# An Alternative to the Logit-Wald Method for Inference under Models for Proportions

A.C. Singh[1], C.J.J.Nadeau[2]
[1]Statistics Canada, 16G, R.H. Coats, 100 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6
avi.singh@statcan.gc.ca
[2]Statistics Canada, 16F, R.H. Coats, 100 Tunney's Pasture Driveway, Ottawa, ON K1A 0T6

## Abstract

The quasi-likelihood estimating functions (EFs) properly weighted for survey data can be used to fit logit models. In practice, an asymmetric logit-Wald interval estimate (IE) is preferable to the usual Wald IE due to skewness in the distribution of estimates for low or high prevalence outcomes, and due to the possibility of the IE boundaries not being in the feasible range. However, IE may have poor coverage properties because of inadequate normality of estimated model parameters for small or moderate sample sizes (due to nonlinearity of estimates) which can be serious for low or high prevalence outcomes. We propose an alternative based on the Randomly Recentered Estimating Equations (RREE) methodology of Singh (2007) for the multiparameter case. In RREE, replicate parameter estimates are created by equating the vector of standardized EFs (or pivotals) to random vector values drawn from the standard multivariate normal distribution. These replicates are used to compute new point, variance, and interval estimates. Simulations are used to evaluate the performance of RREE relative to logit-Wald in terms of bias and variance of point and variance estimators, and average length and coverage probability of interval estimators.

**Key Words:** Estimating Functions; Confidence intervals; Variance Estimation; Survey data

## 1. Introduction

We consider the problem of modeling outcome prevalence for binary data using logit models. For example, for data from the Canadian Community Health Survey (CCHS), one may be interested in modeling the prevalence of healthy life style behavior outcomes such as smoking habit (a binary variable) using covariates such as age, gender, and education. The parameters of interest may be the conditional and marginal predictive means (Korn and Graubard, 1999, Ch. 3) or dose for a given mean response (McCullagh and Nelder, 1989, Ch. 4). A commonly used method is that of quasi-likelihood (*ql*) estimation (see e.g., McCullagh and Nelder, 1989, Ch 9) where the Taylor linearization is used to obtain a (sandwich-type) covariance matrix of model parameter estimates and Wald's method for normality-based interval estimates (IE). However, variance estimates (VE) for prevalence estimates may be unstable for small or moderate sample sizes as well as for low or high prevalence outcomes. Moreover, the IE may have poor coverage properties due to inadequate normality of estimated prevalence especially when the sample size is not large, the main reason being that the skewness in the distribution of estimated prevalence may be marked. As a result, the boundaries of the IE may not even be in the feasible range; i.e., 0 to 1. It would therefore be desirable to have asymmetric IE with appropriate coverage.

To overcome the above problem with IE, logit-Wald (see e.g., Newcombe, 2001) is commonly used where the IE is first constructed in the logit scale using the approximate normality of model parameter estimates under *ql*-estimation (this is the Wald step), and then boundaries of IE are transformed back to the original scale via the inverse logit transformation. The logit-Wald method tends to improve coverage but may be conservative for small or moderate sample sizes again due to inadequate normal approximation of model parameter estimates because of their nonlinearlity. In the case of complex surveys (as opposed to standard simple random sample case), above methods can be extended using weighted quasi-likelihood (*wql*) EFs of Godambe and Thompson (1986) for model parameters, similar to the approach in Binder (1983) for finite population quantities.

The purpose of this paper is to propose an alternative based on the multivariate version of the randomly recentered estimating equations (RREE) method of Singh (2007) in order to improve the performance of IE under *ql*-estimation based on logit-Wald and VE based on Taylor linearization. The RREE method basically consists of creating parameter estimates by solving the standardized EF vector centered at the random vector of values drawn from the pivotal normal distribution. The Monte Carlo distribution of parameter estimates so obtained is used to obtain new PE, VE, and IE. This is described in Section 3 following a background review in Section 2. Empirical results based on a simulation study to compare RREE with *ql*-estimation are presented in Section 4. Application to a complex survey data from the 2001 CCHS is given in Section 5. Here, the concept of generalized design effect (gdeff) (see e.g., Skinner, 1989) is used to adjust the covariance matrix for complex designs. Finally, Section 6 contains concluding remarks.

