COMBINING LINK-TRACING SAMPLING AND CLUSTER SAMPLING TO ESTIMATE THE SIZE OF A HIDDEN POPULATION: A BAYESIAN APPROACH

Martín H. Félix-Medina and Pedro Monjardin, Universidad Autónoma de Sinaloa
Martín H. Félix-Medina, Escuela de Ciencias Físico-Matemáticas,
Ciudad Universitaria, Culiacán Sinaloa, México, 80010 (mhfelix@uas.uasnet.mx)

Key Words: capture-recapture; design-based estimator; finite population.

1. Introduction

Link-tracing sampling (LTS) has been found appropriate for sampling hidden and hard-to-access human populations, such as drug-user, homeless-person, or illegal-worker populations. In this sampling method, an initial sample of people from the target population is selected, and the sample size is increased by asking the people in the initial sample to nominate other members of the population of interest. The nominated people might be asked to nominate other members of the population, and this process might continue until a specified stopping rule is satisfied. (See Spreen, 1992, and Thompson and Frank, 2000, for descriptions and reviews of different variants of this sampling methodology.)

Although LTS allows the sampler to make valid model-based inferences about the population size and other population parameters, the main problem of these inferences is that in practical applications the model assumptions are rarely satisfied. For instance, Frank and Snijders (1994) developed a variant of LTS in which is assumed the following premises: (i) the initial sample is a Bernoulli sample, that is, elements in the initial sample are included independently and with equal probability; and (ii) the nominations of members of the target population made by people in the initial sample are independent and equally probable.

The difficulty of satisfying these assumptions in practical situations motivated Félix-Medina and Thompson (2003) to develop a variant of LTS which does not require the premise of an initial Bernoulli sample. In their variant, those authors assume that a portion of the target population is covered by a sampling frame of accessible sites, such as bars, hospitals, city- blocks or public parks, where members of the target population can be found with high probability. A simple random sample of sites is selected, and the members of the population of interest that belong to each site are identified. Finally, the people in each site are asked to nominate other members of the population.

Using probability models that describe the number of elements found in each site and the probabilities of nomination, the authors derive maximum likelihood estimators (MLEs) of the population size. They also present model-based variance estimators based on the observed Fisher information matrix and design-based variance estimators. The results of a simulation study showed that the MLEs of the population size and their design-based variance estimators are robust to deviations from the assumed model, but that the model-based variance estimators are not robust. However, the main problem of the proposed MLEs is that they tend to overestimate seriously the population size if the nomination probabilities are small.

Those authors indicated that the problem of overestimation that appears when the nomination probabilities are small is caused by the little information in the sample which is not enough to obtain stable estimates of the nomination probabilities. The authors suggested that a possible solution to this problem is to construct estimators which incorporate information about the population parameters by using the Bayesian approach.

Many authors have used the Bayesian approach in Multiple capture-recapture sampling, which resembles the sampling design proposed by Félix-Medina and Thompson (2003). See Fienberg et al. (1999) for a review of capture-recapture sampling. In this work we will use the Bayesian model proposed by Castledine (1981) which allows constructing design-based variance estimators that are robust to deviations from the assumed models.

This research was supported by PROMEP, Grant UASIN-01-01.

2. Sampling design and notation

The sampling design considered in this paper is the same as that proposed by Félix- Medina and Thompson (2003). Therefore, we will suppose a finite hidden human population $U = \{u_1, \ldots, u_\tau\}$ of unknown size τ. In addition, we will suppose that a subset U_1 of an unknown number τ_1 of people of U can be found in different accessible sites, such as bars, public parks, or block-streets, and that a sampling frame of N of those sites can be constructed. In addition, we will assume that we can determine whether or not a person belongs to a site in the frame and, in the affirmative case, the site to which that person belongs to. (A person can belong to only one site.) We will denote by A_i the i-th site or cluster in the list and by m_i the number of people who belong to A_i, $i = 1, \ldots, N$. Notice that $\tau_1 = \sum_1^N m_i$. Finally, let $U_2 = U - U_1$ be the portion of U not covered by the frame and let $\tau_2 = \tau_- \tau_1$ be its size.

The sampling design is as follows. A sample $S_0 = \{A_1, \ldots, A_n\}$ of n clusters is selected by simple random sampling without replacement from the frame, and the m_i persons who belong to $A_i \in S_0$ are identified. Then, the people who belong to each cluster $A_i \in S_0$ are asked to nominate members in $U - A_i$. As a convention, we will say that a person is nominated by a cluster A_i if at least one member of that cluster nominates him or her. Nominations from different clusters are carried out independently, and different nomination strategies can be used in different sites. For instance, in site A_i the m_i members, as a group, could carried out the nominations; whereas in site A_j each of the m_j members could separately make the nominations. From each nominated person the following information is obtained: the sites that nominated him or her; and whether the person belong to a site in S_0, or to site in $U_1 - S_0$ (a nonsampled cluster), or to U_2 (the portion not covered by the frame).

The nomination of people by clusters will be indicated by the sets of variables $\{X_{ij}^{(1)}\}$ and $\{X_{ij}^{(2)}\}$, where $X_{ij}^{(1)} = 1$ if person $u_j \in U_1 - A_i$ is nominated by cluster A_i, and $X_{ij}^{(1)} = 0$ otherwise. Similarly, $X_{ij}^{(2)} = 1$ if person $u_j \in U_2$ is nominated by cluster A_i, and $X_{ij}^{(2)} = 0$ otherwise. Finally, by $\mathbf{y}$ will be denoted the set of counts y_ω, $\omega \subseteq \Omega = \{1, \ldots, n\}$ of the people who are nominated by every sampled cluster A_i with i in the set $\omega \neq \emptyset$, but not otherwise.

3. Estimators of the population size based on posterior modes

As in Félix-Medina and Thompson (2003), we will suppose that the sizes $m_1, \ldots, m_N$ of the clusters $A_1, \ldots, A_N$ are realizations of independent Poisson random variables $M_1, \ldots, M_N$ with mean λ; and that conditionally on the sizes $m_1, \ldots, m_n$ of the sampled clusters $A_1, \ldots, A_n$, the sizes τ_1 and τ_2, and the nomination probabilities $p_i^{(k)}$'s, $k = 1, 2$, the indicator random variables $X_{ij}^{(k)}$'s are independent Bernoulli random variables with means $p_i^{(k)}$'s, $k = 1, 2$, and $i = 1, \ldots, n$.

