Survey Methodology, June 1998
Vol. 24, No. 1, pp. 51-55
Statistics Canada

Logistic Generalized Regression Estimators

RISTO LEHTONEN and ARI VEIJANEN[1]

ABSTRACT

In this paper we study the model-assisted estimation of class frequencies of a discrete response variable by a new survey estimation method, which is closely related to generalized regression estimation. In generalized regression estimation the available auxiliary data are incorporated in the estimation procedure by a linear model fit. Instead of using a linear model for the class indicators, we describe the joint distribution of the class indicators by a multinomial logistic model. Logistic generalized regression estimators are introduced for class frequencies in a population and domains. Monte Carlo experiments were carried out for simulated data and for real data taken from the Labour Force Survey conducted monthly by Statistics Finland. The logistic generalized regression estimation yielded better results than the ordinary regression estimation for small domains and particularly for small class frequencies.

KEY WORDS: Auxiliary information; Class frequencies; Generalized linear models; Labour force survey; Model-assisted estimation; Regression estimators.

1. INTRODUCTION

Consider the estimation of class frequencies of a discrete response variable in a sample survey. The number of individuals in a class equals the class indicator's sum over the population, the total of the indicator. Therefore, the problem can be solved by methods designed for the estimation of population totals. To improve the accuracy of the estimation, a survey statistician often makes use of the available auxiliary data. If the expectation of the response variable can be assumed to depend linearly on the auxiliary variables as can be the case for continuous response variables, it is advisable to use the generalized regression estimator (Särndal, Swensson and Wretman 1992; Estevao, Hidiroglou and Särndal 1995). Generalized regression estimation can improve the efficiency and reduce the bias due to unit nonresponse if the auxiliary variables correlate strongly with the response variable.

From a modeler's perspective, a linear model is quite restrictive and might not be the best choice for binary response variables, such as employment status of a person (employed, unemployed), or more generally for discrete response variables, such as a person's status in the labour market (employed, unemployed, not in labour force). For such variables we introduce a class of logistic generalized regression estimators based on a multinomial logistic model describing the joint distribution of the class indicators. The motivation for the selection of this specific model type thus is similar to that used in the context of generalized linear models (McCullagh and Nelder 1989).

The parameters of the logistic model are here estimated by maximizing a sample-based weighted loglikelihood, the Horvitz-Thompson estimator of the population loglikelihood function (Godambe and Thompson 1986; Nordberg 1989; Skinner, Holt and Smith 1989; Särndal *et al.* 1992, p. 517).

As an application, we consider the estimation of the unemployment rate in the Labour Force Survey conducted monthly by Statistics Finland. Administrative records indicating whether a person is registered jobseeker in local employment office are available as register-based auxiliary data, and these records were merged with the survey data on individual basis using personal identification numbers which are unique in both data sources. The corresponding auxiliary variable correlates strongly with the survey measurement on person's unemployment. Thus, improvement in efficiency and reduction of bias can be expected by making use of these administrative data in the estimation procedure. Additional auxiliary data (sex, age, regional data) were gathered from the Population Register. Also these auxiliary data were merged with the survey data on individual basis.

The properties of the generalized regression estimators were studied by Monte Carlo simulation methods where SRSWOR samples were repeatedly drawn from a population constructed from the Labour Force Survey data. We use incomplete poststratification or raking based on a main effects ANOVA model. The experiments indicate that the logistic formulation yields better results than the linear formulation for small domains. We obtained good results also when there was only one continuous auxiliary variable.

This paper is organized as follows. Section 2 defines the multinomial logistic model and basic concepts used. In Section 3 we introduce generalized regression estimators of class frequencies in a population and domains, and discuss the estimation of the model parameters by weighted loglikelihood. Variance estimators are presented. Monte Carlo experiments are discussed in Section 4. Conclusions are drawn in Section 5.

[1] Risto Lehtonen and Ari Veijanen, Statistics Finland, P.O. Box 5A, FIN-00022 Statistics Finland, Finland.

