

SMALL AREA ESTIMATION OF THE NUMBER OF VISITS TO THE SPECIALIST IN 2002/2003 FRENCH HEALTH SURVEY

Marius STEFAN*

*University Politehnica of Bucharest; Université Libre de Bruxelles
E-mail: mastefan@gmail.com, mastefan@ulb.ac.be

In this paper we model one variable of the 2002/2003 French Health Survey (*The Number of Visits to the Specialist in the last twelve months*) in order to obtain sub national estimations (county and regional). We construct the model, we derive the theoretical estimations for the parameters of interest . Then we test the fit of the model to the data and after deciding that the model is good enough we compute the estimations and their precisions. We conclude the paper with some directions for future research

Key words: .small area, direct and indirect estimations, borrow strength, Markov chains

1. INTRODUCTION

The French Health Survey (FHS) is a large survey (almost 30000 observations) taking place every ten years and collecting information on a large number of health variables (more than 200). The Direction de la Recherche, des Etudes, de l'Evaluation et des Statistiques (DREES) is responsible for the statistical exploitation of the FHS data. Part of its job is to obtain national estimations for some parameters linked to a series of FHS variables. In doing this DREES is helped by INSEE, the French Statistical Institute. INSEE computes these estimations by using its well established methodology based on the classical survey sampling theory. The classical survey sampling theory centres its inference on the survey sampling distribution which is generated by the survey design, the way the sample is selected.

Recently there has been a growing demand for estimations at sub-national level. For instance, the French regional authorities are interested in estimating regional and county parameters (the French territory is divided into 22 regions, every region incorporating several counties, resulting in a total number of 95 counties). Generally, the national surveys like FHS are designed to insure an adequate level of precision at national level. When it comes to regional/county parameters one can still use the classical survey sampling theory resulting in the same formulas for the estimators and their precisions as at the national level but using the regional/county samples. These samples are composed of the observations from the national survey that come from the region or the county of interest. For a lot of such sub populations called alternatively areas or domains these observations are not numerous. This is why they are called *small areas* or *small domains*. As a consequence the estimators based on the classical survey sampling theory called *direct estimators* have not an adequate level of precision and alternative methods should be used.

To tackle this problem, authorities from five regions (see below) decided to spend more money to increase their regional sample size so that the regional estimations based on INSEE methodology be more precise. This resulted in approximately 10000 more individuals interviewed in the five regions. Using this additional sample INSEE computed and published national estimates plus five regional estimates together with their standard errors for a number of variables in FHS. However, these estimates now more precise due to the increased number of observations are still based on the INSEE methodology.

The small area estimation is the new theory trying to improve the classical design-based survey sampling theory when it comes to estimations parameters at sub national levels. The key of the modern small area estimation is the modelling of the variable of interest population values and then using the model to make inference. The model acts like a link between observations coming from different areas of the

population. This is why when model-based an estimator for a sub population called *indirect estimator* uses the entire national sample, not only the sample coming from the sub population. Thus the indirect estimator is generally more precise than the direct estimator by *borrowing strength* from related areas. A detailed account of the small area estimation is given in Rao[3].

INSEE has not a methodology using small area estimation techniques. This is why DREES financed a research aimed at finding a small area methodology for regional and county parameters related to a number of variables in FHS. The results presented in this paper are part of this research. In section 2 we show how to construct a model which will be used to estimate the regional and county means for the study variable *Number of Visits to the Specialist in the last twelve months* coded as R04AM in FHS. In section 3 we test the fit of the model. In section 4 we obtain the theoretical formulas of the estimators and their standard errors. Then we use the theoretical formulas to compute the estimations and their standard errors. Finally, in the last section we draw some conclusions and specify directions for future research.

2. COUNSTRUCTION OF THE MODEL

The variable R04AM is a count variable just as the variable R02AM - the *Number of visits to the generalist in the last twelve months* – for which we showed how to obtain regional estimations in Stefan[5]. The methodology for R04AM follow closely that for R02AM so in this paper we will not go into details preferring to underline the differences between the two variables instead.