## 2. Background Review

For simplicity, we first consider data from a simple survey; i.e., the sample is simple random with replacement or it could be without replacement but with a negligible sampling fraction. For a random sample of *n* observations $i$=1,2,…,$n$, conditional on covariates $x_i$ 's consider a logit model given by

$$\begin{aligned} &y_i = \mu_i(\theta) + \varepsilon_i, \quad y_i \mid x_i \sim_{ind} Ber(\mu_i(\theta)) \\ &\text{logit}(\mu_i(\theta)) = x_i'\theta_{p\times 1} \end{aligned}, \qquad (2.1)$$

where $x_i$ is a *p*-vector of covariates, and the parameter $\theta$ is of dimension *p.* The *ql*-estimator of $\theta$ solves the equation obtained by setting the *ql*-EF vector $\psi_{ql(\theta)}$ to 0 where

$$\psi_{ql(\theta)} = (-\partial\mu(\theta)/\partial\theta')' V_\varepsilon^{-1}(y-\mu(\theta)) = \sum_{i=1}^{n} x_i(y_i - \mu_i(\theta)), \qquad (2.2)$$

$V_\varepsilon$ being the covariance matrix $\text{diag}(\mu_i(\theta)(1-\mu_i(\theta))_{1\le i\le n}$ of the observation errors $\varepsilon_i$ 's,. The PE $\hat{\theta}^{ql}$ can be obtained by solving (2.2) using an iterative method such as Newton-Raphson. The estimated covariance $\hat{\Sigma}_{ql(\theta)}$ of $\hat{\theta}^{ql}$ after Taylor linearization is obtained in a sandwich form as

$$\hat{\Sigma}_{ql(\theta)} = J_{ql(\theta)}^{-1} V_{ql(\psi)} J_{ql(\theta)}'^{-1}\Big|_{\hat{\theta}^{ql}}; \quad J_{ql(\theta)} = (-\partial\psi_{ql(\theta)}/\partial\theta')', \qquad (2.3)$$

where $J_{ql(\theta)}$ is the $p\times p$ observed *ql*-information matrix computed as $\sum_1^n u_i(\theta)x_i x_i'$ where $u_i(\theta) = \mu_i(\theta)(1-\mu_i(\theta))$ and $V_{ql(\psi)}$ is the $p\times p$ covariance matrix of the *ql*-EF vector $\psi_{ql(\theta)}$. Here $V_{ql(\psi)}$ coincides with $J_{ql(\theta)}$ because the link function is canonical, and the EF is optimal for the model.

Under regularity conditions, the estimator $\hat{\theta}^{ql}$ is consistent, and gives rise to a consistent estimator $\mu_i(\hat{\theta}^{ql})$ of the prevalence or the conditional mean given the covariate value. In this paper we will mainly be concerned with estimating the prevalence for a given covariate level. Using Taylor, a consistent variance estimator $v(\mu_i(\hat{\theta}^{ql}))$ for $\mu_i(\hat{\theta}^{ql})$ can be obtained as

$$v(\mu_i(\hat{\theta}^{ql})) = (\partial\mu_i(\theta)/\partial\theta')\Sigma_{ql(\theta)}(\partial\mu_i(\theta)/\partial\theta')'\Big|_{\hat{\theta}^{ql}}; \partial\mu_i(\theta)/\partial\theta' = u_i(\theta)x_i' \ . \qquad (2.4)$$

The VE $v(\mu_i(\hat{\theta}^{ql}))$ is expected to be unstable (i.e., with high relative variance) for small *n* or for low or high prevalence outcomes.

For IE under *ql*-estimation, we can use Wald or logit-Wald. The Wald IE for $\mu_i(\theta)$ is a normality based symmetric interval given by

$$\mu_i(\hat{\theta}^{ql}) \pm z_{\alpha/2}\sqrt{v(\mu_i(\hat{\theta}^{ql}))} \qquad (2.5)$$

which is expected to have undercoverage for small *n* or for low or high prevalence outcomes due to inadequate (i.e., unbalanced tails) normal approximation caused by nonlinearity of $\mu_i(\hat{\theta}^{ql})$. The logit-Wald provides an improved alternative with asymmetric IE such that interval boundaries lie in the admissible range of (0,1). To obtain logit-Wald IE, first a normal IE on the logit scale is constructed for $\mu_i(\theta)$, and then it is transformed back to the original scale of $\mu_i(\theta)$. For example, suppose has one x-variable resulting in two $\theta$-parameters (intercept and slope), we are interested in predicting $\mu(\theta, x=5)$ which equals $(1+\exp(\theta_0+5\theta_1))^{-1}\exp(\theta_0+5\theta_1)$. The Wald (symmetric) IE for $\theta_0+5\theta_1$ is given by (*L,U*) where the lower bound *L* is $(\hat{\theta}_0^{ql}+5\hat{\theta}_1^{ql})-z_{\alpha/2}\sqrt{v(\hat{\theta}_0^{ql}+5\hat{\theta}_1^{ql})}$ and the upper bound *U* is $(\hat{\theta}_0^{ql}+5\hat{\theta}_1^{ql})+z_{\alpha/2}\sqrt{v(\hat{\theta}_0^{ql}+5\hat{\theta}_1^{ql})}$. Next, the logit-Wald IE for $\mu(\theta, x=5)$ is given by $(L^*, U^*)$ where $L^*$ is $(1+\exp(L))^{-1}\exp(L)$ and $U^*$ is $(1+\exp(U))^{-1}\exp(U)$. Notice that $(L^*, U^*)$ is not symmetric but indeed lies inside (0,1). Although the logit-Wald IE under *ql*-estimation improves coverage, it tends to be conservative. In the next section, we propose an alternative based on RREE (Singh, 2007) for the multi-parameter case.