2. MODEL

Consider discrete m-valued random variables Y_k associated with N elements k in a finite population U. We observe their realized values y_k only in a sample $s \subset U$ of size n. Our goal is to estimate the frequency distribution of the y_k's in the population; in classification problems, we estimate the class proportions. Suppose we know the vector of auxiliary variables $\boldsymbol{x}_k$ for every element in the population. We impose a multinomial logistic model

$$P\{Y_k=i\} = \frac{\exp\{\boldsymbol{x}_k'\boldsymbol{\beta}_i\}}{\sum_{r=1}^{m}\exp\{\boldsymbol{x}_k'\boldsymbol{\beta}_r\}} \quad (i = 1, 2, ..., m) \qquad (1)$$

and assume that the Y_k's are conditionally independent given the $\boldsymbol{x}_k$'s. In the binary case, this is the model used in logistic regression. The parameter vector $\boldsymbol{\beta}$ is composed of vectors $\boldsymbol{\beta}_i\,(i = 1, 2, ..., m)$ with components $\beta_{ij}\,(j = 1, 2, ..., q)$. The parameters are assumed identifiable, that is, no two parameter values yield identical probabilities (1) for every k. This implies that the auxiliary variables $x_{kj}\,(j = 1, 2, ..., q)$ are linearly independent. To avoid identifiability problems, we set $\boldsymbol{\beta}_1 = 0$. It is straightforward to generalize (1) so that different auxiliary variables can be assigned for the m classes (Lehtonen and Veijanen 1998).

The sampling design specifies the inclusion probabilities of population elements. The k-th element is drawn with inclusion probability π_k and elements k and p are simultaneously in the sample s with probability $\pi_{kp} > 0$ $(\pi_{kk} = \pi_k)$. As usual, the sample membership indicators $I_k = I\{k \in s\}$ are assumed conditionally independent of the Y_k's given the $\boldsymbol{x}_k$'s, but the inclusion probabilities may correlate with the auxiliary variables.

Under unit nonresponse, if element k responds with probability θ_k independently of the I_p's and Y_p's $(p \in U)$, then we substitute $\pi_k\theta_k$ for π_k. Correspondingly, π_{kp} is replaced by $\pi_{kp}\,\theta_k\,\theta_p$ when the elements respond independently of each other.

3. LOGISTIC GENERALIZED REGRESSION ESTIMATION

3.1 Definition of LGREG

To estimate the frequency distribution of the y_k's, we define class indicators $Z_{ki} = I\{Y_k = i\}$ with realizations z_{ki} and estimate the totals $t_i = \sum_{k\in U} z_{ki}$. The Horvitz-Thompson (HT) estimator of t_i is $\hat{t}_i^{\text{HT}} = \sum_{k\in s} a_k z_{ki}$, where the sampling weights are $a_k = 1/\pi_k$. Generalized regression estimation (GREG) is assisted by a regression model $Z_{ki} = \boldsymbol{x}_k'\boldsymbol{\beta}_i^G + \varepsilon_{ki}$ with $\text{Var}(\varepsilon_{ki}) = \sigma_{ki}^2$ (Särndal *et al.* 1992; Estevao *et al.* 1995). The parameter $\boldsymbol{\beta}_i^G$ is estimated by

$$\hat{\boldsymbol{\beta}}_i^G = \left(\sum_{k\in s} a_k \frac{\boldsymbol{x}_k\boldsymbol{x}_k'}{\sigma_{ki}^2}\right)^{-1}\left(\sum_{k\in s} a_k \frac{\boldsymbol{x}_k z_{ki}}{\sigma_{ki}^2}\right) \quad (i = 1, 2, ..., m) \qquad (2)$$

and the fitted values $\hat{z}_{ki} = \boldsymbol{x}_k'\hat{\boldsymbol{\beta}}_i^G$ are incorporated in the GREG estimator

$$\hat{t}_i^G = \sum_{k\in U}\hat{z}_{ki} + \sum_{k\in s} a_k(z_{ki} - \hat{z}_{ki}) \quad (i = 1, 2, ..., m). \qquad (3)$$

The selection of a linear model for a GREG estimator (3) is fully justified for a continuous response variable. For binary measurements Z_{ki}, a linear model might be unrealistic. Ordinarily, we would prefer a logistic model to a linear one. In the logistic formulation, the predicted value always lies in [0,1], whereas in the linear formulation, the predicted value can exceed these natural limits. If the probability of $Z_{ki} = 1$ is close to 0 or 1, then the two models yield different results. Moreover, when there are $m > 2$ classes, it appears sensible to describe the joint distribution of the Z_{ki}'s $(i = 1, 2, ..., m)$ by the multinomial logistic model (1). To apply the model (1) in generalized regression estimation, we estimate the expectations $\mu_{ki} = E(Z_{ki}\,|\,\boldsymbol{x}_k; \boldsymbol{\beta}) = P\{Y_k = i\,|\,\boldsymbol{x}_k; \boldsymbol{\beta}\}$ by