Then, under the previous assumptions,

$$f(\mathbf{y} \big| \mathbf{m}_s,\ \tau_1, \tau_2, \mathbf{p}_1, \mathbf{p}_2) \propto \frac{\tau_1!}{(\tau_1 - m - r_1)!} \left(1 - \frac{n}{N}\right)^{\tau_1} \times \prod_{i=1}^{n} \left[p_i^{(1)}\right]^{z_i^{(1)}} \left[1 - p_i^{(1)}\right]^{\tau_1 - m_i - z_i^{(1)}} \times \frac{\tau_2!}{(\tau_2 - r_2)!} \prod_{i=1}^{n} \left[p_i^{(2)}\right]^{z_i^{(2)}} \left[1 - p_i^{(2)}\right]^{\tau_2 - z_i^{(2)}},$$

where $\mathbf{m}_s = \{m_i\}_1^n$, $\mathbf{p}_1 = \{p_i^{(1)}\}_1^n$, $\mathbf{p}_2 = \{p_i^{(2)}\}_1^n$, $m = \sum_1^n m_i$ is the observed value of the random variable M that indicates the number of people in the initial sample S_0; $z_i^{(1)}$ and $z_i^{(2)}$ are the observed values of the random variables $Z_i^{(1)}$ and $Z_i^{(2)}$ that indicate the numbers of people in $U_1 - A_i$ and U_2, respectively, that are nominated by cluster A_i; and r_1 and r_2 are the observed values of the random variables R_1 and R_2 that indicate the numbers of distinct people in U_1 and U_2 that are nominated by at least one of the clusters in S_0.

We will now focus on the problem of defining the initial distributions of τ_1, τ_2, $\mathbf{p}_1$ and $\mathbf{p}_2$. In the case of τ_1 and τ_2, we will consider the following two models for the initial distributions:

Model I

$\pi(\tau_1|\lambda_1) \propto (N\lambda_1)^{\tau_1}/\tau_1!$ and $\pi(\lambda_1) \propto \lambda_1^{a_1-1}e^{-b_1\lambda_1}$,
$\pi(\tau_2|\lambda_2) \propto \lambda_2^{\tau_2}/\tau_2!$ and $\pi(\lambda_2) \propto \lambda_2^{a_2-1}e^{-b_2\lambda_2}$,

where a_1, b_1, a_2, b_2 are known constants, τ_1 and τ_2 are conditionally independent given λ_1 and λ_2, and λ_1 and λ_2 are also independent.

Model II

$f(\tau_k^{(l)}) \propto 1/\tau_k^{(l)}$, where $l = 0, 1$, $k = 1, 2$, and τ_1 and τ_2 are independent random variables.

The initial distribution for τ_1 defined in Model I is motivated by the fact that $\tau_1 = \sum_1^N M_i$, and that M_i is a Poisson variable with mean λ_1. Notice that Model I allows the researcher to incorporate information about τ_1 and τ_2 which is known prior to the observation of the sample. On the other hand, the distributions defined in Model II are not informative. When $l = 0$, τ_1 and τ_2 have improper uniform distributions, and when $l = 1$, these parameters have Jeffreys' distributions.

In the case of the nomination probabilities $p_i^{(k)}$'s, we will use a two-stage normal model proposed by Castledine (1981) for the logits $\alpha_i^{(k)} = \log[p_i^{(k)}/(1-p_i^{(k)})]$ of the probabilities $p_i^{(k)}$'s. Thus, we will consider the model:

$$\alpha_i^{(k)} \mid \theta_k \sim N(\theta_k, \sigma_k^2), \quad \text{and} \quad \theta_k \sim N(\mu_k, \gamma_k^2);$$

$i = 1\ldots, n$, $k = 1, 2$, where $N(\theta_k, \sigma_k^2)$ stands for the normal distribution with mean θ_k and variance σ_k^2; σ_k^2, μ_k and γ_k^2 are known constants; and the $\alpha_i^{(k)}$'s are conditionally independent given θ_k.

Finally, we will suppose that all the random variables τ_k, λ_k, $\alpha_i^{(k)}$ and θ_k, $k = 1, 2$, $i = 1, \ldots, n$, are independent.

3.1 Model I

In this case, the posterior joint distribution for τ_1, τ_2, $\{\alpha_i^{(1)}\}_1^n$, and $\{\alpha_i^{(2)}\}_1^n$ is

$$\begin{aligned} & f(\tau_1, \tau_2, \{\alpha_i^{(1)}\}, \{\alpha_i^{(2)}\} \Big| \text{data}) \propto h_1(\tau_1) \\ & \times \prod_{i=1}^{n} \frac{\exp[\alpha_i^{(1)} z_i^{(1)}]}{\left[1+\exp[\alpha_i^{(1)}]\right]^{\tau_1 - m_i}} \\ & \times \exp\left[-\frac{\sum_{i=1}^n \left(\alpha_i^{(1)} - \overline{\alpha}^{(1)}\right)^2}{2\sigma_1^2} - \frac{\left(\overline{\alpha}^{(1)} - \mu_1\right)^2}{2\upsilon_1}\right] \\ & \times h_2(\tau_2) \prod_{i=1}^{n} \frac{\exp[\alpha_i^{(2)} z_i^{(2)}]}{\left[1+\exp[\alpha_i^{(2)}]\right]^{\tau_2}} \\ & \times \exp\left[-\frac{\sum_{i=1}^n \left(\alpha_i^{(2)} - \overline{\alpha}^{(2)}\right)^2}{2\sigma_2^2} - \frac{\left(\overline{\alpha}^{(2)} - \mu_2\right)^2}{2\upsilon_2}\right], \end{aligned} \quad (1)$$

where

$$h_1(\tau_1) = \frac{(N-n)^{\tau_1}(\tau_1 + a_1 - 1)!}{(\tau_1 - m - r_1)!(N+b_1)^{\tau_1}},$$
$$h_2(\tau_2) = \frac{(\tau_2 + a_2 - 1)!}{(\tau_2 - r_2)!(b_2+1)^{\tau_2}},$$

$\overline{\alpha}^{(k)}$ is the arithmetic mean of the $\alpha_i^{(k)}$, and $\upsilon_k = \gamma_k^2 + \sigma_k^2/n$, $k = 1, 2$.

