survey carried out in all 50 states. For some applications, sample sizes were considered insufficient to evaluate state level estimates. The purpose of the Dual Frame NHIS/RDD Methodology and Field Test was to evaluate the feasibility of supplementing NHIS face-to-face interviews with RDD telephone interviews (Biemer, 1997). This study was conducted in two states, here labeled States A and B. These states were selected for the study due to their relatively large NHIS sample sizes (Biemer, 1997). In NHIS data, the initial interview was conducted face-to-face and the reinterview was conducted by telephone. For the RDD data, both interviews were conducted over the telephone.

From the questionnaire used for NHIS and RDD, we selected question $G1$, "Are any firearms now kept in or around your home?", with possible responses "yes" or "no". The hypothesis in which we are interested is $H_0 : P(G1 = \text{Yes}|\text{State A}) - P(G1 = \text{Yes}|\text{State B}) = 0$. We combined NHIS and RDD data; and for purposes of this analysis we considered the second interviews to give the true responses.

Table 1: Explanatory indicator variables for the logistic regression model.

Variable	Group Indicated
(Baseline Gender)	(Female respondent)
Male	Male respondent
(Baseline Mode)	(NHIS)
RDD	RDD
Fire2	Second Interview of G1=Yes
Fire2_RDD	Fire2×RDD
(Baseline Age)	(Age∈ [18, 39])
Age40	Age ≥ 40
Age40_Fire2	Age40×Fire2

Table 2: Logistic regression coefficient point estimates, standard errors, approximate 95% confidence intervals and p-values for $H_0 : \beta_i = 0$.

Predictor	$\hat{\beta}_i$	$se(\hat{\beta}_i)$	$(\hat{\beta}_{iL},\hat{\beta}_{iU})$
Constant	-4.0272	0.3062	(-4.6278,-3.4266)
Male	0.4805	0.1812	(0.1251,0.8359)
RDD	-0.5973	0.2796	(-1.1457,-0.0488)
Fire2	5.8095	0.3328	(5.1568,6.4622)
Fire2_RDD	1.4697	0.3843	(0.7159,2.2235)
Age40	1.5805	0.3137	(0.9654,2.1957)
Age40_Fire2	-1.6857	0.3676	(-2.4066,-0.9648)

6.2 Effect of Heterogeneous Misclassification Probabilities

To examine whether there are any auxiliary variables associated with probability of saying "yes" on question $G1$ on the second interview, we estimated the coefficients for the logistic regression model in (11). Some potentially important explanatory variables are a person's state of residence, gender, age and first interview modes; specific explanatory indicator variables are reported in Table 6.1. Exploratory analysis led to the final model coefficient estimates reported

Table 3: Estimates of cell proportions and their variances under heterogeneous misclassification probabilities.

Data	Point Estimate	State A	State B
NHIS	$\hat{\pi}_i$	0.50604215	0.27126436
	$\hat{V}(\hat{\pi}_i)$	0.00218004	0.0004723
Combined	$\hat{\pi}_i$	0.46793823	0.27091443
	$\hat{V}(\hat{\pi}_i)$	0.00071217	0.00026642

Table 4: Estimates of cell proportions and their variances under homogeneous misclassification probabilities.

Data	Point Estimate	State A	State B
NHIS	$\hat{\pi}_i^*$	0.48371183	0.28573574
	$\hat{V}(\hat{\pi}_i^*)$	0.00080075	0.00038814
Combined	$\hat{\pi}_i^*$	0.45726409	0.27793431
	$\hat{V}(\hat{\pi}_i^*)$	0.00034553	0.00023340

in Table 6.1. Based on Table (6.1), we constructed eight groups of respondents based on the combination of binary classification by gender, mode and Age40. For each group, estimates $\hat{A}_{ic}$ are obtained for both States A and B, $i = 1, 2$, $c = 1, \cdots, 8$, for the combined data. For the NHIS and RDD data, there are only four groups within each state, that is, $c = 1, \cdots, 4$. We considered the estimates $\hat{A}_{ic}$ in the evaluation of test powers. The estimator of variance of $\hat{e}_{i\cdot}$, $\hat{V}(\hat{e}_{i\cdot})$, was obtained by the linearization method (StataCorp, 1997, Reference P-Z, p. 418). Table 3 shows point estimates of $\hat{\pi}_i$ and $\hat{V}(\hat{\pi}_i)$ for the NHIS and combined data, respectively.

