Parametric Estimation and Prediction under Informative Sampling and Nonignorable Nonresponse (in)

Abdulhakeem A.A. Eideh
Department of Mathematics
Al-Quds University
Abu-Dees Campus, Al-Quds, Palestine
E-mail: msabdul@staff.alquds.edu

ITACOSM 2019
UNIVERSITÀ DEGLI STUDI FIRENZE
FLORENCE 5-7 June 2019 | 6th ITAlian COnference on Survey Methodology

Thursday 6 June 2019
15:30-16:40
Room 209
Session: Nonsampling Error

Outline

- Introduction
- Response distribution
- Parametric Estimation under in
- Parametric Prediction under in
- Test of in
- Simple Ratio Population Model (in)
- Conclusions

Introduction

- **Informative sampling - the sample selection probabilities depend on the values of the model outcome variable (or the model outcome variable) is correlated with design variables not included in the model). See Pfeffermann et. al (1998).**

- **In addition to the effect of complex sample design, one of the major problems in the analysis of survey data is that of missing values or nonresponse.**

- **For inference problem, Little (1982) classify the nonresponse mechanism as ignorable and nonignorable.**

Cross-Classification on Sampling Design and Nonresponse Mechanism

Sampling Design	Nonresponse Mechanism	
	ignorable	nonignorable
informative	ii-Observed	in- Missing
noninformative	ni- Observed	nn-Observed

- **Brick (2013, JOS) …Thus, bias is often the largest component of mean square error of the estimates**

Response Distribution

- $U = \{1, \dots, N\}$ **- finite population**
- y_i **- value of** y **for the** $i\text{th}$ **population unit**
- x_i**,** $i \in U$ **- value** x **- auxiliary variable**
- $\mathbf{z} = \{z_1, \dots, z_N\}$ **values of known design variables**
- $\pi_i = \Pr(i \in s | x, y, z) > 0$
- $w_i = 1/\pi_i$ **sampling weight ;** $i = 1, \dots, N$
- $y_1, \dots, y_N$ **ind r.vs with pdf** $f_p(y_i \mid x_i, \theta)$
- $I_i = 1$ **if unit** $i \in U$ **is sampled** $I_i = 0$ **W.O**
- $s = \{i \mid i \in U, I_i = 1\}$ **- sample size** n
- $\bar{s} = \{i \mid i \in U, I_i = 0\}$ **- nonsampled unit**
- $R_i = 1$ **if unit** $i \in s$ **is observed** $R_i = 0$ **W.O**
- $r = \{i \in s \mid R_i = 1\}$**-response set**
- $\bar{r} = \{i \in s \mid R_i = 0\}$**- nonresponse set**
- $\psi_i = \Pr(i \in r \mid x, y, z^*) > 0$ **response probability**
- $\phi_i = 1/\psi_i$ **- response weight**

- **Sample distribution, Pfeffermann et al (1998)**

$$f_s(y_i \mid x_i,\theta,\gamma) = \frac{E_p(\pi_i \mid x_i, y_i, \gamma) f_p(y_i \mid x_i,\theta)}{E_p(\pi_i \mid x_i,\theta,\gamma)}$$

- **Nonsampled dist. Sverchkov and Pfeffermann (2004)**

$$f_{\bar{s}}(y_i \mid x_i,\theta,\gamma) = \frac{E_p(1-\pi_i \mid x_i, y_i, \gamma) f_p(y_i \mid x_i,\theta)}{E_p(1-\pi_i \mid x_i,\theta,\gamma)}$$

- **Response dist. Eideh (2002PhD, 2007):**

$$f_r(y_i \mid x_i,\theta,\eta,\gamma) = f_s(y_i \mid x_i,\theta,\eta,\gamma, i \in r)$$
$$= \frac{E_s(\psi_i \mid x_i, y_i, \gamma) E_p(\pi_i \mid x_i, y_i, \eta) f_p(y_i \mid x_i,\theta)}{E_s(\psi_i \mid x_i,\theta,\eta,\gamma) E_p(\pi_i \mid x_i,\theta,\eta)}$$

- **Nonresponse dist.**

$$f_{\bar{r}}(y_i \mid x_i) = \frac{\{1-E_s(\psi_i \mid x_i, y_i)\} f_s(y_i \mid x_i)}{\{1-E_s(\psi_i \mid x_i)\}}$$