First we undertook an exploratory analysis to see which are the variables R04AM depends on. We retained four variables: the Region, the Sex and the Age indexed respectively by i , s and k . The Region has 22 values, the Sex has 2 values and the Age has 8 values because we transformed the Age from a continuous to a categorical variable with 8 values corresponding to the intervals $[0,1]$, $[2,12]$, $[13,23]$, ... $[56,67]$ and $[68,104]$. We decided to drop the variable Stratum because its influence on the fit of the model as well as on the final results was found to be small. As in Stefan[5] l represents an individual. As a result the population individuals are cross classified in $22 \times 2 \times 8$ cells with y_{iskl} being the value of R04AM for an individual l in cell (i, s, k) . y_{iskl} should then verify $E(y_{iskl}) \approx \mu_{isk}$ with μ_{isk} denoting the population mean of cell (i, s, k) . R04AM is a count variable so we will use the Poisson distribution to model it meaning that $y_{iskl} \sim \text{Poisson}(\mu_{isk})$. The Poisson distribution has its mean equal to its variance. We have to test if R04AM verify this condition. We computed the sample means and variances of the cells (i, s, k) . Figure 1 represents the cell mean versus the cell variance. The solid line is of equation $y = x$. Most of the points are above it suggesting that the variance is larger than the mean. Such a situation is called over dispersion and is frequently met in practice. The second line is of equation $y = 4.68x$ (we anticipated by plugging in the value 4.68 which comes from the estimation of the model for R04AM presented later in this paper) which is a much closer approximation of the relationship between the variance and the mean of R04AM. Clearly, the Poisson distribution is not appropriate for R04AM and something has to be done to handle the over dispersion. We also had over dispersion in the data for the variable R02AM presented in Stefan[5] and refer the reader to that paper for a detailed account of the way we adopted to handle the over dispersion. We give here a short report.

If we don't take into account the over dispersion than the resulting model will be like model 1 in Stefan[5]. The main idea is to take random the parameters μ_{isk} of $\text{Poisson}(\mu_{isk})$. This is model 2 in Stefan[5] but we showed there that a far better model is model 3 achieved if we take μ_{iskl} instead of μ_{isk} . In Stefan[5] we compared the fit of the three models and concluded that model 3 presents a much better fit to the R02AM data than that of models 1 and 2. The same holds true for R04AM so we directly take model 3 as the model which we will use to get the regional estimations. The model bellow is model 3 for R02AM in Stefan[5]:

Model for R02AM

$$y_{iskl} | v_{iskl} \underset{ind}{\sim} \text{Poisson}(v_{iskl}),$$

$$v_{iskl} | \mu_{isk}, \underset{ind}{\alpha}, \underset{ind}{\kappa} \sim \text{Gamma}\left(\frac{\mu_{isk}^{1-\kappa}}{\alpha}, \frac{\mu_{isk}^{-\kappa}}{\alpha}\right),$$

$$\begin{aligned} \log(\mu_{isk}) &= \beta_{li} + \beta_{3s} + \beta_{4k}, \\ \alpha &\sim \text{Unif}(0,100), \quad \kappa \sim \text{Unif}(-1,100), \\ \beta_{li} &\sim \text{Unif}(-10,10), \quad \beta_{3s} \sim \text{Unif}(-10,10), \quad \beta_{4k} \sim \text{Unif}(-10,10) \end{aligned}$$