## 3. RREE: an Alternative to the Logit-Wald Method

In RREE, we start with a pivotal $H_{ql(\psi)}^{-1}\psi_{ql(\theta)}$ based on the *ql*-EF $\psi_{ql(\theta)}$ such that even for moderate *n,*

$$H_{ql(\psi)}^{-1}\psi_{ql(\theta)} \sim_{approx} N(0, I_{p\times p}), \tag{3.1}$$

where $H_{ql(\psi)}$ is the Cholesky root of the covariance matrix $V_{ql(\psi)}$; i.e., $V_{ql(\psi)} = H_{ql(\psi)}H'_{ql(\psi)}$. Under *ql*-estimation, $\hat{\theta}^{ql}$ is obtained by solving $H_{ql(\psi)}^{-1}\psi_{ql(\theta)}=0$, or $\psi_{ql(\theta)}=0$, while under RREE, a large number *R* of replicate parameter estimates $\{\tilde{\theta}_r\}_{1\le r\le R}$ are obtained by randomly recentering the estimating equations; i.e., by solving the *p*-equations

$$H_{ql(\psi)}^{-1}\psi_{ql(\theta)} = \varepsilon_r; \quad \varepsilon_r \sim_{iid} N_p(0, I) \quad . \tag{3.2}$$

The empirical distribution $\{\tilde{\theta}_r\}_{1\le r\le R}$ so obtained gives rise to new PE, VE, and IE for $\theta$, and any function of it; see Singh (2007) for theoretical details. In fact, in certain special cases, one can show analytically that both PE and VE based on RREE have higher relative biases than that under *ql*-estimation but with smaller relative MSE (mean square error); the bias decreasing with larger sample sizes as expected. However, it is IE where RREE has in general much improved finite sample property in terms of central and tail coverage probabilities; the reason being that the pivotal $H_{ql(\psi)}^{-1}\psi_{ql(\theta)}$ is expected to be closer to normal than the commonly used pivotal $H_{ql(\psi)}^{*-1}\psi_{ql(\theta)}$ under *ql*-estimation where $H_{ql(\psi)}^{*}$ is $H_{ql(\psi)}$ evaluated at $\hat{\theta}^{ql}$; see e.g., McCullagh (1991), and Godambe and Thompson (1999). For example, in the case of estimating a proportion $\theta$, the superior performance of the Wilson IE based on the pivotal $\sqrt{n}(\bar{y}-\theta)\big/\sqrt{\theta(1-\theta)}$ over the commonly used Wald IE based on the pivotal $\sqrt{n}(\bar{y}-\theta)\big/\sqrt{\bar{y}(1-\bar{y})}$ is well documented in Cai, Brown and Dasgupta (2001).

We remark that the empirical distribution $\{\tilde{\theta}_r\}_{1\le r\le R}$ is essentially invariant to any choice of the root of the covariance matrix $V_{ql(\psi)}$ such as the Cholesky root $H_{ql(\psi)}$ (already mentioned) or the one (denote by $\tilde{H}_{ql(\psi)}$) obtained from the spectral decomposition of $V_{ql(\psi)}$. This is so because the random recenters transformed by a root of $V_{ql(\psi)}$ has the same distribution; i.e., both $H_{ql(\psi)}\varepsilon_r$ and $\tilde{H}_{ql(\psi)}\varepsilon_r$ are distributed as $N_p(0, V_{ql(\psi)})$. Now, with RREE, it is possible to truncate the empirical distribution of $\hat{\theta}^{ql}$ (although the theoretical distribution is approximately normal) by discarding those recenters $\varepsilon_r$'s that give rise to nonfeasible or nonexistent solutions to the estimating equations. Moreover, before computing PE and VE, it is generally desirable in practice to trim extreme replicate values in $\{\tilde{\theta}_r\}_{1\le r\le R}$ for which at least one element of the *p*-vector lies outside the interval given by $\text{median} \pm 2.5(IQR)$; $IQR$ denoting the inter-quartile range. Such trimming helps to robustify PE and VE against extreme values. However, trimming is not needed for IE as it is generally not affected by extreme values. The empirical distribution $\{\tilde{\theta}_r\}_{1\le r\le R}$ gives rise to empirical distributions of other functions of $\hat{\theta}^{ql}$ such as the conditional predictive mean or the prevalence $\mu_i(\hat{\theta}^{ql})$ as $\{\mu_i(\tilde{\theta}_r)\}_{1\le r\le R}$ at the covariate value $x_i$.