Relationship between $E_p, E_s, E_{\bar{s}}, E_r, E_{\bar{r}}$

- **Pfeffermann and Sverchkov (1999)**

$$E_p(y_i \mid x_i) = \{E_s(w_i \mid x_i)\}^{-1} E_s(w_i y_i \mid x_i)$$

- **Sverchkov an Pfeffermann(2004)**

$$E_{\bar{s}}(y_i \mid x_i) = \frac{E_s\{(w_i - 1)y_i \mid x_i\}}{E_s\{(w_i - 1) \mid x_i\}}$$

- **Eideh (2002, 2007, 2016):**

$$E_p(y_i \mid x_i) = \frac{E_r(\phi_i w_i y_i \mid x_i)}{E_r(\phi_i w_i \mid x_i)}$$

$$E_{\bar{r}}(y_i \mid x_i) = \frac{E_r\{(\phi_i - 1)y_i \mid x_i\}}{E_r\{(\phi_i - 1) \mid x_i\}}$$

$$E_s(\psi_i \mid y_i) = \frac{1}{E_r(\phi_i \mid y_i)}$$

$$E_{\bar{s}}(y_i \mid x_i) = \frac{E_r\{\phi_i(w_i - 1)y_i \mid x_i\}}{E_r\{\phi_i(w_i - 1) \mid x_i\}}$$

- **Also, we can show the following:**

$$f_r(y_i \mid x_i) = \frac{E_r(\phi_i w_i \mid x_i)}{E_r(\phi_i w_i \mid x_i, y_i)} f_p(y_i \mid x_i)$$

$$f_{\bar{r}}(y_i \mid x_i) = \frac{E_r\{(\phi_i - 1) \mid x_i, y_i\}}{E_r\{(\phi_i - 1) \mid x_i\}} f_r(y_i \mid x_i)$$

$$f_{\bar{s}}(y_i \mid x_i) = \frac{E_r\{(\phi_i(w_i - 1)) \mid x_i, y_i\}}{E_r\{(\phi_i(w_i - 1)) \mid x_i\}} f_r(y_i \mid x_i)$$

- **Example:** **Method of Moments Estimator of Census Log-Likelihood**

$$U(\theta) = \partial l(\theta)/\partial\theta = \sum_{i=1}^{N} \partial \log f_p(y_i | x_i, \theta)/\partial\theta = 0$$

$$\hat{U}(\theta) = \sum_{i \in r} q_i^* \left\{ \frac{\partial \log f_p(y_i | x_i, \theta)}{\partial\theta} \right\} = 0$$

$$q_i^* = \frac{w_i \hat{\phi}_i}{E_r(w_i \hat{\phi}_i | x_i)}$$

Estimation of response weights $\phi_i = 1/\psi_i$ for all $i \in r$

- **If the nonresponse mechanism is NMAR, then values of y_i for $i \in r$ is available, but for $i \notin r$ are not available, so we cannot fit the following model:**

$$\psi_i(x_i, y_i, \gamma) = \Pr(R_i = 1 \mid i \in s, x_i, y_i)$$
$$= \frac{\exp(\gamma_0 + \gamma_1 x_i + \gamma_2 y_i)}{1 + \exp\exp(\gamma_0 + \gamma_1 x_i + \gamma_2 y_i)}$$

- **Sverchkov(2008)**

$$R_i : Bernoulli(\psi_i(x_i, y_i, \gamma))$$
$$f(r_i \mid x_i, y_i) = (\psi_i(x_i, y_i, \gamma))^{r_i} (1 - \psi_i(x_i, y_i, \gamma))^{1-r_i}$$

- **MLE γ:**

$$\frac{\partial l(\gamma)}{\partial \gamma} = \sum_{i \in r} \frac{\partial \log(\psi_i(x_i, y_i, \gamma))}{\partial \gamma} + \sum_{i \in \bar{r}} \frac{\partial \log(1 - \psi_i(x_i, y_i, \gamma))}{\partial \gamma} = 0$$