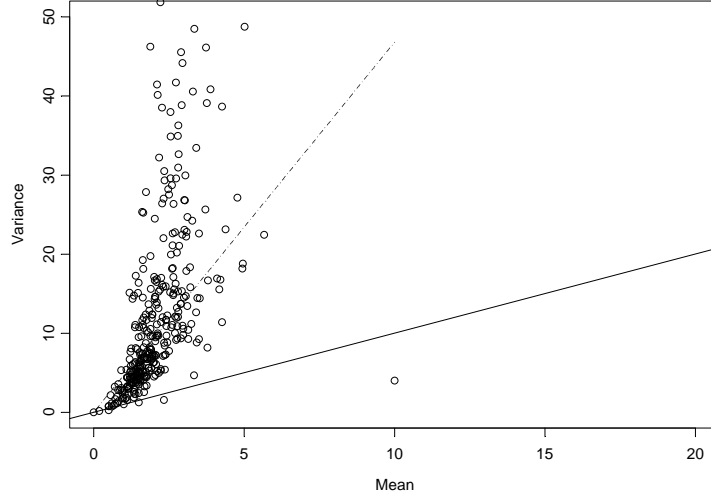


Figure 1. Cell mean versus cell variance of R04AM

The choice of the second line of the model is motivated by the relationship between mean and variance of R04AM that can be seen in Figure 1. It is of the same type than that for R02AM, that is $V(y_{iskl}) = \mu_{isk} + \alpha \times \mu_{isk}^{1+\kappa}$. It can be proved (see Stefan[5]) that we can model this relationship by taking v_{iskl} following Gamma distributions of parameters $\mu_{isk}^{1-\kappa} / \alpha$ and $\mu_{isk}^{-\kappa} / \alpha$. So the model above should also work for R04AM. However, Figure 1 seems to indicate that in the case of R04AM κ is zero. This is confirmed by estimating the model above with the data of R04AM (this was not the case for R02AM when κ was estimated at 0.5).

The third line is the usual link function for the mean of the Poisson distribution. We also tested the fit of a model with $\log(\mu_{isk}) = \beta_{li} + \beta_{3s} + \beta_{4k} + \varepsilon_{isk}$ instead of $\log(\mu_{isk}) = \beta_{li} + \beta_{3s} + \beta_{4k}$ but the result was a model with approximately the same fit as the model above. Finally the last two lines are the a priori laws for the hyper parameters of the model. As for R02AM we took non informative uniform distributions with the intervals large enough to mark the absence of a priori information about these parameters. We also undertook a sensitivity analysis to the choice of the a priori laws but we found no evidence of influence on the final estimations confirming a well known fact that when the sample size is large the a priori information has little or no influence on the final results. Given what we discussed above the model for R04AM will be:

$$\begin{aligned} &\text{Model for R04AM} \\ &y_{iskl} | v_{iskl, ind} \sim \text{Poisson}(v_{iskl}), \\ &v_{iskl} | \mu_{isk}, \alpha \sim \text{Gamma}\left(\frac{\mu_{isk}}{\alpha}, \frac{1}{\alpha}\right), \\ &\log(\mu_{isk}) = \beta_{li} + \beta_{3s} + \beta_{4k}, \\ &\alpha \sim \text{Unif}(0,100), \\ &\beta_{li} \sim \text{Unif}(-10,10), \quad \beta_{3s} \sim \text{Unif}(-10,10), \quad \beta_{4k} \sim \text{Unif}(-10,10) \end{aligned}$$

3. MODEL FIT

In the previous section we showed how to construct a model for R04AM. We now test the fit of this model to the data. We will use the sample without extension having 28259 observations and the same tools as in Stefan[5]. There are two categories of measures of fit: those that help selecting between several models and those telling if a model is good enough for making inference. In Stefan[5] the first measures were used to select between models 1, 2 and 3, the last measures to check if model 3 is good enough. For R04AM in order to use save space we will not show the numerical results of comparing models 1 and 2 to model 3 because the same conclusions hold true for R04AM (model 3 or his adapted form for R04AM has a much better fit to the data than models 1 and 2) and because the comparison serves only to see the effect of not taking into account the over dispersion. We will focus on the question to see how good is the model above. In this section, for notation facility i designates an individual.