Since the analytical integration of (1) with respect to $\alpha_i^{(1)}$ and $\alpha_i^{(2)}$ seems to be not possible, we will not try to obtain analytical expressions for the posterior distributions of τ_1 and τ_2, but, as in Castledine (1981), we will use the mode of $f(\tau_1, \tau_2, \{\alpha_i^{(1)}\}, \{\alpha_i^{(2)}\} \Big| \text{data})$ as an estimator of $(\tau_1, \tau_2, \{\alpha_i^{(1)}\}, \{\alpha_i^{(2)}\})$. Using this strategy, we have that the proposed estimator is obtained as the solution to the following system of equations:

$$\hat{\tau}_1 = \frac{M + R_1 + (1 - n/N)\frac{N(a_1-1)}{N+b_1} \prod_{i=1}^{n} \left(1 - \hat{p}_i^{(1)}\right)}{1 - (1-n/N)\frac{N}{N+b_1} \prod_{i=1}^{n} \left(1 - \hat{p}_i^{(1)}\right)};$$

$$\begin{aligned} \hat{p}_i^{(1)} &= \frac{\exp\{\hat{\alpha}_i^{(1)}\}}{1+\exp\{\hat{\alpha}_i^{(1)}\}} \\ &= \frac{Z_i^{(1)}}{\hat{\tau}_1 - M_i} - \frac{\hat{\alpha}_i^{(1)} - \hat{\bar{\alpha}}^{(1)}}{(\hat{\tau}_1 - M_i)\sigma_1^2} - \frac{\hat{\bar{\alpha}}_i^{(1)} - \mu_1}{n(\hat{\tau}_1 - M_i)\nu_1}; \quad (2) \end{aligned}$$

$i = 1, \ldots, n;$

$$\hat{\tau}_2 = \frac{R_2 + \frac{a_2-1}{1+b_2} \prod_{i=1}^{n} \left(1 - \hat{p}_i^{(2)}\right)}{1 - \frac{1}{1+b_2} \prod_{i=1}^{n} \left(1 - \hat{p}_i^{(2)}\right)};$$

$$\begin{aligned} \hat{p}_i^{(2)} &= \frac{\exp\{\hat{\alpha}_i^{(2)}\}}{1+\exp\{\hat{\alpha}_i^{(2)}\}} \\ &= \frac{Z_i^{(2)}}{\hat{\tau}_2} - \frac{\hat{\alpha}_i^{(2)} - \hat{\bar{\alpha}}^{(2)}}{\hat{\tau}_2\sigma_2^2} - \frac{\hat{\bar{\alpha}}^{(2)} - \mu_2}{n\hat{\tau}_2\nu_2}; \quad (3) \end{aligned}$$

$i = 1, \ldots, n;$

where $\hat{\bar{\alpha}}^{(1)} = \sum_1^n \hat{\alpha}_i^{(1)}/n$ and $\hat{\bar{\alpha}}^{(2)} = \sum_1^n \hat{\alpha}_i^{(1)}/n$. From this, we have that an estimator of τ is $\hat{\tau} = \hat{\tau}_1 + \hat{\tau}_2$.

It is worth noting that the forms of these estimators are very similar to the forms of the estimators proposed by Félix-Medina and Thompson (2003). However, our estimators incorporate the initial information about the parameters τ_k and $\alpha_i^{(k)}$, $i = 1, \ldots, n$; $k = 1, 2$, and consequently they should surpass the performances of the estimators proposed by Félix-Medina and Thompson (2003).

To obtain estimators of the variances of $\hat{\tau}_1$, $\hat{\tau}_2$ and $\hat{\tau}$, which are robust to deviations from the assumed Poisson distribution for the M_i's, we will use the same strategy as that used by Félix-Medina and Thompson (2003). In this strategy, the distribution of the cluster sizes is not used, but it is replaced by the design-based distribution used to select the initial sample S_0. This is carried out by means of the formula:

$$\mathbf{V}(\hat{\tau}_k) = \mathbf{V}_p\left[\mathbf{E}_\xi(\hat{\tau}_k \Big| \mathbf{m}_s)\right] + \mathbf{E}_p\left[\mathbf{V}_\xi(\hat{\tau}_k \Big| \mathbf{m}_s)\right], \quad (4)$$

where $\mathbf{E}_\xi(\hat{\tau}_k \Big| \mathbf{m}_s)$ and $\mathbf{V}_\xi(\hat{\tau}_k \Big| \mathbf{m}_s)$ denote the conditional model-based expectation and variance operators, given that $\{M_i = m_i\}$, and $\mathbf{E}_p(\cdot)$ and $\mathbf{V}_p(\cdot)$ denote the design-based expectation and variance operators. Thus, the variance estimators are obtained by applying (4) to the first order Taylor's approximations of $\hat{\tau}_1$ and $\hat{\tau}_2$ about the model-based expectations of $c_s^{(1)} = \left(\{M_i\}, \{Z_i^{(1)}\}, R_1\right)$ and $c_s^{(2)} = \left(\{z_i^{(2)}\}, R_2\right)$, respectively.

By means of this strategy we obtain that an estimator of $\mathbf{V}_p\left[\mathbf{E}_\xi(\hat{\tau}_k \middle| \mathbf{m}_s)\right]$ is

$$\hat{\mathbf{V}}_{11} = n(1-n/N)\hat{K}^2 \frac{1}{n-1}\sum_{i=1}^{n}(m_i - \bar{m})^2, \qquad (5)$$

where $\bar{m} = n^{-1}\sum_1^n m_i$; $\hat{K} = -\hat{Q}_1/[\hat{A}_1(\hat{\tau}_1 - m - r_1)]$; $\hat{Q}_1 = \prod_{i=1}^n (1-\hat{p}_i^{(1)})$;

$$\hat{A}_1 = \sum_{i=1}^{n}\frac{(\hat{p}_i^{(1)})^2}{\hat{B}_i^{(1)}} - \hat{C}_1 + \frac{1}{\hat{\tau}_1 + a_1 - 1} - \frac{1}{\hat{\tau}_1 - m - r_1};$$

$$\hat{B}_i^{(1)} = (\hat{\tau}_1 - m_i)\hat{p}_i^{(1)}(1-\hat{p}_i^{(1)}) + \sigma_1^{-2},$$

$i = 1, \ldots, n$; and

$$\hat{C}_1 = \frac{(\nu_1^{-1} - n\sigma_1^{-2})\left[n^{-1}\sum_{i=1}^n \hat{p}_i^{(1)}/\hat{B}_i^{(1)}\right]^2}{1 + n^{-1}(\nu_1^{-1} - n\sigma_1^{-2})n^{-1}\sum_{i=1}^n 1/\hat{B}_i^{(1)}}.$$