For a simple example of a multi-parameter RREE, consider the logit model with a single covariate and the intercept. Then

$$\mu_i(\theta) = \exp(\theta_0 + \theta_1 x_i)[1 + \exp(\theta_0 + \theta_1 x_i)]^{-1} \qquad (3.3)$$

and the corresponding *ql-* EFs based on a random sample of *n* observations are

$$\psi_{ql(\theta_0)} = \sum_{i=1}^{n} (y_i - \mu_i(\theta)), \quad \psi_{ql(\theta_1)} = \sum_{i=1}^{n} x_i (y_i - \mu_i(\theta)) \,. \qquad (3.4)$$

The covariance matrix $V_{ql(\psi)}$ is given by

$$V_{ql(\psi)} = \begin{pmatrix} \sum_1^n u_i(\theta) & \sum_1^n x_i u_i(\theta) \\ \sum_1^n x_i u_i(\theta) & \sum_1^n x_i^2 u_i(\theta) \end{pmatrix}. \qquad (3.5)$$

Now, the replicate values of the vector parameter estimate $\hat{\theta}^{ql}$ are obtained by solving iteratively

$$\begin{pmatrix} \psi_{ql(\theta_0)} \\ \psi_{ql(\theta_1)} \end{pmatrix} = H_{ql(\psi)} \begin{pmatrix} \varepsilon_{1,r} \\ \varepsilon_{2,r} \end{pmatrix}, \qquad (3.6)$$

where $\{\varepsilon_{1,r}, \varepsilon_{2,r} : 1 \le r \le R\}$ are independent standard normal deviates. For computational convenience, one can first evaluate $H_{ql(\psi)}$ at the initial value $\hat{\theta}^{ql}$ to compute the next iterative value of $\theta$ serving as the current value. Next $H_{ql(\psi)}$ is evaluated at the current value of $\theta$, and the process is repeated until convergence to obtain $\tilde{\theta}_r$.

## 4. Simulation Study

A simulation study was conducted for a two parameter (intercept and slope) logit model for binary data where the covariate *x* for *i*=1, 2, …,*n* was defined as $x_i = \min\{1, (\mathrm{mod}(i,10) + 0.5)/10\}$; i.e., it takes values from .05 to .95 in increments of .10 and then repeats itself. The $\theta$-vector was chosen as $(-2.25, 3)'$ which gives rise to $\mu(\theta, x_i = 0.2)$ as $(1 + \exp(-1.65))^{-1} \exp(-1.65)$ or 16.11% as the chosen parameter of interest. The sample size *n* was set at 10, 20, 30, 50, and 100, the number *M* of simulation runs at 4000, and the number *R* of recenters for each simulation at 2000. Tables 1 and 2 compare results for three methods, QL (quasi-likelihood consisting of the solution of *ql*-EF at 0 for PE, Taylor linearization for VE, and Wald or logit Wald for IE), RREE* and RREE. Extreme replicate estimates were trimmed using the rule of median ± 2.5(IQR) for RREE methods. The abbreviation ME in Table 1 stands for MSE estimator which is the same as VE but the RB and RRMSE are computed by treating it as ME.

It is seen from Table 1 that the RB for PE does go up for RREE but it decreases considerably larger *n* as expected, but the corresponding RRMSE are quite comparable. However, for ME, although RREE shows more RB, its RRMSE is substantially less. In other words, the VE of RREE as a measure of MSE is considerably more stable than that for QL. The performance of RREE* relative to RREE is generally poor as expected. Turning to Table 2, QL-Logit Wald does improve considerably the coverage over QL-Wald for the two-sided interval, but it is conservative. The non-coverage for one-sided intervals is also improved under QL-Logit Wald but still poor with unbalanced tails. The RREE method, on the other hand, provides good coverage for the two-sided interval, and shows somewhat less imbalance in the tails in comparison to QL-Logit Wald, but its non-coverage for one-sided intervals is still not satisfactory.