- **Using the missing information principle (see Cepillini et al 1955 and Orchard and Woodbury 1972). The idea is to take the expectation of likelihood equation with respect to the superpoulation model, and use the relationship between the moment under response and nonresponse distributions.**
- **Let** $O = O_s \bigcup O_r$ **be the available information, from the sample and response set**
- **So that, the observed log-likelihood equation is:**

$$\sum_{i \in r} \frac{\partial \log(\psi_i(x_i, y_i, \gamma))}{\partial \gamma} + \sum_{i \in \bar{r}} \frac{\left\{ E_r \left(\frac{[\phi_i(x_i, y_i, \gamma) - 1] \partial \log(1 - \psi_i(x_i, y_i, \gamma))}{\partial \gamma} \middle| O \right\} \right.}{E_r[(\phi_i(x_i, y_i, \gamma) - 1) | O]} = 0$$

- **Hence**

$$\hat{\psi}_i = \psi_i(\hat{\gamma}) = \psi_i(x_i, y_i, \hat{\gamma}) = \Pr(i \in r | x_i, y_i, \hat{\gamma}).$$

- **For more see Reddles, Kim (2016).**

Parametric Estimation under in

- **Assume response measurements are independent,**

$$L_{r,in}(\theta,\eta,\gamma)=\prod_{i=1}^{m} f_r(y_i \mid x_i,\theta,\eta,\gamma)$$

$$=\prod_{i=1}^{m}\frac{E_s(\psi_i \mid x_i,y_i,\gamma)E_p(\pi_i \mid x_i,y_i,\eta)f_p(y_i \mid x_i,\theta)}{E_s(\psi_i \mid x_i,\theta,\eta,\gamma)E_p(\pi_i \mid x_i,\theta,\eta)}$$

- **Weighted response likelihood - weights**
- **Estimation: Four steps method**
- **Step 1: Estimation of** ψ_i **:** $\hat{\phi}_i = 1/\hat{\psi}_i$
- **Step 2: Estimation of the effect of nonresponse mechanism**

$$E_s(\psi_i \mid x_i,y_i,\gamma)=\frac{1}{E_r(\phi_i \mid x_i,y_i,\gamma)}$$

- γ **can be estimated by regressing** $\hat{\phi}_i$ **on** (x_i,y_i) **using the data set** $\{\hat{\phi}_i, y_i, x_i,\ i\in r\}$**.**
- **Step 3:**

$$E_p(y_i \mid x_i,\eta)=\frac{E_r(\phi_i w_i y_i \mid x_i)}{E_r(\phi_i w_i \mid \mathbf{x}_i)}=E_r(l_i y_i | x_i)$$

$$l_i=\frac{\phi_i w_i}{E_r(\phi_i w_i \mid x_i)}$$

- η **can be estimated using regression analysis.**

- **Step 4: maximizing**

$$\tilde{l}_{r.in}(\theta)=\sum_{i=1}^{m}\log f_p(y_i \mid x_i,\theta)$$
$$-\sum_{i=1}^{m}\log E_p(\pi_i \mid x_i,\theta,\tilde{\eta})$$
$$-\sum_{i=1}^{m}\log E_s(\psi_i \mid x_i,\theta,\tilde{\eta},\tilde{\gamma})$$

Daniel Bonnéry, F. Jay Breidt, and François Coquet (2018):

Asymptotics for the maximum sample likelihood estimator under informative selection from a finite population

- **Abstarct: …While the sample likelihood methodology has been widely applied, its theoretical foundation has been less developed. A precise asymptotic description of a wide range of informative selection mechanisms is proposed.**
- **Under this framework, consistency and asymptotic normality of the maximum sample likelihood estimators are established. The theory allows for the possibility of nuisance parameters that describe the selection mechanism.**
- **The proposed regularity conditions are verifiable for various sample schemes, motivated by real problems in surveys. Simulation results for these examples illustrate the quality of the asymptotic approximations, and demonstrate a practical approach to variance estimation that combines aspects of model-based information theory and design-based variance estimation.**

New Test of Nonignorable Nonresponse

$$f_r(y_i \mid x_i) = \frac{E_r(\phi_i w_i \mid x_i)}{E_r(\phi_i w_i \mid x_i, y_i)} f_p(y_i \mid x_i)$$

1.Note that $f_r(y_i \mid x_i) \neq f_p(y_i \mid x_i)$ **unless**

$$E_r(\phi_i w_i \mid x_i, y_i) = E_r(\phi_i w_i \mid x_i)$$

- **Regress** $\hat{\phi}_i w_i$ **on** x_i, y_i
- **Test coefficient of** y_i **is zero**