An estimator of $\mathbf{E}_p\left[\mathbf{V}_\xi(\hat{\tau}_1 \middle| \mathbf{m}_s)\right]$ is

$$\hat{\mathbf{V}}_{12} = \hat{A}_1^{-2}\left\{\sum_{i=1}^{n}\left(\frac{\hat{p}_i^{(1)}}{\hat{B}_i^{(1)}} - \frac{\hat{D}_1}{\hat{B}_i^{(1)}}\right)^2 (\hat{\tau}_1 - m_i)\hat{p}_i^{(1)} \right.$$
$$\times(1-\hat{p}_i^{(1)}) + \frac{\hat{\tau}_1 - m)\hat{Q}_1(1-\hat{Q}_1)}{(\hat{\tau}_1 - m - r_1)^2}$$
$$\left. -\frac{2(\hat{\tau}_1 - m)\hat{Q}_1}{\hat{\tau}_1 - m - r_1}\sum_{i=1}^{n}\left(\frac{\hat{p}_i^{(1)}}{\hat{B}_i^{(1)}} - \frac{\hat{D}_1}{\hat{B}_i^{(1)}}\right)\hat{p}_i^{(1)}\right\}, \qquad (6)$$

where

$$\hat{D}_1 = \frac{n^{-1}(\nu_1^{-1} - n\sigma_1^{-2})n^{-1}\sum_{i=1}^n \hat{p}_i^{(1)}/\hat{B}_i^{(1)}}{1 + n^{-1}(\nu_1^{-1} - n\sigma_1^{-2})n^{-1}\sum_{i=1}^n 1/\hat{B}_i^{(1)}}.$$

Therefore, a design-based estimator of $\mathbf{V}(\hat{\tau}_1)$ is $\hat{\mathbf{V}}(\hat{\tau}_1) = \hat{\mathbf{V}}_{11} + \hat{\mathbf{V}}_{12}$.

In the case of $\hat{\tau}_2$, since $\mathbf{E}_\xi(\hat{\tau}_2 \middle| \mathbf{m}_s) \approx \tau_2$, it follows that $\mathbf{V}_p\left[\mathbf{E}_\xi(\hat{\tau}_2 \middle| \mathbf{m}_s)\right] \approx 0$. Therefore, an estimator of $\mathbf{V}(\hat{\tau}_2)$ is

$$\hat{\mathbf{V}}(\hat{\tau}_2) = \hat{A}_2^{-2}\left\{\sum_{i=1}^{n}\left(\frac{\hat{p}_i^{(2)}}{\hat{B}_i^{(2)}} - \frac{\hat{D}_2}{\hat{B}_i^{(2)}}\right)^2 \hat{\tau}_2\hat{p}_i^{(2)} \right.$$
$$\times(1-\hat{p}_i^{(2)}) + \frac{\hat{\tau}_2\hat{Q}_2(1-\hat{Q}_2)}{(\hat{\tau}_2 - r_2)^2}$$
$$\left. -\frac{2\hat{\tau}_2\hat{Q}_2}{\hat{\tau}_2 - r_2}\sum_{i=1}^{n}\left(\frac{\hat{p}_i^{(2)}}{\hat{B}_i^{(2)}} - \frac{\hat{D}_2}{\hat{B}_i^{(2)}}\right)\hat{p}_i^{(2)}\right\} \qquad (7)$$

where $\hat{Q}_2 = \prod_{i=1}^n (1-\hat{p}_i^{(2)})$,

$$\hat{A}_2 = \sum_{i=1}^{n}\frac{(\hat{p}_i^{(2)})^2}{\hat{B}_i^{(2)}} - \hat{C}_2 + \frac{1}{\hat{\tau}_2 + a_2 - 1} - \frac{1}{\hat{\tau}_2 - r_2},$$

$$\hat{B}_i^{(2)} = \hat{\tau}_2\hat{p}_i^{(2)}(1-\hat{p}_i^{(2)}) + \sigma_2^{-2},$$

$$\hat{C}_2 = \frac{(\nu_2^{-1} - n\sigma_2^{-2})\left[n^{-1}\sum_{i=1}^n \hat{p}_i^{(2)}/\hat{B}_i^{(2)}\right]^2}{1 + n^{-1}(\nu_2^{-1} - n\sigma_2^{-2})n^{-1}\sum_{i=1}^n 1/\hat{B}_i^{(2)}}, \quad \text{and}$$

$$\hat{D}_2 = \frac{n^{-1}(\nu_2^{-1} - n\sigma_2^{-2})n^{-1}\sum_{i=1}^n \hat{p}_i^{(2)}/\hat{B}_i^{(2)}}{1 + n^{-1}(\nu_2^{-1} - n\sigma_2^{-2})n^{-1}\sum_{i=1}^n 1/\hat{B}_i^{(2)}}.$$

To obtain a variance estimator of $\hat{\tau}$, notice that $\mathbf{E}_\xi(\hat{\tau}_2|\mathbf{m}_s) \approx \tau_2$ implies that $\mathbf{Cov}(\hat{\tau}_1, \hat{\tau}_2) \approx 0$, and consequently, a variance estimator of $\hat{\tau}$ is $\hat{\mathbf{V}}(\hat{\tau}) = \hat{\mathbf{V}}(\hat{\tau}_1) + \hat{\mathbf{V}}(\hat{\tau}_2)$.