Let \mathbf{y}_{obs} be the vector of all the observations. One way in which a model fit can be tested is to generate for every individual i in the sample new observations $y_{new,i}$ from the posterior predictive density $f(y_i | \mathbf{y}_{obs})$ and to compare the vector of these new observations \mathbf{y}_{new} to \mathbf{y}_{obs} . It can be shown that a new value $y_{new,i}$ can be sampled from $f(y_i | \mathbf{y}_{obs})$ as follows: for each individual i we have the Markov chain $\{\nu_i^g\}$ corresponding to ν_i obtained by estimating the model. After the burn-in period the values $\{\nu_i^g\}$ come from $f(\nu_i | \mathbf{y}_{obs})$; we considered a burn-in period of 2000 iterations after which the chains reach convergence and used the next $G=1000$ iterations ; for each of the 1000 iterations we generated $y_{new,i}^g$ by sampling $\text{Poisson}(\nu_i^g)$.

An alternative distribution to generate new data is from the cross validation prediction densities denoted by $f(y_i | \mathbf{y}_{(i)})$ where $\mathbf{y}_{(i)}$ is the vector of all the observations except i . For an individual i $f(y_i | \mathbf{y}_{(i)})$ suggests what values of y_i are likely when the model is fitted to $\mathbf{y}_{(i)}$. Fortunately there is a way to sample $f(y_i | \mathbf{y}_{(i)})$ without having to rerun the Gibbs sampling for every $\mathbf{y}_{(i)}$ and then sample the corresponding posterior predictive density to generate new values as we showed above for $f(y_i | \mathbf{y}_{obs})$. Under our model the algorithm is equivalent to the following: from each vector $\mathbf{v}_i = (\nu_i^g)$ draw a sample with replacement and with probabilities proportional to $1/f(y_{obs,i} | \nu_i^g)$ and let $\mathbf{v}_i^* = (\nu_i^{g*})$, $g=1, \dots, G$ be the new vector; for each element ν_i^{g*} sample a value $y_{new,i}^{g*}$ from $\text{Poisson}(\nu_i^{g*})$. The vector $\mathbf{y}_{new,i}^*$ will be composed of G values sampled from $f(y_i | \mathbf{y}_{(i)})$.

In order to compare \mathbf{y}_{obs} to \mathbf{y}_{new} and \mathbf{y}_{new}^* we represented in Figure 2 their empirical distributions.

If the distribution of \mathbf{y}_{obs} is similar to that of \mathbf{y}_{new} or \mathbf{y}_{new}^* than the model generates the same values as those actually observed and the conclusion is that the model is well fit to the data. In the literature \mathbf{y}_{new} was criticized as being unreliable because the values ν_i^g used to generate $y_{new,i}^g$ come from fitting the model to the data, that is to \mathbf{y}_{obs} . So \mathbf{y}_{new} tends to resemble \mathbf{y}_{obs} even if we have a bad model. This is why it was proposed the vector \mathbf{y}_{new}^* whose components $y_{new,i}^*$ were each generated from $\mathbf{y}_{(i)}$ not containing y_i . In our case Figure 2 shows that distributions of \mathbf{y}_{new} and \mathbf{y}_{new}^* are almost the same and both of them are close to \mathbf{y}_{obs} . This fact let us conclude that the model for R04AM is well fitted to the data.

In Stefan[5] we also used some p-values based on two measures of discrepancy between \mathbf{y} and \mathbf{v} : the Deviance and the Dis measure defined in Stefan[5]. The associated p-values are the probabilities $P[\text{Deviance}(\mathbf{y}_{new}, \mathbf{v}) \geq \text{Deviance}(\mathbf{y}_{obs}, \mathbf{v})]$ and respectively $P[\text{Dis}(\mathbf{y}_{new}, \mathbf{v}) \geq \text{Dis}(\mathbf{y}_{obs}, \mathbf{v})]$. A p-value close to 0.5 indicates a good fit while extreme values close to 0 or 1 indicates a model poorly fit to the data. In Stefan[5] we explained how these p-values can be estimated. Under the adapted model for R04AM we obtained 0.46 for $P[\text{Deviance}(\mathbf{y}_{new}, \mathbf{v}) \geq \text{Deviance}(\mathbf{y}_{obs}, \mathbf{v})]$ and 0.35 for $P[\text{Dis}(\mathbf{y}_{new}, \mathbf{v}) \geq \text{Dis}(\mathbf{y}_{obs}, \mathbf{v})]$ indicating again that the model is well adapted to the data.