2. Use <u>KL distance</u> between $f_r(y_i \mid x_i)$ and $f_p(y_i \mid x_i)$

- **<u>Suggestion:</u> New measure of representativeness (Call it Generalized measure) of a response set:**

  variance of $\hat{\phi}_i w_i$

- **Existing measure assume the sample is srs and MAR nonresponse mechanism.**
- **Furthermore! Regress** $\hat{\phi}_i$ **on** w_i **(or** w_i **and** x_i**) to predict** $\hat{\phi}_i$ **for** $i \in \bar{r}$ **and** $i \in \bar{s}$**, se we have** $\hat{\phi}_i$ $i \in U$ **!!!**
- **Bias reuction postaratification based on** $\hat{\phi}_i w_i$

Parametric Prediction of Finite Population Parameter under in

- **General Theory: Single-stage population model (two-stage in progress)**

$$T = \sum_{i=1}^{N} y_i = \sum_{i \in s} y_i + \sum_{i \in \bar{s}} y_i = \sum_{i \in r} y_i + \sum_{i \in \bar{r}} y_i + \sum_{i \in \bar{s}} y_i$$

- **For the prediction process we have the following available information:**

$$O = O_s \bigcup O_r$$

$$O_s = \left[\left\{ \left(x_i, I_i \right), i \in U \right\}, \left\{ \pi_i, R_i, i \in s \right\} \right]$$

$$O_r = \left[\left\{ \left(y_i, \hat{\psi}_i, x_i \right), i \in r \right\} \middle| i \in s \right], N, n, \text{and } m$$

- $\hat{T} = \hat{T}(O)$ **- predictor of** T **based on** O**.**
- **The mean square error (MSE)**

$$MSE_p\left(\hat{T}\right) = E_p\left\{ \left(\hat{T} - T \right)^2 \mid O \right\}$$
$$= \left\{ \hat{T} - E_p\left(T \mid O \right) \right\}^2 + Var_p\left(T \mid O \right)$$

- **is minimized when** $\hat{T} = E_P\left(T \mid O \right)$

- **Now we consider the following:**

$$
\begin{aligned}
T^* &= E_p(T \mid O) \\
&= \sum_{i\in r} y_i + \sum_{i\in \bar{r}} E_{\bar{r}}\left(y_i \middle| O\right) + \sum_{i\in \bar{s}} E_{\bar{s}}\left(y_i \middle| O\right)
\end{aligned}
$$

- **The empirical predictor for T:**

$$
\begin{aligned}
\hat{T} &= \hat{E}_p(T \mid O) \\
&= \sum_{i\in r} y_i + \sum_{i\in \bar{r}} \hat{E}_{\bar{r}}\left(y_i \middle| O\right) + \sum_{i\in \bar{s}} \hat{E}_{\bar{s}}\left(y_i \middle| O\right)
\end{aligned}
$$

- **Based on relationships between moments:**

$$
\begin{aligned}
T^* = \sum_{i\in r} y_i &+ \sum_{i\in \bar{r}} \frac{E_r\{(\phi_i - 1)y_i\}}{E_r\{(\phi_i - 1)\}} \\
&+ \sum_{i\in \bar{s}} \frac{E_r\{\phi_i(w_i - 1)y_i\}}{E_r\{\phi_i(w_i - 1)\}}
\end{aligned}
$$

- **Hence, T^* can be estimated based only on the data in the response set $\{y_i, \hat{\phi}_i, w_i : i \in r\}$.**

- **Using method of moments estimator, we can show that, the best linear unbiased predictor for *T* is:**

$$\hat{T}_{in}^{*} = \sum_{i \in r} y_i + (n-m)\frac{\sum_{i \in r}(\phi_i - 1)y_i}{\sum_{i \in r}(\phi_i - 1)} +$$

$$(N-n)\frac{\sum_{i \in r}\phi_i(w_i - 1)y_i}{\sum_{i \in r}\phi_i(w_i - 1)}$$

$$= \sum_{i \in r} w_i^{in} y_i$$

$$w_i^{in} = 1 + (n-m)\frac{(\phi_i - 1)}{\sum_{i \in r}(\phi_i - 1)} + (N-n)\frac{\phi_i(w_i - 1)}{\sum_{i \in r}\phi_i(w_i - 1)}$$