3.2 Model II

In this case, we have two different joint posterior distributions for τ_1, τ_2, $\{\alpha_i^{(1)}\}$, and $\{\alpha_i^{(2)}\}$. Both posterior distributions are given by (1), but now $h_1(\tau_1)$ and $h_2(\tau_2)$, which will be denoted by $h_1^{(l)}(\tau_1)$ and $h_2^{(l)}(\tau_2)$, are given by

$$h_1^{(l)}(\tau_1) = \frac{(\tau_1 - l)!\,(1-n/N)^{\tau_1}}{(\tau_1 - m - r_1)!} \quad \text{and}$$
$$h_2^{(l)}(\tau_2) = \frac{(\tau_2 - l)!}{(\tau_2 - r_2)!}, \quad l = 0, 1. \qquad (8)$$

As in Model I, maximizing the logarithms of the posterior distributions, we obtain two estimators $(\hat{\tau}_1^{(l)}, \hat{\tau}_2^{(l)}, \{\hat{\alpha}_{(l)i}^{(1)}\}, \{\hat{\alpha}_{(l)i}^{(2)}\})$, $l = 0, 1$, of $(\tau_1, \tau_2, \{\alpha_i^{(1)}\}, \{\alpha_i^{(2)}\})$, which are defined as the solutions to the following systems of equations:

$$\hat{\tau}_1^{(l)} = \frac{m + r_1 - l(1-n/N)\prod_{i=1}^n (1-\hat{p}_{(l)i}^{(1)})}{1 - (1-n/N)\prod_{i=1}^n (1-\hat{p}_{(l)i}^{(1)})},$$
$$\hat{\tau}_2^{(l)} = \frac{r_2}{1 - \prod_{i=1}^n (1-\hat{p}_{(l)i}^{(2)})},$$

and $\hat{p}_{(l)i}^{(1)}$ and $\hat{p}_{(l)i}^{(2)}$, $i = 1, \ldots, n$, $l = 0, 1$, are given by (2) and (3).

Estimators of the variances of $\hat{\tau}_1^{(l)}$ and $\hat{\tau}_2^{(l)}$, $l = 0, 1$, can be obtained using the same strategy as that used in Model I. Thus, an estimator $\hat{\mathbf{V}}_{11}^{(l)}$ of

$\mathbf{V}_p\left[\mathbf{E}_\xi(\hat{\tau}_1^{(l)}\middle|\mathbf{m}_s)\right]$ is given by (5), but in this case $\hat{A}_1$, which will be denoted by $\hat{A}_1^{(l)}$, is given by

$$\hat{A}_1^{(l)} = \sum_{i=1}^{n}\frac{(\hat{p}_{(l)i}^{(1)})^2}{\hat{B}_{(l)i}^{(1)}} - \hat{C}_1^{(l)} - \frac{m+r_1-l}{(\hat{\tau}_1^{(l)}-l)(\hat{\tau}_1^{(l)}-m-r_1)},$$

$l = 0, 1$, and the estimators $\hat{B}_{(l)i}^{(1)}$, $\hat{C}_1^{(l)}$, and $\hat{Q}_1^{(l)}$ are given by the expressions for $\hat{B}_i^{(1)}$, $\hat{C}_1$, y $\hat{Q}_1$, but replacing $\hat{\tau}_1$ by $\hat{\tau}_1^{(l)}$, and $\hat{p}_i^{(l)}$ by $\hat{p}_{(l)i}^{(1)}$, $i = 1,\ldots,n$.

An estimator $\hat{\mathbf{V}}_{12}^{(l)}$ of $\mathbf{E}_p\left[\mathbf{V}_\xi(\hat{\tau}_1^{(l)}\middle|\mathbf{m}_s)\right]$ is given by (6), but using the estimators $\hat{\tau}_1^{(l)}$, $\hat{p}_i^{(l)}$, $\hat{Q}_1^{(l)}$, $\hat{A}_1^{(l)}$, $\hat{B}_{(l)i}$, $\hat{C}_1^{(l)}$, and $\hat{D}_1^{(l)}$. Therefore, a design-based estimator of $\mathbf{V}(\hat{\tau}_1^{(l)})$ is $\hat{\mathbf{V}}(\hat{\tau}_1^{(l)}) = \hat{\mathbf{V}}_{11}^{(l)} + \hat{\mathbf{V}}_{12}^{(l)}$.

In the case of $\hat{\tau}_2^{(l)}$, an estimator $\hat{\mathbf{V}}(\hat{\tau}_2^{(l)})$ of $\mathbf{V}(\hat{\tau}_2^{(l)})$ is given by (7), but now $\hat{A}_2$, which will be denoted by $\hat{A}_2^{(l)}$, is given by

$$\hat{A}_2^{(l)} = \sum_{i=1}^{n}\frac{(\hat{p}_{(l)i}^{(2)})^2}{\hat{B}_{(l)i}^{(2)}} - \hat{C}_2^{(l)} - \frac{r_2-l}{(\hat{\tau}_2^{(l)}-l)(\hat{\tau}_2-r_2)},$$

and the estimators $\hat{B}_{(l)i}^{(2)}$, $\hat{C}_2^{(l)}$, $\hat{D}_2^{(l)}$, and $\hat{Q}_2^{(l)}$ are given by the expressions for $\hat{B}_i^{(2)}$, $\hat{C}_2$, $\hat{D}_2$, and $\hat{Q}_2$, but replacing $\hat{\tau}_2$ by $\hat{\tau}_2^{(l)}$, and $\hat{p}_i^{(2)}$ by $\hat{p}_{(l)i}^{(2)}$, $i = 1,\ldots,n$.

As in the case of Model I, a variance estimator of $\hat{\tau}^{(l)}$ is $\hat{\mathbf{V}}(\hat{\tau}^{(l)}) = \hat{\mathbf{V}}(\hat{\tau}_1^{(l)}) + \hat{\mathbf{V}}(\hat{\tau}_2^{(l)})$.

4. Monte Carlo study

To observe the performances of the proposed estimators, the estimators of their variances, and the corresponding 95% normal-based confidence intervals, as well as to compare their performances with those of the MLEs proposed by Félix-Medina and Thompson (2003), we carried out a simulation study.

We considered two finite populations; a description of each population is presented in Table 1. Notice that the assumption of the Poisson distribution of the M_i's is only satisfied in Population I. The nomination probabilities $p_i^{(k)}$, $k = 1, 2$, were generated using the model $p_i^{(k)} = 1 - \exp(-\beta_k m_i)$, where the values of β_k were set so that the following values of $\mathbf{E}(p_i^{(k)})$ were obtained. Case 1: $(\mathbf{E}(p_i^{(1)}), \mathbf{E}(p_i^{(2)})) = (0.05, 0.03)$, and Case 2: $(\mathbf{E}(p_i^{(1)}), \mathbf{E}(p_i^{(2)})) = (0.01, 0.006)$.