For homogeneous misclassification probabilities, each $a_{i,jk}$ in matrix A_i is estimated by

$$\hat{a}_{i,jk} = \hat{M}_{ij}^{-1} \sum_{t \in s_{ij}} w_t I_{tj}$$

where $\hat{M}_{ij} = \sum_{t \in s_{ij}} w_t$ and s_{ij} is the set of sample units in class j and population i. We consider these estimates $\hat{A}_i$ as known. The variance of $\hat{p}_i$ is estimated by the linearization method. Table 4 reports point estimates of $\hat{\pi}_i$ and $\hat{V}(\hat{\pi}_i)$ for the NHIS and combined data, respectively. The bias of $\hat{\pi}_1^* - \hat{\pi}_2^*$ is

$$B = b_1 - b_2 = \{\mathrm{E}(\hat{\pi}_1^*) - \mathrm{E}(\hat{\pi}_2^*)\} - (\pi_1 - \pi_2)$$

and is estimated by

$$\hat{B} = (\hat{\pi}_1^* - \hat{\pi}_2^*) - (\hat{\pi}_1 - \hat{\pi}_2)$$

since $\hat{\pi}_i$ is an unbiased estimator of π_i. These estimated biases are used to evaluate the power of a test based on an incorrect assumption of homogeneous misclassification probabilities when it is not true. Figure 1 shows powers from tests adjusted with homogeneous (dotted line) and with heterogeneous (solid line) misclassification probabilities. The upper panel shows powers from the NHIS data and the lower panel shows powers from the combined NHIS and RDD data. Both plots display a similar pattern. In both graph, the test based on assuming $A_{ic} = A_i$ appears to have a positive bias, and the type I error rate is inflated accordingly. On the other hand, the inflation of variance due to accounting for heterogeneity of misclassification probabilities is nontrivial relative to the biasedness caused by incorrectly assuming their equality. The loss of power due to accounting for heterogeneous misclassification probabilities appears to be more severe for the NHIS data. Thus, the loss of power attributable to adjustments for heterogeneity is of serious concern.

For the RDD data the difference between the two power curves is relatively small when it is compared to the NHIS and combined data, even though there is some positive biasedness exhibited when homogeneity is assumed.

References

Biemer, P. P. (1997). Dual frame NHIS/RDD methodology and field test. Analysis report, Research Triangle Institute, Research Triangle Park, NC.

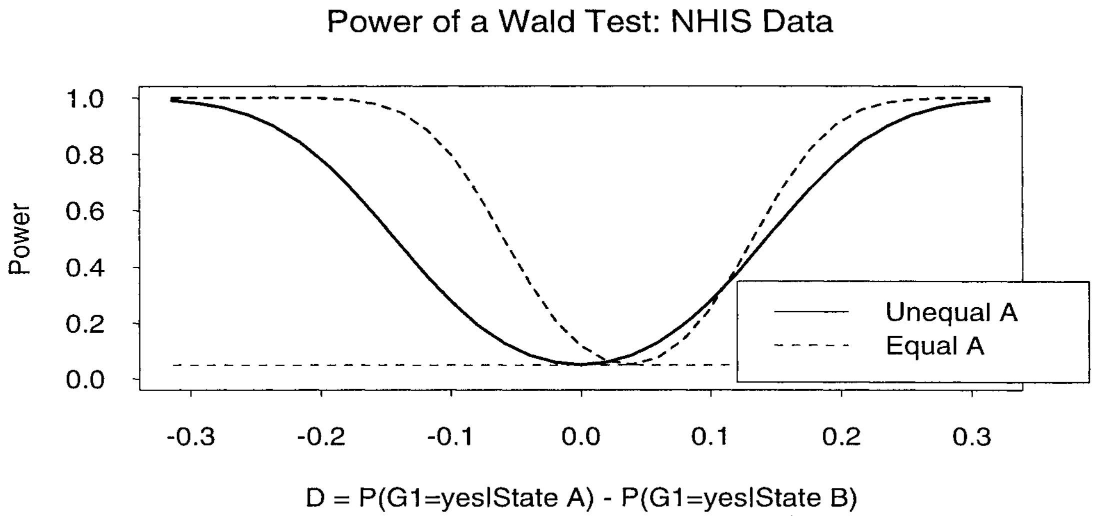
Heo, S. (1999). *Diagnostics for Survey Inference Accounting for Incomplete Data and Measurement Error.* Unpublished Ph.D. dissertation, Department of Statistics, Texas A&M University, College Station, TX.