Note that

(a) $\sum_{i \in r}(\phi_i - 1)y_i$ **is the Horvitz Thompson estimator of** $\sum_{i \in \bar{r}} y_i$ **.**

(b) $\sum_{i \in r}\phi_i(w_i - 1)y_i$ **is the Horvitz Thompson estimator of** $\sum_{i \in \bar{s}} y_i$

(c) $\dfrac{n-m}{\sum_{i \in r}(\phi_i - 1)}$ **is the "Hajek type correction" for controlling the variability of the response weights.**

(d) $\dfrac{(N-n)}{\sum_{i\in r}\phi_i\left(w_i-1\right)}$ **is the "Hajek type correction" for controlling the variability of the product of response weights and sampling weights.**

- **It is easy to verify that**

(a) Under nn:

$$w_i^{nn}=1+\left(n-m\right)\frac{\left(\phi_i-1\right)}{\sum_{i\in r}\left(\phi_i-1\right)}+\left(N-n\right)\frac{\phi_i}{\sum_{i\in r}\phi_i}$$

(b) Under ni:

$$w_i^{ni}=1+\frac{\left(n-m\right)}{m}+\frac{\left(N-n\right)}{m}=\frac{N}{m}$$

(c) Under ii:

$$w_i^{ii}=1+\frac{\left(n-m\right)}{m}+\left(N-n\right)\frac{w_i-1}{\sum_{i\in r}\left(w_i-1\right)}$$

- **Also, we can show that**

$$T_{in}^{*}=\sum_{i\in r}y_i+\sum_{i\in\bar{r}}E_s\left(y_i\middle|O\right)+\sum_{i\in\bar{s}}E_p\left(y_i\middle|O\right)-$$

$$\left\{\sum_{i\in\bar{r}}\frac{Cov_s\left[\left(\psi_i,y_i\right)\middle|O\right]}{1-E_s\left(\psi_i\middle|O\right)}+\sum_{i\in\bar{s}}\frac{Cov_p\left[\left(\pi_i,y_i\right)\middle|O\right]}{1-E_p\left(\pi_i\middle|O\right)}\right\}$$

and

$$B\left(T_{in}^{*}\right)=E_p\left(T_{in}^{*}-T\right)$$

$$=-\left\{\begin{array}{l}\sum_{i\in\bar{r}}\left[\left(E_p\left(y_i\right)-E_s\left(y_i\right)\right)+\frac{Cov_s\left(\psi_i,y_i\right)}{E_s\left(1-\psi_i\right)}\right]+\\ \sum_{i\in\bar{s}}\frac{Cov_p\left(\pi_i,y_i\right)}{1-E_p\left(\pi_i\right)}\end{array}\right\}$$

Therefore, the predictor T_{in}^{*} is unbiased T if:

(a) $Cov_s(\psi_i, y_i) = 0$, **(or** $Cov_r(\phi_i, y_i) = 0$ **)**
that is, there is no correlation between the study variable and the response probabilities ψ_i **, consequently, nonresponse mechanism is ignorable, and**

(b)$Cov_p(\pi_i, y_i) = 0$,

(or $E_r(\phi_i)E_r(\phi_i w_i y_i) = E_r(\phi_i y_i)E_r(\phi_i w_i)$

that is, there is no correlation between the study variable and the first order inclusion probabilities π_i **, so sampling design is noninformative.**

If (a) is satisfied then (b) becomes $Cov_r(\phi_i w_i, y_i) = 0$

In other words, if the sampling design is noninformative and response mechanism is ignorable, so that $E_p(y_i) = E_{\bar{r}}(y_i) = E_{\bar{s}}(y_i)$, **then** T_{in}^* **in unbiased of** T **.**

Note that, the stronger the relationship between the study variable and the response probability, and the study variable and first order inclusion probabilities, the larger the bias.