The simulation experiment was executed as follows. From each population of $N = 250$ values of m_i's, a simple random sample without replacement of $n = 25$ values was selected. From cluster A_i in

Table 1: Parameters of simulated populations

Population I	Population II
$N = 250$	$N = 250$
M_i Poisson	M_i Neg. Binomial
$\mathbf{E}(M_i) = 7.2$	$\mathbf{E}(M_i) = 7.2$
$\mathbf{V}(M_i) = 7.2$	$\mathbf{V}(M_i) = 24.48$
$\tau_1 = 1897$	$\tau_1 = 1740$
$\tau_2 = 700$	$\tau_2 = 700$
$\tau = 2597$	$\tau = 2440$
$\tau_1/\tau = .72$	$\tau_1/\tau = .71$

the sample, the values of $X_{ij}^{(1)}$ were generated using $\tau_1 - m_i$ independent identically distributed Bernoulli random variables with mean $p_i^{(1)}$, and those of $X_{ij}^{(2)}$ using τ_2 independent Bernoulli random variables with mean $p_i^{(2)}$. A total of $r = 5000$ samples were selected from each population using the previous procedure. The values of the parameters of the initial distributions were the following: $\sigma_k^2 = 9$, $\mu_k = -3.5$, $\gamma_k^2 = 9$, $k = 1, 2$, $a_1 = 1$, $b_1 = 0.1$, $a_2 = 8$, $b_2 = 0.01$, so that $\mathbf{E}(\lambda_1) = 10$, $\mathbf{V}(\lambda_1) = 100$, $\mathbf{E}(\lambda_2) = 800$, and $\mathbf{V}(\lambda_2) = 80000$.

We observed the performances of the Bayesian estimators $\hat{\tau}_1$, $\hat{\tau}_2$, $\hat{\tau} = \hat{\tau}_1 + \hat{\tau}_2$, $\hat{\tau}_1^{(l)}$, $\hat{\tau}_2^{(l)}$, $\hat{\tau}^{(l)} = \hat{\tau}_1^{(l)} + \hat{\tau}_2^{(l)}$, $l = 0, 1$, and those of the MLEs $\tilde{\tau}_1$, $\tilde{\tau}_2$, and $\tilde{\tau} = \tilde{\tau}_1 + \tilde{\tau}_2$ proposed by Félix- Medina and Thompson (2003). The performance of an estimator $\hat{\tau}$, say, was evaluated by its relative bias and square root of its relative mean square error, defined as r-bias $= \sum_1^r(\hat{\tau}_i - \tau)/(r\tau)$ and $\sqrt{\text{r-mse}} = \sqrt{\sum_1^r(\hat{\tau}_i-\tau)^2/(r\tau^2)}$, where $\hat{\tau}_i$ was the value of $\hat{\tau}$ obtained in the i-th replication.

In addition, we observed the performances of the design-based variance estimators of the estimators of the population sizes. In the case of the MLEs, the model-based estimators of their variances obtained from the observed Fisher information matrix were also considered in the study. The performance of a variance estimator was also evaluated by its relative bias and the square root of its relative mean square error, which were similarly defined to those of an estimator of the population size.

Finally, the performance of a 95% normal-based confidence interval, defined as $\hat{\tau} \pm 1.96\sqrt{\hat{\mathbf{V}}(\hat{\tau})}$, was evaluated by its relative frequency of coverage and its semi-length.

The results of the simulation study, Tables 2-4, showed that when the nomination probabilities were large (Case 1) and regardless of whether or not the Poisson distribution of the M_i's was satisfied, the performance of every one of the estimators of the population size was good, that is, the relative biases

Table 2: Results of the estimators of the population sizes

	Population I				Population II			
	$\mathbf{E}(p_i^{(1)}) = 0.05$ $\mathbf{E}(p_i^{(2)}) = 0.03$		$\mathbf{E}(p_i^{(1)}) = 0.01$ $\mathbf{E}(p_i^{(2)}) = 0.006$		$\mathbf{E}(p_i^{(1)}) = 0.05$ $\mathbf{E}(p_i^{(2)}) = 0.03$		$\mathbf{E}(p_i^{(1)}) = 0.01$ $\mathbf{E}(p_i^{(1)}) = 0.006$	
	r-bias	$\sqrt{\text{r-mse}}$	r-bias	$\sqrt{\text{r-mse}}$	r-bias	$\sqrt{\text{r-mse}}$	r-bias	$\sqrt{\text{r-mse}}$
$\tilde{\tau}_1$	-0.0007	0.0169	-0.0019	0.0529	-0.0025	0.0221	-0.0050	0.0914
$\tilde{\tau}_2$	0.0018	0.0599	78916.3	5.56×10^6	0.0043	0.0692	4.3×10^7	4.3×10^9
$\tilde{\tau}$	-6.08×10^{-5}	0.0204	21271.2	1.5×10^6	-0.0006	0.0253	1.2×10^7	1.2×10^9
$\hat{\tau}_1^{(0)}$	-0.0007	0.0169	-0.0022	0.0529	-0.0025	0.0221	-0.0055	0.0915
$\hat{\tau}_2^{(0)}$	0.0018	0.0599	0.1175	0.5855	0.0043	0.0691	0.1447	0.7180
$\hat{\tau}^{(0)}$	-3.32×10^{-5}	0.0204	0.0300	0.1611	-0.0005	0.0253	0.0376	0.2116
$\hat{\tau}_1^{(1)}$	-0.0009	0.0170	-0.0052	0.0532	-0.0029	0.0222	-0.0088	0.0918
$\hat{\tau}_2^{(1)}$	-0.0017	0.0593	-0.0325	0.3595	-0.0003	0.0682	-0.0511	0.3703
$\hat{\tau}^{(1)}$	-0.0012	0.0203	-0.0125	0.1050	-0.0021	0.0253	-0.0209	0.1283
$\hat{\tau}_1$	-0.0031	0.0172	-0.0261	0.0580	-0.0055	0.0229	-0.0308	0.0948
$\hat{\tau}_2$	0.0016	0.0216	0.0033	0.0673	0.0038	0.0270	-0.0009	0.0721
$\hat{\tau}$	-0.0018	0.0201	-0.0182	0.0648	-0.0028	0.0249	-0.0222	0.0859

Notes: $\tilde{\tau}_1$, $\tilde{\tau}_2$ and $\tilde{\tau}$, MLE's; $\hat{\tau}_1^{(0)}$, $\hat{\tau}_2^{(0)}$ and $\hat{\tau}^{(0)}$, Bayesian estimators based on the initial improper uniform distributions; $\hat{\tau}_1^{(1)}$, $\hat{\tau}_2^{(1)}$ and $\hat{\tau}^{(1)}$, Bayesian estimators based on the initial Jeffreys' distributions; $\hat{\tau}_1$, $\hat{\tau}_2$ and $\hat{\tau}$, Bayesian estimators based on the initial two-stage Poisson-Gamma distributions.

were less than 0.01 and the square roots of their relative mean square error were less than 0.1. Thus, the estimators were robust to deviations from the Poisson assumption. However, when the nomination probabilities were small (Case 2), the performances of the MLEs of τ_2 and τ were very badly. The performances of the corresponding Bayesian estimators were much better than those of the MLEs, but not as good as those obtained in Case 1. The best performances were obtained by the Bayesian estimators based on the initial Poisson distributions for τ_1 and τ_2.