Figures 1 and 2 depict in graphical form the information contained in Tables 1 and 2 but for all sample sizes considered in the simulation study. We note that differences between the various inferential procedures are less important as the sample size increases.

**Table 1: % Relative Bias and Relative Root MSE of PE and ME for $\mu(\theta, x = 0.2)$**

| Method | n=10 | | | | n=30 | | | |
|---|---|---|---|---|---|---|---|---|
| | PE | | ME | | PE | | ME | |
| | RB | RRMSE | RB | RRMSE | RB | RRMSE | RB | RRMSE |
| **QL** | 13.56 | 97.64 | 6.10 | 73.03 | 1.60 | 59.12 | -3.55 | 49.05 |
| **RREE*** | 30.33 | 97.15 | -20.43 | 64.39 | 4.27 | 57.48 | -8.96 | 52.96 |
| **RREE** | 33.53 | 97.55 | -15.53 | 49.95 | 11.51 | 59.34 | -12.65 | 42.84 |

**Table 2: 95% Coverage Probabilities of IE for $\mu(\theta, x = 0.2)$**

| Method | n=10 | | | | n=30 | | | |
|---|---|---|---|---|---|---|---|---|
| | L (2.5%) | C (95%) | R (2.5%) | Av. Len. (SD) | L (2.5%) | C (95%) | R (2.5%) | Av. Len. (SD) |
| **QL-Wald** | 1.53 | 81.50 | 16.97 | 0.5747 (0.2705) | 1.08 | 87.60 | 11.32 | 0.3502 (0.1086) |
| **QL- Logit Wald** | 3.03 | 96.97 | 0.00 | 0.7142 (0.0703) | 3.12 | 96.88 | 0.00 | 0.3801 (0.0576) |
| **RREE*** | 1.75 | 84.06 | 14.19 | 0.4681 (0.2341) | 0.95 | 88.95 | 10.10 | 0.3195 (0.1144) |
| **RREE** | 4.76 | 95.24 | 0.00 | 0.5579 (0.0882) | 3.65 | 95.20 | 1.15 | 0.3510 (0.0680) |