- **We can show that**

$$B\left(T_{in}^{*}\right)=-\left\{\begin{array}{l}\sum_{i\in\bar{r}}\left\{\begin{array}{l}-\dfrac{Cov_r\left(\phi_i w_i,y_i\right)}{E_r\left(\phi_i w_i\right)E_r\left\{\left(\phi_i-1\right)\right\}}+\\ \dfrac{E_r\left(\phi_i w_i y_i\right)E_r\left(\phi_i\right)-E_r\left(\phi_i y_i\right)E_r\left(\phi_i w_i\right)}{E_r\left(\phi_i w_i\right)E_r\left\{\left(\phi_i-1\right)\right\}}\end{array}\right\}\\ -\sum_{i\in\bar{s}}\dfrac{E_r\left(\phi_i\right)E_r\left(\phi_i w_i y_i\right)-E_r\left(\phi_i y_i\right)E_r\left(\phi_i w_i\right)}{E_r\left(\phi_i w_i\right)\left[E_r\left(\phi_i w_i\right)-E_r\left(\phi_i\right)\right]}\end{array}\right\}$$

- **Hence, the bias $B\left(T_{in}^{*}\right)$ can be estimated based only on the data in the response set, $\left\{y_i,\phi_i,w_i : i\in r\right\}$, using method of moments estimates technique, that is, replace the moment under the response distribution by the average over the response set, for example $\hat{E}_r\left(a_i\right)=m^{-1}\sum_{i\in r}a_i$.**
- **To test the informativness of sampling design, see Pfeffermann and Sverchkov (1999) and Eideh and Nathan (2006). Moreover, for testing the ignorability of nonresponse mechanism, see Eideh (2012).**

Particular cases:

Case 1: nn

$$T_{nn}^{*}=\sum_{i\in r}y_i+\sum_{i\in\bar{r}}E_p\left(y_i\middle|O\right)+\sum_{i\in\bar{s}}E_p\left(y_i\middle|O\right)$$

$$-\sum_{i\in\bar{r}}\frac{Cov_p\left[\left(\psi_i,y_i\right)\middle|O\right]}{1-E_p\left(\psi_i\middle|O\right)}$$

$$B\left(T_{nn}^{*}\right)=-\sum_{i\in\bar{r}}\left\{\begin{array}{l}-\dfrac{Cov_r\left(\phi_i w_i,y_i\right)}{E_r\left(\phi_i w_i\right)E_r\left\{\left(\phi_i-1\right)\right\}}+\\ \dfrac{E_r\left(\phi_i w_i y_i\right)E_r\left(\phi_i\right)-E_r\left(\phi_i y_i\right)E_r\left(\phi_i w_i\right)}{E_r\left(\phi_i w_i\right)E_r\left\{\left(\phi_i-1\right)\right\}}\end{array}\right\}$$

Case 2: ni

$$T_{ni}^{*}=\sum_{i\in r}y_i+\sum_{i\in\bar{r}}E_p\left(y_i\middle|O\right)+\sum_{i\in\bar{s}}E_p\left(y_i\middle|O\right)$$

$$B\left(T_{ni}^{*}\right)=-\left\{\sum_{i\in\bar{r}}\left[E_p\left(y_i\right)-E_{\bar{r}}\left(y_i\right)\right]+\sum_{i\in\bar{s}}\left[E_p\left(y_i\right)-E_{\bar{s}}\left(y_i\right)\right]\right\}=0$$

Case 3: ii

$$T_{ii}^{*}=\sum_{i\in r}y_i+\sum_{i\in\bar{r}}E_s\left(y_i\middle|O\right)+\sum_{i\in\bar{s}}E_p\left(y_i\middle|O\right)-\sum_{i\in\bar{s}}\frac{Cov_p\left[\left(\pi_i,y_i\right)\middle|O\right]}{1-E_p\left(\pi_i\middle|O\right)}$$

$$B\left(T_{ii}^{*}\right)=-\left\{\begin{array}{l}\displaystyle\sum_{i\in\bar{r}}\frac{E_r\left(\phi_i\right)E_r\left(\phi_i w_i y_i\right)-E_r\left(\phi_i y_i\right)E_r\left(\phi_i w_i\right)}{E_r\left(\phi_i\right)E_r\left(\phi_i w_i\right)}-\\ \displaystyle\sum_{i\in\bar{s}}\frac{E_r\left(\phi_i\right)E_r\left(\phi_i w_i y_i\right)-E_r\left(\phi_i y_i\right)E_r\left(\phi_i w_i\right)}{E_r\left(\phi_i w_i\right)\left[E_r\left(\phi_i w_i\right)-E_r\left(\phi_i\right)\right]}\end{array}\right\}$$

Simple Ratio Population Model (in)