The performance of every one of the variance estimators was acceptable when the nomination probabilities were large and the M_i's were Poisson distributed (the estimators were practically unbiased and the values of the $\sqrt{\text{r-mse}}$ fell between 0.15 and 0.29). When the M_i's were not Poisson distributed and the nomination probabilities were large, the model-based variance estimator of the MLEs of τ_1 tended to subestimate the variance (r-bias$=-0.2$). The performances of the design-based variance estimators were not seriously affected by the deviation from the Poisson assumption. However, when the nomination probabilities were small, every one of the estimators of the variances of the estimators of τ_2 and τ, except the estimators of the variances of $\hat{\tau}_2$ and $\hat{\tau}$, seriously overestimated the variances. The estimators $\hat{\mathbf{V}}(\hat{\tau}_2)$ and $\hat{\mathbf{V}}(\hat{\tau})$ performed very satisfactorily.

The performance of each one of the confidence interval estimators was influenced by the performances of the corresponding estimator of the population size and its variance estimator. Thus, when both of these estimators performed well, so did the confidence interval estimator; but when one of them perfomed badly, so did the interval estimator.

5. Conclusions

In this work we have used the design proposed by Félix-Medina and Thompson (2003) to sample a hidden human population, and we have developed estimators of the population size using the Bayesian approach. From the frequentist approach, we have developed design-based variance estimators and normal-based confidence intervals of the population sizes.

According to the simulation results, the proposed Bayesian estimators are robust to deviations from the assumed Poisson distribution of the cluster sizes. They perform similarly to the MLEs proposed by Félix-Medina and Thompson (2003) when the nomination probabilities are large, and better than them when the nomination probabilities are small.

The best estimators were the Bayesian estimators obtained by using the informative Poisson initial distributions for τ_1 and τ_2. These estimators, as well

Table 3: Results of the variance estimators.

	Population I				Population II			
	$\mathbf{E}(p_i^{(1)}) = 0.05$ $\mathbf{E}(p_i^{(2)}) = 0.03$		$\mathbf{E}(p_i^{(1)}) = 0.01$ $\mathbf{E}(p_i^{(2)}) = 0.006$		$\mathbf{E}(p_i^{(1)}) = 0.05$ $\mathbf{E}(p_i^{(2)}) = 0.03$		$\mathbf{E}(p_i^{(1)}) = 0.01$ $\mathbf{E}(p_i^{(1)}) = 0.006$	
	r-bias	$\sqrt{\text{r-mse}}$	r-bias	$\sqrt{\text{r-mse}}$	r-bias	$\sqrt{\text{r-mse}}$	r-bias	$\sqrt{\text{r-mse}}$
$\tilde{\mathbf{V}}_M(\tilde{\tau}_1)$	-0.0197	0.1559	0.0399	0.0792	-0.2016	0.2929	-0.5988	0.5998
$\tilde{\mathbf{V}}_M(\tilde{\tau}_2)$	0.0071	0.2852	3.1×10^7	2.8×10^9	-0.0119	0.4362	5.1×10^{10}	5.1×10^{12}
$\tilde{\mathbf{V}}_M(\tilde{\tau})$	-0.0073	0.2124	3.1×10^7	2.8×10^9	-0.0844	0.3412	5.1×10^{10}	5.1×10^{12}
$\tilde{\mathbf{V}}_D(\tilde{\tau}_1)$	-0.0348	0.1523	-0.1096	0.1769	-0.1179	0.2588	-0.1846	0.2966
$\tilde{\mathbf{V}}_D(\tilde{\tau}_2)$	0.0071	0.2852	3.1×10^7	2.8×10^9	-0.0119	0.4362	5.1×10^{10}	5.1×10^{12}
$\tilde{\mathbf{V}}_D(\tilde{\tau})$	-0.0129	0.2109	3.1×10^7	2.8×10^9	-0.0522	0.3386	5.1×10^{10}	5.1×10^{12}
$\hat{\mathbf{V}}_D(\hat{\tau}_1^{(0)})$	-0.0242	0.1529	0.0046	0.1697	-0.0827	0.2575	-0.0282	0.2854
$\hat{\mathbf{V}}_D(\hat{\tau}_2^{(0)})$	0.0071	0.2849	0.2867	15.4462	-0.0117	0.4348	0.5912	14.915
$\hat{\mathbf{V}}_D(\hat{\tau}^{(0)})$	-0.0091	0.2111	0.2877	14.738	-0.0393	0.3394	0.5895	14.008
$\hat{\mathbf{V}}_D(\hat{\tau}_1^{(1)})$	-0.0259	0.1528	0.0005	0.1688	-0.0878	0.2575	-0.0342	0.2854
$\hat{\mathbf{V}}_D(\hat{\tau}_2^{(1)})$	0.0059	0.2815	0.2057	3.1670	-0.0122	0.4262	0.3677	3.6031
$\hat{\mathbf{V}}_D(\hat{\tau}^{(1)})$	-0.0129	0.2079	0.1697	2.7105	-0.0518	0.3306	0.2009	2.4809
$\hat{\mathbf{V}}_D(\hat{\tau}_1)$	-0.0384	0.1517	-0.0293	0.1652	-0.1221	0.2598	-0.0734	0.2799
$\hat{\mathbf{V}}_D(\hat{\tau}_2)$	0.0059	0.2641	0.0183	0.2051	-0.0139	0.3867	0.0438	0.1936
$\hat{\mathbf{V}}_D(\hat{\tau})$	-0.0160	0.1964	-0.0059	0.1411	-0.0598	0.3061	-0.0489	0.1830

Notes: $\tilde{\mathbf{V}}_M(\cdot)$, model-based variance estimators; $\tilde{\mathbf{V}}_D(\cdot)$, design-based variance estimators.

as the estimators of their variances and the confidence intervals based on them performed very well. From this, we conclude that if initial information were available, it could be worth of using it.