$$\hat{\mu}_{ki} = P\{Y_k = i\,|\,\boldsymbol{x}_k; \hat{\boldsymbol{\beta}}\} = \frac{\exp\{\boldsymbol{x}_k'\hat{\boldsymbol{\beta}}_i\}}{1 + \sum_{r=2}^{m}\exp\{\boldsymbol{x}_k'\hat{\boldsymbol{\beta}}_r\}},$$

which depend nonlinearly on the auxiliary variables. We define a logistic generalized regression (LGREG) estimator by

$$\hat{t}_i = \sum_{k\in U}\hat{\mu}_{ki} + \sum_{k\in s} a_k(z_{ki} - \hat{\mu}_{ki}) \quad (i = 1, 2, ..., m). \qquad (4)$$

The GREG and LGREG estimators (3) and (4) include a sum of predicted values over the population. However, it is not actually necessary to have information about the $\boldsymbol{x}_k$'s for every element in the population U. In GREG (3), it is enough to know the auxiliary totals $\sum_{k\in U}\boldsymbol{x}_k$, because (3) can also be expressed in the form $\hat{t}_i^G = \hat{t}_i^{\text{HT}} + (\sum_{k\in U}\boldsymbol{x}_k - \sum_{k\in s} a_k\boldsymbol{x}_k)'\hat{\boldsymbol{\beta}}_i^G$. For the special case of complete poststratification, the information required in LGREG is similar to that needed in GREG. For other cases, such as incomplete poststratification, we cannot compute $\sum_{k\in U}\hat{\mu}_{ki}$ in (4) without knowing the frequency of each value of $\boldsymbol{x}_k$ in the population. For example, if we have two discrete auxiliary variables, then in GREG we need the marginal frequencies, but in LGREG we need the cell frequencies.

In addition to estimates for the whole population, estimates are usually calculated for subpopulations. The population U is partitioned into *domains* $U_{(d)} \subset U$ of size

$N_{(d)}$. The set s of respondents is composed of corresponding subsets $s_{(d)} = s \cap U_{(d)}$ with $n_{(d)}$ elements. As in GREG estimation (Särndal *et al.* 1992), we apply LGREG estimator

$$\hat{t}_{(d)i} = \sum_{k\in U_{(d)}} \hat{\mu}_{ki} + \sum_{k\in s_{(d)}} a_k (z_{ki} - \hat{\mu}_{ki}). \quad (5)$$

These estimators are additive: $\sum_i \hat{t}_{(d)i} = N_{(d)}$. If we combine two nonoverlapping domains d_1 and d_2, the LGREG estimate for $d = d_1 \cup d_2$ is $\hat{t}_{(d)i} = \hat{t}_{(d_1)i} + \hat{t}_{(d_2)i}$. Hence, $\sum_d \hat{t}_{(d)i} = \hat{t}_i$ for nonoverlapping domains and $\sum_i \hat{t}_i = N$.

In generalized regression estimation, an estimate (3) or (4) can be negative, when negative residuals coincide with large values of a_k. Negative GREG estimates become more common, as the number of auxiliary variables increases (Chambers 1996). In LGREG estimation, in contrast, this is not so, because $\hat{\mu}_{ki}$ is bounded by the model formulation. In our experiments, LGREG estimates were negative only for small domains in certain cases. In many cases, LGREG estimate equals the sum of estimated expectations and then it is always positive (see Section 3.2).

If the model (1) includes an auxiliary indicator variable, its total over the population is exactly estimated by LGREG. This calibration property is desirable in many applications.

3.2 ML Estimation by π-Weighted Loglikelihood

We estimate the parameter $\boldsymbol{\beta}$ in the model (1) by maximizing a π-weighted loglikelihood

$$L_s(\boldsymbol{\beta}_2, ..., \boldsymbol{\beta}_m) =$$

$$\sum_{k\in s} \pi_k^{-1} \left\{ I\{Y_k = 1\} \log \left(1 - \sum_{i=2}^{m} \mu_{ki} \right) + \sum_{i=2}^{m} I\{Y_k = i\} \log \mu_{ki} \right\}$$

(Godambe and Thompson 1986; Nordberg 1989; Särndal *et al.* 1992, p. 517). In general, we maximize the likelihood function numerically by appropriate numerical methods such as a Newton-Raphson algorithm.