- $y_i | x_i \underset{p}{\sim} N(\beta x_i, \sigma^2 x_i), i = 1, \ldots, N$ **independent**

$$E_p(\pi_i | y_i, x_i) = \exp(\eta_0 x_i + \eta_1 y_i)$$

$$E_s(\psi_i | y_i, x_i) = \exp(\gamma_0 x_i + \gamma_1 y_i)$$

$$y_i \mid x_i \underset{s}{\sim} N((\eta_1 \sigma^2 + \beta) x_i, \sigma^2 x_i), i = 1, \ldots, n$$

$$y_i \mid x_i \underset{r}{\sim} N((\eta_1 \sigma^2 + \gamma_1 \sigma^2 + \beta) x_i, \sigma^2 x_i) i = 1, m$$

$$T^*_{in,R} = \sum_{i \in r} y_i + (\sum_{i \in U} x_i - \sum_{i \in r} x_i)\beta +$$
$$\sum_{i \in \bar{r}} \left\{ \eta_1 \sigma^2 x_i - (\gamma_1 \sigma^2 x_i) \frac{\exp(\gamma_0 x_i) M_s(\gamma_1)}{1 - \exp(\gamma_0 x_i) M_s(\gamma_1)} \right\} +$$
$$\sum_{i \in \bar{s}} \left\{ -(\eta_1 \sigma^2 x_i) \frac{\exp(\eta_0 x_i) M_p(\eta_1)}{1 - \exp(\eta_0 x_i) M_p(\eta_1)} \right\}$$

$$
\begin{aligned}
&B\left(T^*_{in,R}\right) = E_p\left(T^*_{in,R} - T\right)\\
&= -\left\{\begin{array}{l}\sum_{i\in\bar{r}} E_p\left[y_i - E_{\bar{r}}\left(y_i\right)\right] + \\ \sum_{i\in\bar{s}} E_p\left[y_i - E_{\bar{s}}\left(y_i\right)\right]\end{array}\right\}\\
&= -\left\{\begin{array}{l}\sum_{i\in\bar{r}}\left[-\left(\eta_1\sigma^2 x_i\right) + \left(\gamma_1\sigma^2 x_i\right)\frac{\exp\left(\gamma_0 x_i\right)M_s\left(\gamma_1\right)}{1-\exp\left(\gamma_0 x_i\right)M_s\left(\gamma_1\right)}\right] + \\ \sum_{i\in\bar{s}}\left[\left(\eta_1\sigma^2 x_i\right)\frac{\exp\left(\eta_0 x_i\right)M_p\left(\eta_1\right)}{1-\exp\left(\eta_0 x_i\right)M_p\left(\eta_1\right)}\right]\end{array}\right\}
\end{aligned}
$$

Particular cases:

Case 1: nn, ($\eta = 0$)

$$T^*_{nn,R} = \sum_{i \in r} y_i + (\sum_{i \in U} x_i - \sum_{i \in r} x_i)\beta +$$

$$\sum_{i \in \bar{r}} \left\{ -\left(\gamma_1 \sigma^2 x_i\right) \frac{\exp(\gamma_0 x_i) M_p(\gamma_1)}{1 - \exp(\gamma_0 x_i) M_p(\gamma_1)} \right\}$$

$$B\left(T^*_{nn,R}\right) = -\sum_{i \in \bar{r}} \left[\left(\gamma_1 \sigma^2 x_i\right) \frac{\exp(\gamma_0 x_i) M_p(\gamma_1)}{1 - \exp(\gamma_0 x_i) M_p(\gamma_1)} \right]$$

Case 2: nn ($\eta = 0$), ($\gamma = 0$):

$$\begin{aligned} T^*_{ni,R} &= \sum_{i \in r} y_i + \sum_{i \in \bar{r}} \beta x_i + \sum_{i \in \bar{s}} \beta x_i \\ &= \sum_{i \in r} y_i + (\sum_{i \in U} x_i - \sum_{i \in r} x_i)\beta \end{aligned} \quad , B\left(T_{ni,R}\right) = 0$$