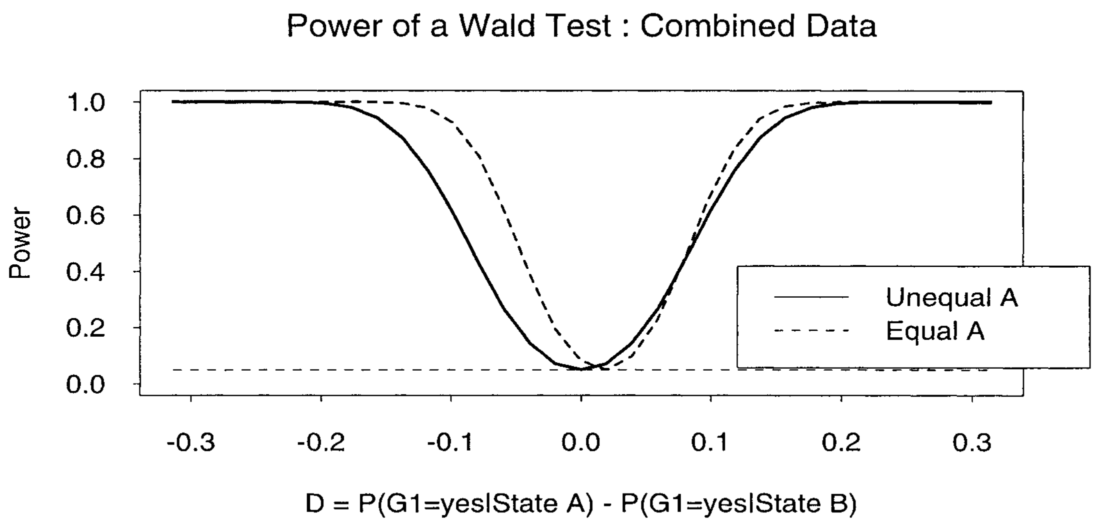
Hochberg, Y. and Tenenbein, A. (1983). On triple sampling schemes for estimating from binomial data with misclassification errors. *Communications in Statistics—Series A, Theory and Methods* **12**, 1523–1533.

Mote, V. L. and Anderson, R. L. (1965). An investigation of the effect of misclassification on the properties of chi-square tests in the analysis of categorical data. *Biometrika* **52**, 95–109.

Rao, J. N. K. and Thomas, D. R. (1991). Chi-squared tests with complex survey data subject to misclassification error. In P. P. Biemer, R. M.

Figure 1: Power of a Wald test statistic with one degree of freedom for the NHIS data (upper panel) and the combined RDD and NHIS data (lower panel), allowing for possibly unequal misclassification probabilities.





Groves, L. E. Lyberg, N. A. Mathiowetz, and S. Sudman (eds.), *Measurement Errors in Surveys*, pp. 637–663. New York: John Wiley & Sons.

Scott, A. J. and Rao, J. N. K. (1981). Chi-squared tests for contingency tables with proportions estimated from survey data. In D. Krewski, R. Platek, and J. N. K. Rao (eds.), *Current Topics in Survey Sampling*, pp. 247–265. New York: Academic Press.

Selén, J. (1986). Adjusting for errors in classification and measurement in the analysis of partly and purely categorical data. *Journal of the American Statistical Association* **81**, 75–81.

Shao, J. (1996). Resampling methods in sample surveys (with discussion). *Statistics* **27**, 203–254.

StataCorp (ed.). (1997). *Stata User's Guide, Release 5.* College Station, TX: Stata Press.

Tenenbein, A. (1972). A double sampling scheme for estimating from misclassified multinomial data with application to sampling inspection. *Technometrics* **14**, 187–202.