6. References

Castledine, B.J. (1981). A bayesian analysis of multiple- recapture sampling for a closed population. *Biometrika*, **67**, 197-210.

Félix-Medina, M.H., and Thompson, S.K. (2003). Combining cluster sampling and link-tracing sampling to estimate the size of hidden populations. To appear in *Journal of Official Statistics.*

Fienberg, S.E., Johnson, M.S. and Junker, B.W. (1999). Classical multilevel and bayesian approaches to population size estimation using multiple lists. *Journal of the Royal Statistical Society, A*, **162**, 383-405.

Frank, O. and Snijders, T.A.B. (1994). Estimating the size of hidden populations using snowball sampling. *Journal of Official Statistics*, **10**, 53-67.

Spreen, M. (1992). Rare populations, hidden populations, and link-tracing designs: what and why?. *Bulletin de Méthodologie Sociologique*, **36**, 59-70.

Thompson, S.K. and Frank, O. (2000). Model-based estimation with link-tracing sampling designs. *Survey Methodology*, **26**, 87-98.

Table 4: Results of the 95% confidence intervals.

	Population I				Population II			
	$\mathbf{E}(p_i^{(1)}) = 0.05$ $\mathbf{E}(p_i^{(2)}) = 0.03$		$\mathbf{E}(p_i^{(1)}) = 0.01$ $\mathbf{E}(p_i^{(2)}) = 0.006$		$\mathbf{E}(p_i^{(1)}) = 0.05$ $\mathbf{E}(p_i^{(2)}) = 0.03$		$\mathbf{E}(p_i^{(1)}) = 0.01$ $\mathbf{E}(p_i^{(1)}) = 0.006$	
	Coverage	Semi-length	Coverage	Semi-length	Coverage	Semi-length	Coverage	Semi-length
$\tilde{\tau}_1 \pm 1.96\sqrt{\tilde{\mathbf{V}}_M(\tilde{\tau}_1)}$	0.9477	62.29	0.9530	200.67	0.9221	66.36	0.7827	197.08
$\tilde{\tau}_2 \pm 1.96\sqrt{\tilde{\mathbf{V}}_M(\tilde{\tau}_2)}$	0.9469	81.64	0.9303	5.6×10^{11}	0.9484	92.25	0.940	1.3×10^{16}
$\tilde{\tau} \pm 1.96\sqrt{\tilde{\mathbf{V}}_M(\tilde{\tau})}$	0.9489	102.86	0.9546	5.6×10^{11}	0.9410	113.88	0.9288	1.3×10^{16}
$\tilde{\tau}_1 \pm 1.96\sqrt{\tilde{\mathbf{V}}_D(\tilde{\tau}_1)}$	0.9456	61.81	0.9294	185.22	0.9347	69.78	0.9079	278.37
$\tilde{\tau}_2 \pm 1.96\sqrt{\tilde{\mathbf{V}}_D(\tilde{\tau}_2)}$	0.9469	81.64	0.9303	5.6×10^{11}	0.9484	92.25	0.940	1.3×10^{16}
$\tilde{\tau} \pm 1.96\sqrt{\tilde{\mathbf{V}}_D(\tilde{\tau})}$	0.9484	102.57	0.9497	5.6×10^{11}	0.9467	115.93	0.9484	1.3×10^{16}
$\hat{\tau}_1^{(0)} \pm 1.96\sqrt{\hat{\mathbf{V}}_D(\hat{\tau}_1^{(0)})}$	0.9477	62.16	0.9451	196.66	0.9395	71.13	0.9288	303.79
$\hat{\tau}_2^{(0)} \pm 1.96\sqrt{\hat{\mathbf{V}}_D(\hat{\tau}_2^{(0)})}$	0.9468	81.61	0.9235	591.20	0.9485	92.19	0.9150	698.11
$\hat{\tau}^{(0)} \pm 1.96\sqrt{\hat{\mathbf{V}}_D(\hat{\tau}^{(0)})}$	0.9491	102.75	0.9473	634.10	0.9483	116.71	0.9469	788.56
$\hat{\tau}_1^{(1)} \pm 1.96\sqrt{\hat{\mathbf{V}}_D(\hat{\tau}_1^{(1)})}$	0.9464	62.10	0.9439	196.24	0.9384	71.04	0.9267	303.03
$\hat{\tau}_2^{(1)} \pm 1.96\sqrt{\hat{\mathbf{V}}_D(\hat{\tau}_2^{(1)})}$	0.9433	80.89	0.8667	450.61	0.9457	91.14	0.8515	483.35
$\hat{\tau}^{(1)} \pm 1.96\sqrt{\hat{\mathbf{V}}_D(\hat{\tau}^{(1)})}$	0.9457	102.14	0.9036	500.12	0.9445	115.82	0.9054	588.39
$\hat{\tau}_1 \pm 1.96\sqrt{\hat{\mathbf{V}}_D(\hat{\tau}_1)}$	0.9413	61.57	0.9037	189.17	0.9314	70.29	0.8974	291.29
$\hat{\tau}_2 \pm 1.96\sqrt{\hat{\mathbf{V}}_D(\hat{\tau}_2)}$	0.9470	79.53	0.9384	250.89	0.9483	89.02	0.9407	250.22
$\hat{\tau} \pm 1.96\sqrt{\hat{\mathbf{V}}_D(\hat{\tau})}$	0.9445	100.72	0.9289	314.88	0.9430	113.64	0.9234	385.45

Notes: $\tilde{\mathbf{V}}_M(\cdot)$, model-based variance estimators; $\tilde{\mathbf{V}}_D(\cdot)$, design-based variance estimators. $\tilde{\tau}_1$, $\tilde{\tau}_2$ and $\tilde{\tau}$, MLE's; $\hat{\tau}_1^{(0)}$, $\hat{\tau}_2^{(0)}$ and $\hat{\tau}^{(0)}$, Bayesian estimators based on the initial improper uniform distributions; $\hat{\tau}_1^{(1)}$, $\hat{\tau}_2^{(1)}$ and $\hat{\tau}^{(1)}$, Bayesian estimators based on the initial Jeffreys' distributions; $\hat{\tau}_1$, $\hat{\tau}_2$ and $\hat{\tau}$, Bayesian estimators based on the initial two-stage Poisson-Gamma distributions.