Case 3: ii ($\gamma = 0$):

$$T^*_{ii,R} = \sum_{i \in r} y_i + (\sum_{i \in U} x_i - \sum_{i \in r} x_i)\beta +$$

$$\sum_{i \in \bar{s}} \left\{ -\left(\eta_1 \sigma^2 x_i\right) \frac{\exp(\eta_0 x_i) M_p(\eta_1)}{1 - \exp(\eta_0 x_i) M_p(\eta_1)} \right\}$$

$$B\left(T^*_{ii,R}\right) = -\left\{ \begin{array}{l} \sum_{i \in \bar{r}} \left[-\left(\eta_1 \sigma^2 x_i\right) \right] + \\ \sum_{i \in \bar{s}} \left[\left(\eta_1 \sigma^2 x_i\right) \frac{\exp(\eta_0 x_i) M_p(\eta_1)}{1 - \exp(\eta_0 x_i) M_p(\eta_1)} \right] \end{array} \right\} \tag{58}$$

Prediction of Census Likelihood

$$L_p(\theta) = \prod_{i=1}^{N} \log f_p(y_i \mid x_i, \theta)$$

$$l_p(\theta) = \log L_p(\theta) = \sum_{i \in r} \log f_p(y_i \mid x_i, \theta) +$$

$$\sum_{i \in \bar{r}} \log f_p(y_i \mid x_i, \theta) + \sum_{i \in \bar{s}} \log f_p(y_i \mid x_i, \theta)$$

MLE

$$\frac{\partial l_p(\theta)}{\partial \theta} = \sum_{i \in r} \frac{\partial \log f_p(y_i \mid x_i, \theta)}{\partial \theta} +$$

$$\sum_{i \in \bar{r}} \frac{\partial \log f_p(y_i \mid x_i, \theta)}{\partial \theta} + \sum_{i \in \bar{s}} \frac{\partial \log f_p(y_i \mid x_i, \theta)}{\partial \theta} = 0$$

But y_i not observed for $i \in \bar{r}, i \in \bar{s}$ and, so use missing information principle, we can show that MLE SATISFIES

$$\frac{\partial l_p(\theta)}{\partial \theta} = \sum_{i \in r} \frac{\partial \log f_p(y_i \mid x_i, \theta)}{\partial \theta} +$$

$$\sum_{i \in \bar{r}} E_{\bar{r}} \left(\frac{\partial \log f_p(y_i \mid x_i, \theta)}{\partial \theta} \middle| O \right) +$$

$$\sum_{i \in \bar{s}} E_{\bar{s}} \left(\frac{\partial \log f_p(y_i \mid x_i, \theta)}{\partial \theta} \middle| O \right) = 0$$

USING

$$E_{\bar{r}}(y_i \mid x_i) = \frac{E_r\{(\phi_i - 1) y_i \mid x_i\}}{E_r\{(\phi_i - 1) \mid x_i\}}$$

$$E_{\bar{s}}(y_i \mid x_i) = \frac{E_r\{\phi_i (w_i - 1) y_i \mid x_i\}}{E_r\{\phi_i (w_i - 1) \mid x_i\}}$$

$$\frac{\partial l_p(\theta)}{\partial \theta} = \sum_{i \in r} \frac{\partial \log f_p(y_i \mid x_i, \theta)}{\partial \theta} +$$

$$\sum_{i \in \bar{r}} \frac{E_r\left\{(\phi_i - 1)\frac{\partial \log f_p(y_i \mid x_i, \theta)}{\partial \theta} \mid x_i\right\}}{E_r\{(\phi_i - 1) \mid x_i\}} +$$

$$\sum_{i \in \bar{s}} \frac{E_r\left\{\phi_i(w_i - 1)\frac{\partial \log f_p(y_i \mid x_i, \theta)}{\partial \theta} \mid x_i\right\}}{E_r\{\phi_i(w_i - 1) \mid x_i\}} = 0$$

If we assume that $E_r\{\bullet \mid x_i\} \approx E_r\{\bullet\}$, we get

$$\sum_{i \in r} \phi_i^r \frac{\partial \log f_p(y_i \mid x_i, \theta)}{\partial \theta} = 0$$

$$\phi_i^r \approx 1 + (n - m)\frac{\hat{\phi}_i - 1}{\hat{\bar{\phi}} - 1} + (N - n)\frac{\hat{\phi}_i(w_i - 1)}{av(\widehat{\phi w}) - \hat{\bar{\phi}}}$$

Pseudo MLE with adjusted response weights ϕ_i^r

Conclusions

- **I hope that the new mathematical statistical results obtained will encourage further theoretical, empirical and practical research in these directions.**


STOP