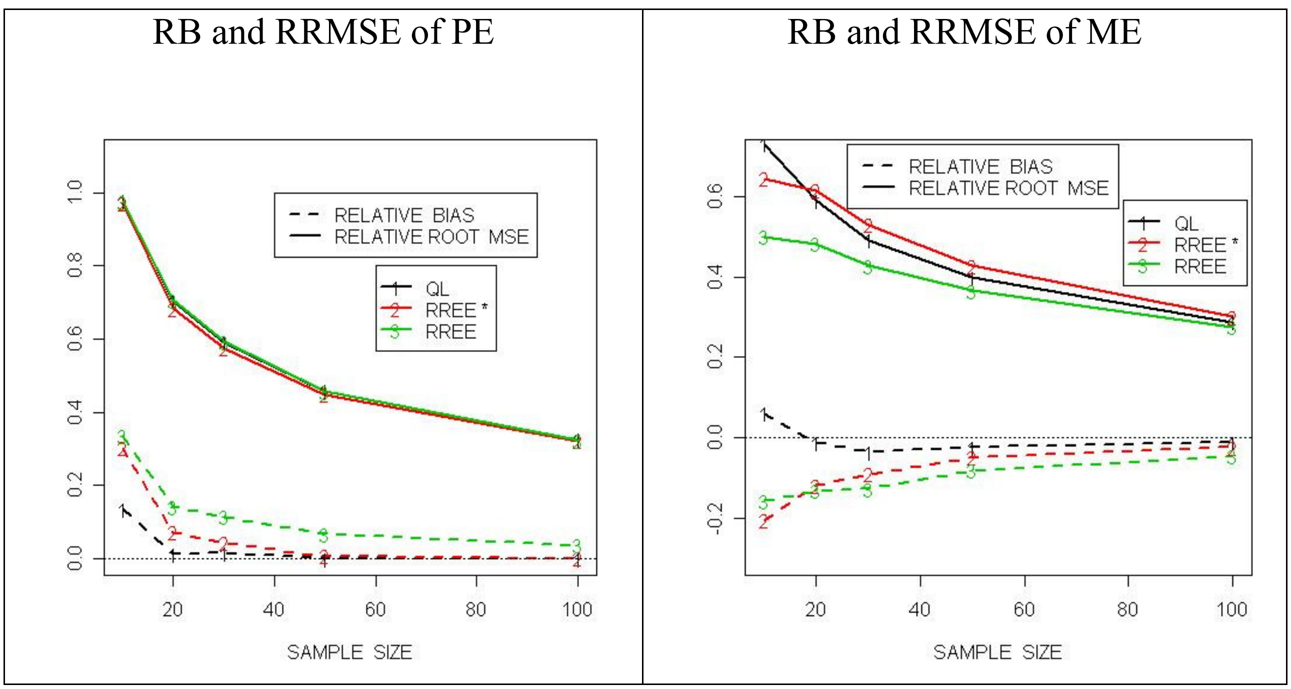


**Figure 1: % Relative Bias and Relative Root MSE of PE and ME for $\mu(\theta, x = 0.2)$**

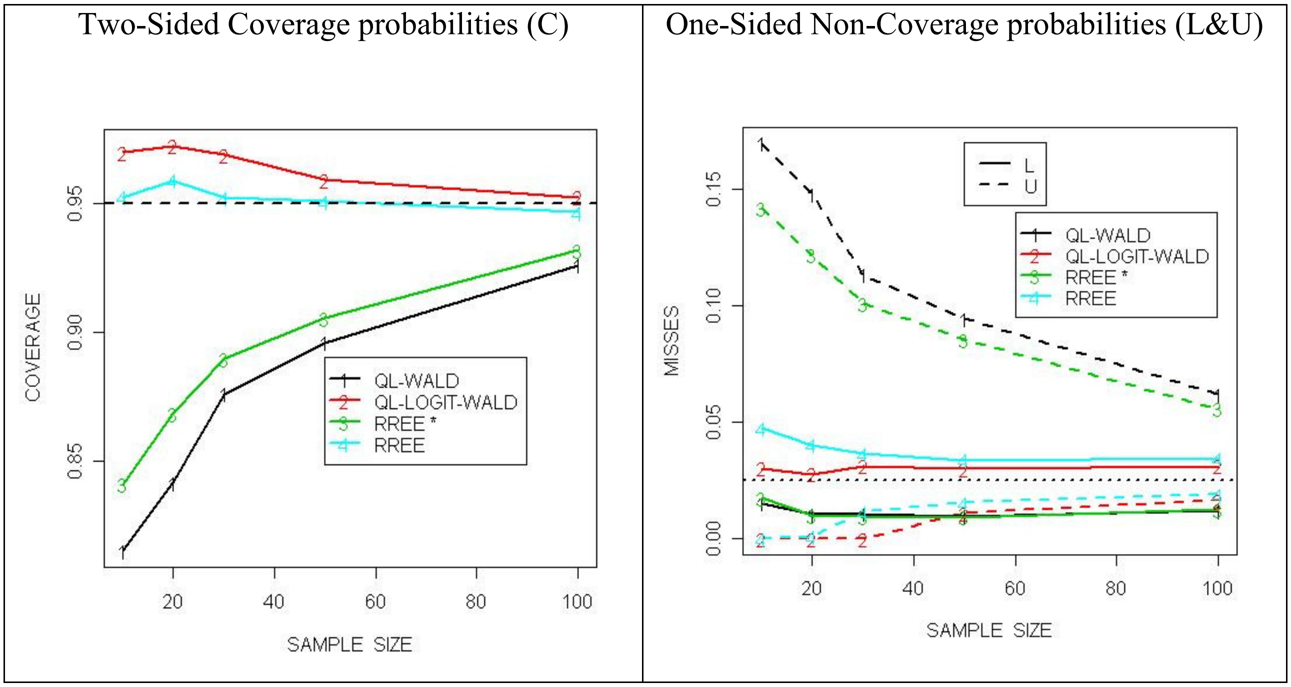


**Figure 2: 95% Coverage Probabilities of IE for *μ(θ, x = 0.2)***

## 5. Application to Survey Data

We illustrate an application of RREE to the data of cycle 1.1 of the Canadian Community Health Survey (CCHS) conducted in 2000-2001 whose goal was to collect general health information at the health Region level, a sub-provincial level of geography (Béland, 2002). The target population is all persons aged 12 years or older living in private dwellings in the ten provinces and three territories. The sample design is fairly complex involving stratified multi-stage cluster sampling. For the RREE application, suppose the parameter of interest is the proportion of smokers among 18 years old for the Yukon territory. For Yukon, the sample size was 809 while the population size was 24937. Using the optimal weighted quasi-likelihood (*wql*) method of Godambe and Thompson (1986), a two-parameter logit model with a single covariate of age (treated as a continuous variable measured in single years of age) was fit to this data. The *wql*-EFs for the parameters $(\theta_0, \theta_1)$ are given by