It can be shown that for complete poststratification, the fitted values $\hat{z}_{ki}$ in GREG are equal to the estimates $\hat{\mu}_{ki}$ in LGREG. Thus, when there are no missing cells in complete poststratification, the GREG and LGREG estimators are identical (Lehtonen and Veijanen 1998). This does not hold for other models such as incomplete poststratification.

The LGREG estimator (4) has two parts: a sum of estimated expectations over the population and an adjustment term $\sum_{k\in s} a_k (z_{ki} - \hat{\mu}_{ki})$. It can be shown that if the model contains an intercept, the adjustment term vanishes and the frequency t_i is estimated by $\sum_{k\in U} \hat{\mu}_{ki}$ (Lehtonen and Veijanen 1998).

In our experiments, we apply a ratio estimator $\hat{R} = \hat{t}_i / (\hat{t}_i + \hat{t}_j)$. Its variance is estimated by Taylor linearization techniques (Särndal *et al.* 1992, p. 179):

$$\hat{V}(\hat{R}) = \frac{1}{(\hat{t}_i + \hat{t}_j)^2} \left[(1 - \hat{R})^2 \hat{C}_{ii} + 2\hat{R}(\hat{R} - 1)\hat{C}_{ij} + \hat{R}^2 \hat{C}_{jj} \right], \quad (6)$$

where C_{ij}, the covariance of $\hat{t}_i$ and $\hat{t}_j$, is estimated by

$$\hat{C}_{ij} = \sum_{k,p\in s} \frac{\Delta_{kp}}{\pi_{kp}} \frac{e_{ki}}{\pi_k} \frac{e_{pj}}{\pi_p}. \quad (7)$$

In (7), $e_{ki} = z_{ki} - \hat{\mu}_{ki}$ and $\Delta_{kp} = \mathrm{Cov}(I_k, I_p) = \pi_{kp} - \pi_k \pi_p$. Similar derivations hold for the corresponding domain estimators.

4. EXPERIMENTS

4.1 Details of Simulation Studies

In all the simulation experiments, $K = 1{,}000$ samples were drawn from a population with simple random sampling without replacement (SRSWOR). Monte Carlo means and standard errors of the estimates were calculated from the simulated samples. The design effect for an estimator $\hat{t}_{(d)i}$ was calculated as a ratio of estimated variances: $\mathrm{Deff}(\hat{t}_{(d)i}) = \hat{V}_{mc}(\hat{t}_{(d)i}) / \hat{V}_{mc}(\hat{t}_{(d)i}^{\mathrm{HT}})$, where $\hat{V}_{mc}(\hat{t}_{(d)i}^{\mathrm{HT}})$ denotes the Monte Carlo variance estimate of the HT estimator (Lehtonen and Pahkinen 1996). We measured the overall accuracy of domain estimates by the mean absolute relative domain error over D domains and K samples s_j:

$$\mathrm{MARDE}(i) = \frac{1}{D} \sum_{p=1}^{D} \frac{1}{K} \sum_{j=1}^{K} \frac{100 \left| \hat{t}_{(d_p)i}(s_j) - t_{(d_p)i} \right|}{t_{(d_p)i}}.$$

In the GREG estimates (2), the variance was a constant $\sigma_{ki}^2 = \sigma^2$, which cancelled out. For LGREG, domain frequencies were estimated by (5) and variances by (7). For GREG and HT, see Särndal *et al.* (1992, p. 401). Confidence intervals for the frequencies were computed as if the class indicators were independent. At the nominal significance level of 95%, an acceptable coverage rate lies in [93.65%, 96.35%] for $K = 1{,}000$ simulated samples.

4.2 An Experiment With Simulated Data

To compare LGREG with GREG, we simulated a data set, in which the auxiliary variable X was a continuous random variable uniformly distributed in (-3,3). The variable of interest, Y, representing three classes followed distribution (1) specified by $\boldsymbol{x}_k' \boldsymbol{\beta}_1 = 0$, $\boldsymbol{x}_k' \boldsymbol{\beta}_2 = 3X_k - 1$, and $\boldsymbol{x}_k' \boldsymbol{\beta}_3 = -2X_k$ for $N = 10{,}000$ elements ($k = 1, 2, ..., N$). A thousand samples of size n = 1,000 were independently drawn with SRSWOR. X_k and X_k^2 were used as auxiliary variables. All the estimators appeared unbiased (Table 1). The variance estimates had empirical bias smaller than 3% and standard deviation smaller than 5%.