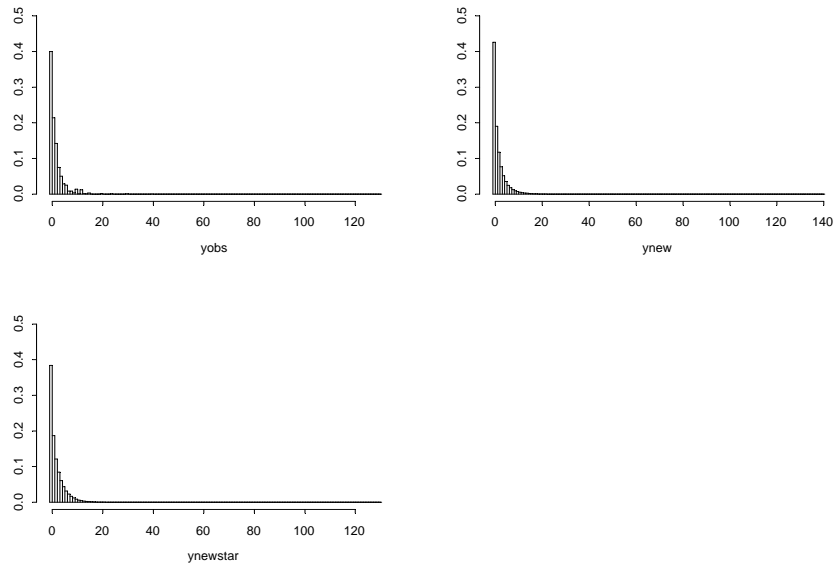


Figure 2. Distributions of y_{obs} , y_{new} and y_{new}^* under the model for R04AM

In Stefan[5] we computed and used the standardized residuals defined as:

$$r_i = \frac{y_{obs,i} - E(y_i | \mathbf{y}_{(i)})}{\sqrt{V(y_i | \mathbf{y}_{(i)})}}$$

In Figure 3 we plotted the r_i . We can notice that most of r_i are between -2 and 2 and that there is no particular pattern in the structure of r_i .