$$\psi_{wq(\theta_0)} = \sum_{i=1}^{n} w_i (y_i - \mu_i(\theta)), \quad \psi_{wq(\theta_1)} = \sum_{i=1}^{n} w_i x_i (y_i - \mu_i(\theta)), \tag{5.1}$$

where $w_i$'s denote the design weights. A consistent estimate $\hat{V}_{wq(\psi)}$ of the covariance matrix $V_{wq(\psi)}$ of the EF-vector $\psi_{wq(\theta)}$ can be obtained using standard survey sampling methods where the unknown $\theta$-parameter is evaluated at the *wql*-estimator $\hat{\theta}^{wql}$ computed as a solution of (5.1). Analogous to (2.3), the estimated covariance matrix $\hat{\Sigma}_{wq(\theta)}$ of $\hat{\theta}^{wql}$ is obtained after Taylor linearization in a sandwich form (Binder, 1983), and is given by

$$\hat{\Sigma}_{wq(\theta)} = J^{-1}_{wq(\theta)} V_{wq(\psi)} J'^{-1}_{wq(\theta)} \Big|_{\hat{\theta}^{wql}}; \quad J_{wq(\theta)} = (-\partial \psi_{wq(\theta)} / \partial \theta')', \tag{5.2}$$

where the observed *wq*-information matrix $J_{wq(\theta)}$ is similar to the matrix in (3.5) except that sampling weights $w_i$'s are appropriately inserted. Now the *wql*-Wald or logit Wald method for IE can be applied using the approximate distribution of $\hat{\theta}^{wql}$ as $N_p(0, \hat{\Sigma}_{wq(\theta)})$. To apply RREE to survey data, we need to express $V_{wq(\psi)}$ as a function of all the parameters under consideration which may not always be the case with $\hat{V}_{wq(\psi)}$ (as, for example, in the case of linear models). Moreover, the estimated covariance matrix $\hat{V}_{wq(\theta)}$ may not be stable for low or high prevalence outcomes.

To alleviate this problem, we can use a smoothed version of $V_{wq(\psi)}$ obtained as follows. Assuming that the design is ignorable for the model (2.1), a working covariance $V^*_{wq(\psi)}$ as an alternative to $V_{wq(\psi)}$, is obtained under the model but conditional on the selected units in the sample as

$$V^*_{wq(\psi)} = \begin{pmatrix} \sum_1^n w_i^2 u_i(\theta) & \sum_1^n w_i^2 x_i u_i(\theta) \\ \sum_1^n w_i^2 x_i u_i(\theta) & \sum_1^n w_i^2 x_i^2 u_i(\theta) \end{pmatrix}, \tag{5.3}$$

Now, the eigenvalues (or the generalized design effects; re: Skinner, 1989), $\lambda_1, \lambda_2$, of the design effect matrix $V^{*-1}_{wq(\psi)} V_{wq(\psi)} \big|_{\hat{\theta}^{wql}}$ are computed, and then following Singh, Folsom, and Vaish (2005), a smoothed version $\bar{V}_{wq(\psi)}$ of $V_{wq(\psi)}$ is obtained as $\bar{\lambda} V^*_{wq(\psi)}$ where $\bar{\lambda}$ denotes the average of eigenvalues. Note that such type of smoothing works well if the coefficient of variation of eigenvalues is small. In our application, $\lambda_1, \lambda_2$ are obtained as 0.85 and 1.52 which are not very close, and so $\bar{V}_{wq(\psi)}$ may not be a good smoothed approximation. Note also that the eigenvalues for smoothing are computed only once using the consistent estimator $\hat{\theta}^{wql}$ and not for each replicate estimate. Now RREE can be easily applied as before using the normal approximation $N_p(0, \bar{V}_{wq(\theta)})$ for the distribution of $\psi_{wq(\theta)}$.

Table 3 illustrates the results obtained under *wql*-estimation and RREE as well as RREE*. It is of interest to compare results under the wrong assumption of a simple design obtained by completely ignoring the sampling weights. The precision of PE and the width of IE under the simple design assumption are misleading in that the estimate appears more precise, and the interval narrower. Notice that results for different methods are very similar because the prevalence parameter of interest is not low and the sample size even after discounting for the design effect is quite large. Nevertheless, this example does illustrate the modifications needed for application to complex surveys.