Table 1
The design effects (Deff) for class frequency estimators and the empirical coverage rates (CR) (%) of 95% confidence intervals for classes $i = 1, 2, 3$

Method	Deff			CR		
	$\hat{t}_1$	$\hat{t}_2$	$\hat{t}_3$	$\hat{t}_1$	$\hat{t}_2$	$\hat{t}_3$
HT	1	1	1	95.2	95.3	94.7
GREG	0.93	0.55	0.57	95.0	94.3	95.6
LGREG	0.89	0.45	0.50	94.9	93.7	95.3

The best results were obtained by LGREG, probably due to the fact that the proportional frequencies of classes varied greatly over the range of the auxiliary variable. The probability of each class was such a function of the continuous auxiliary variable that a linear regression model did not fit the data well.

4.3 An Experiment With the Finnish Labour Force Survey Data

4.3.1 Constructed Population

We studied the estimation of the unemployment rate using the Finnish Labour Force Survey (LFS) data of three consecutive months of the year 1994. The constructed population consisted of 33,329 individuals. From the Population Register we obtained, for each population member, age class (15-24, 25-34, 35-44, 45-54, and 55-64 years), sex and region (three areas). A jobseeker indicator was obtained from the register maintained by Ministry of Labour showing which individuals were registered as unemployed jobseekers. The time lag in this administrative data source is about two weeks. It can thus be expected that the proportion of persons with changes in the actual labour market status is small within this short time interval. It should be noticed that the register-based jobseeker status is defined differently from the employment status measured in the Labour Force Survey. The survey measurement is based on a standard International Labour Office (ILO) definition. All these auxiliary data were merged with the survey data on individual basis.

The nonresponse rate varied by jobseeker status so that among registered jobseekers the rate was 11.4% whereas for the others the rate was 7.6%. The probability of nonresponse was modeled by a logistic ANOVA model and the ML estimates of nonresponse rates (ranging from 2.9% to 22.8%) were used as a nonresponse model in simulations.

For simulation experiments, we constructed an artificial population consisting of N = 30,835 persons. Employment status was defined by three classes: "employed", "unemployed", and "not in labour force" with population frequencies t_1 = 17,373, t_2 = 4,433, and t_3 = 9,029, respectively. The unemployment rate was defined by $R = t_2/(t_1 + t_2) = 20.33\%$. As domains we used the cells in the crosstabulation of age classes, sex, and the register-based unemployment status.

From the artificial population, K = 1,000 independent random samples of size n = 1,000 persons were drawn with simple random sampling without replacement. In each sample, nonresponse was simulated by the nonresponse model fitted to the original population. The response probabilities were then estimated from each sample by logistic regression with the same ANOVA model as in the nonresponse model. We multiplied each probability π_k by the estimated response probability.

Three models were used to compare LGREG with GREG. The components of $\boldsymbol{x}_k$ were dummies corresponding to age (5 classes), sex, region (3 areas) and jobseeker status. In incomplete poststratification, or raking, a main effects ANOVA model was based on classified auxiliary variables. We compared models with and without the jobseeker indicator. The third model also included a fourth-order polynomial of age.

4.3.2 Results

Incorporating no auxiliary information, HT estimators had usually larger variance than the generalized regression estimators (Table 2). Both generalized regression estimators based on a raking model with age, sex, and region yielded some improvement over the HT estimates. Much better results were obtained by models including the jobseeker indicator, which correlates more strongly (r = 0.83) with the ILO unemployment indicator than the other auxiliary variables. Thus these auxiliary data improve the efficiency of estimation (*cf.* Djerf 1997).