**Table 3: PE, VE, and IE for $\mu(\theta, age = 18)$ for the 2001 CCHS Yukon Data**

| **Method** | **Simple Design Assumed** | | | **Complex Design of CCHS** | | |
|---|---|---|---|---|---|---|
| | **PE in %** | **VE x $10^4$** | **IE in%** | **PE in %** | **VE x $10^4$** | **IE in %** |
| **QL-Wald** | 28.95 | 6.91 | (23.79, 34.10) | 24.93 | 10.17 | (18.67, 31.18) |
| **QL- Logit Wald** | 28.95 | 6.91 | (24.08, 34.36) | 24.93 | 10.17 | (19.21, 31.68) |
| **RREE*** | 29.02 | 7.09 | (23.72, 34.30) | 25.02 | 9.76 | (18.81. 31.18) |
| **RREE** | 29.01 | 7.10 | (23.97, 34.32) | 25.04 | 9.68 | (19.24, 31.40) |

## 6. Concluding Remarks

In this paper, we considered the problem of poor coverage of Wald and logit-Wald methods for IE under *ql*-estimation for parameters of logit models for proportions. For small or moderate sample sizes, the Wald method for IE tends to give undercoverage for the two-sided interval while the logit-Wald method tends to yield overcoverage. The proposed alternative method of RREE is expected to improve coverage on theoretical grounds. Based on the limited simulation study, it was observed that RREE does improve considerably the coverage of two-sided intervals but only marginally for the one-sided intervals. However, for PE and VE, RREE for small sample sizes may introduce more bias in comparison to the *ql*-method, although the corresponding MSEs tend to be smaller especially for VE. An application of RREE to real data from the 2001 CCHS for the Yukon territory of Canada was presented. It was shown that after smoothing the covariance matrix of *wql*-EFs via generalized design effects, the application of RREE is essentially similar to that for the case of simple surveys; i.e., simple random samples. In practice, it is desirable to smooth the covariance matrix of *wql*-EFs for improved normal approximation because the design-based estimated covariance matrix for complex surveys may be unstable or may not be a function of all the parameters under consideration. It was observed that it is important to use a suitable pivotal to achieve superior performance of RREE. The RREE* method, in particular, uses the Wald-type pivotal, and exhibited poor coverage properties. Finally, we remark that there is a need to develop an efficient algorithm for RREE for models with many parameters because solving recentered estimating equations for high dimensional parameters would be computationally very tedious. We plan to investigate it in future.

## References

Béland, Y. (2002). Canadian Community Health Survey – Methodological overview. *Health Reports*, Vol 13, No.3, 9-14 (Statistics Canada, Catalogue no. 82-003).

Binder, D.A (1983). On the variances of asymptotically normal estimators from complex surveys. *Int. Statist. Rev.,* 51, 279-292.

Brown, L.D., Cai, T and DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statist. Sci.*, Vol 16, No.2, 101-133.

Godambe, V.P. and Thompson, M.E. (1986) Parameters of superpopulation and finite population: their relationship and estimation. *Int. Statist. Rev.*, 54, 37-59.

Godambe, V.P. and Thompson, M.E. (1999). A new look at confidence intervals in survey sampling. *Survey Methodology*, Vol. 25, No. 2, 161-174.

Korn, E.L. and Graubard, B.I. (1999). *Analysis of Health Surveys*. New York: John Wiley.

McCullagh, P. (1991). Quasi-likelihood and estimating functions. *Statistical Theory and Modeling:* In honor of Sir David Cox, FRS, ed. D.V. Hinkley, N. Reid, and E.J. Snell, Chapman and hall, London, 265-286.

McCullagh, P. and Nelder, J.A. (1989). *Generalized linear models*, 2nd Ed., London: Chapman and Hall.

Newcombe, R.G. (2001). Logit Confidence Intervals and the Inverse Sinh Transformation. Amer. Statist., Vol 55, No.3, 200-202.

Singh, A.C. (2007). A new application of estimating functions to point, variance, and interval estimation for simple and complex surveys. Proc. Fed. Comm. Statist. Meth., Washington, DC, www.fcsm.gov.

Singh, A.C., Folsom, R.E., Jr., and Vaish, A.K. (2005). Small area modeling for survey data with smoothed error covariance structure via generalized design effects. *Federal Committee on Statistical Methods Conference Proceedings*, Washington, DC. (www.fcsm.gov ).

Skinner, C.J. (1989). Introduction to Part A. In *Analysis of Complex Surveys* (C.J. Skinner, D. Holt, and T.M.F. Smith, Eds.), Ch. 2, Chichester: Wiley.