Table 2
Properties of unemployment rate estimates ($\hat{R}(\%)$) for the raking model (R) and the model including age polynomial (P), with (E) or without (N) the jobseeker indicator. SD denotes the standard deviation and CR (%) denotes the coverage rate of 95% confidence intervals

Model	Method	$\hat{R}$	Bias	SD	Deff	CR	MARDE
	HT	20.32	−0.0081	1.461	1	95.7	35.28
RN	GREG	20.30	−0.0262	1.454	0.995	95.3	46.03
RN	LGREG	20.31	−0.0229	1.454	0.995	95.3	45.93
RE	GREG	20.30	−0.0244	0.895	0.612	96.0	35.74
RE	LGREG	20.29	−0.0419	0.901	0.617	94.8	34.80
PE	GREG	20.30	−0.0259	0.887	0.607	95.6	35.41
PE	LGREG	20.29	−0.0421	0.896	0.613	95.1	34.76

Table 3
Mean absolute relative domain errors (MARDE) and mean coverage rates (CR) (%) of 95% confidence intervals for estimated class frequencies in domains with true frequency $t_{(d)i}$ (i = 1, 2, 3) (a) smaller than 100, and (b) at least 100. The model included the age polynomial

	Method	MARDE			CR		
		$\hat{t}_{(d)1}$	$\hat{t}_{(d)2}$	$\hat{t}_{(d)3}$	$\hat{t}_{(d)1}$	$\hat{t}_{(d)2}$	$\hat{t}_{(d)3}$
(a)	GREG	96.92	67.36	121.95	88.2	77.8	84.6
	LGREG	80.28	67.20	104.05	83.9	76.5	51.7
(b)	GREG	6.95	12.31	14.35	94.1	85.9	93.7
	LGREG	6.88	12.34	14.29	93.9	85.4	93.3

The differences between GREG and LGREG were small at the population level (Table 2). LGREG was never inferior to GREG. Domain totals, especially in small domains, were more accurately estimated by LGREG than by GREG (Table 3). When the model included the age as a continuous auxiliary variable, the standard deviation of the unemployment rate estimate was smaller for LGREG than for GREG in 19 of 20 domains. Unfortunately, the confidence intervals obtained by LGREG were often too narrow due to small variance estimates (Table 3).

5. SUMMARY

We introduce a new approach to the model-assisted estimation of population class frequencies of a discrete response variable in survey sampling. Our logistic generalized regression estimation (LGREG) is based on a multinomial logistic model, which might be more realistic for class indicators than the linear model normally used in generalized regression estimation (GREG). LGREG and GREG estimators yield identical results for complete poststratification, but differ for other models such as raking. As compared with GREG, LGREG usually requires more auxiliary information, not only the auxiliary totals. Nevertheless, LGREG appears preferable to GREG when the class probabilities vary greatly over the range of continuous auxiliary variables and when we need estimates for small domains, particularly in the presence of small class frequencies.

ACKNOWLEDGEMENTS

We are grateful to Prof. Carl-Erik Särndal, University of Montreal, for his comments on the previous version of the manuscript. Detailed comments given by an Associate Editor and two referees were very helpful. Thanks are also due to Mr. Timo Koskimäki, Statistics Finland, for providing us with the Labour Force Survey data and to Mr. Kari Djerf for helpful comments.

REFERENCES

CHAMBERS, R.L. (1996). Robust case-weighting for multipurpose establishment surveys. *Journal of Official Statistics*, 12, 3-32.

DJERF, K. (1997). Effects of post-stratification on the estimates of the Finnish Labour Force Survey. *Journal of Official Statistics*, 13, 29-39.

ESTEVAO, V., HIDIROGLOU, M.A., and SÄRNDAL, C.-E. (1995). Methodological principles for a generalized estimation system at Statistics Canada. *Journal of Official Statistics*, 11, 181-204.

GODAMBE, V.P., and THOMPSON, M.E. (1986). Parameters of superpopulation and survey population: their relationships and estimation. *International Statistical Review*, 54, 127-138.

LEHTONEN, R., and PAHKINEN, E.J. (1996). *Practical Methods for Design and Analysis of Complex Surveys*. Revised Edition. Chichester: John Wiley & Sons.

LEHTONEN, R., and VEIJANEN, A. (1998). On Multinomial Logistic Generalized Regression Estimators. Jyreãskylã. Preprints from the Department of Statistics, University of Jyreãskylã, 22.

McCULLAGH, P., and NELDER, J.A. (1989). *Generalized Linear Models*. Second Edition. London: Chapman and Hall.

NORDBERG, L. (1989). Generalized linear modeling of sample survey data. *Journal of Official Statistics*, 5, 223-239.

SKINNER, C.J., HOLT, D., and SMITH, T.M.F. (Eds) (1989). *Analysis of Complex Surveys*. New York: John Wiley & Sons.

SÄRNDAL, C.-E., SWENSSON, B., and WRETMAN, J.H. (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.