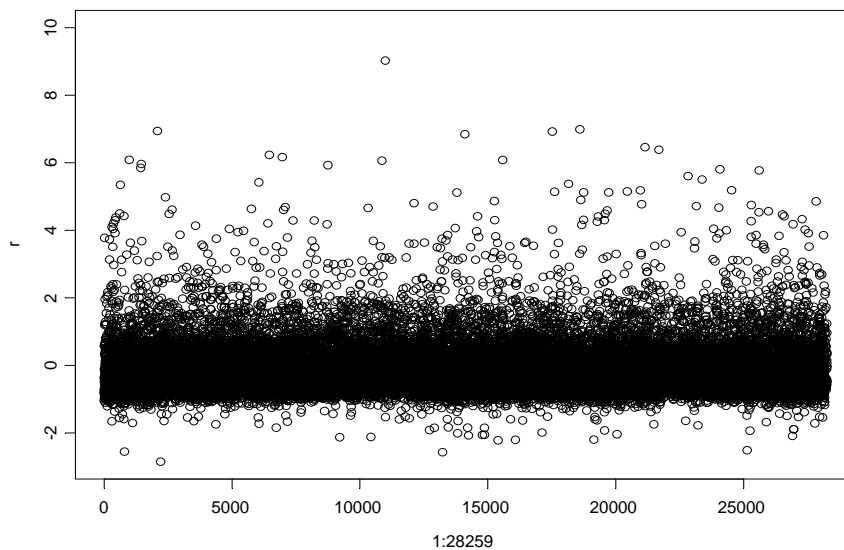


Figure 3. The standardized residuals under the model for R04AM

4. PARAMETERS ESTIMATION

The parameters of interest are the regional means μ_i , $i=1,22$ of the variable R04AM. We will use the principles of Bayesian statistics which consist of estimating a parameter by its posterior mean and the precision of this estimation by its posterior variance or standard error. Thus, our estimators will be $\hat{\mu}_i = E(\mu_i | \mathbf{y}_{obs})$ and their variances $V(\hat{\mu}_i) = V(\mu_i | \mathbf{y}_{obs})$. So the focus of the Bayesian statistics is the posterior distribution of a parameter that is the distribution of the parameter after the sample was selected. Generally this distribution cannot be obtained in a closed form but in order to have $E(\mu_i | \mathbf{y}_{obs})$ and $V(\mu_i | \mathbf{y}_{obs})$ it is sufficient to get a sample from the posterior distribution. Then $E(\mu_i | \mathbf{y}_{obs})$ and $V(\mu_i | \mathbf{y}_{obs})$ will be approximated by the sample mean and the sample variance respectively. The whole difficulty consists of finding a way to sample the posterior distribution. Generally and for a large number of models these samples called Markov chains are obtained by the Gibbs sampling or the Metropolis-Hastings algorithm. We will not go deeper into the Markov chains theory and the problems one has to deal with when working with Markov chains like initial values, convergence, burn-in period, number of iterations to use, Monte Carlo precision, and so on. A somewhat larger account is given in Stefan[5] and Stefan[4]. For more complete account there is Rao[3] or any other book about Bayesian statistics.

Using the well known Metropolis Hastings algorithm within Gibbs sampling (because some of the full conditional distributions resulting from the above hierarchical model for R04AM are known up to a constant) one can get Markov chains of any length for each parameter of the model. We will denote the elements of the chains by μ_{isk}^g , $alpha^g$, ... with $g=1, \dots, G$. Then, each parameter can be estimated by the sample mean and the precision can be measured by the sample variance. But we are interested in μ_i which are not part of the model so first we have find the link between the regional means and the parameters in the model. This is done in Stefan[5]. We present only the formulas for $\hat{\mu}_i$ and $V(\hat{\mu}_i)$ which are given by:

$$\hat{\mu}_i = \frac{1}{N_i} \left[\sum_s \sum_k \sum_{l \in obs_s} y_{iskl} + \frac{1}{G} \sum_g \sum_s \sum_k (N_{isk} - n_{isk}) \mu_{isk}^g \right] \quad (1)$$

$$\begin{aligned} V(\hat{\mu}_i) = & \frac{1}{N_i^2} \left\{ \frac{1}{G} \sum_g \sum_s \sum_k (N_{isk} - n_{isk}) \mu_{isk}^g + \frac{1}{G} \sum_g \sum_s \sum_k (N_{isk} - n_{isk}) \alpha^g \mu_{isk}^g + \right. \\ & \left. + \frac{1}{G} \sum_g \left[\sum_s \sum_k (N_{isk} - n_{isk}) \mu_{isk}^g \right]^2 - \left[\frac{1}{G} \sum_g \sum_s \sum_k (N_{isk} - n_{isk}) \mu_{isk}^g \right]^2 \right\} \quad (2) \end{aligned}$$

where n_{isk} and N_{isk} are the sample and population size of cell $i \times s \times k$, N_i is population size of region i , G is the length or the number of iterations to use after the chain reached convergence. If μ is the R04AM France mean, then similar formulas to (1) and (2) can be obtained for $\hat{\mu}$ and $V(\hat{\mu})$ by merely adding a sum indexed by i after $\sum_s \sum_k$ and by replacing N_i by the France population size denoted by N . N_{isk} , N_i and N

are those of the French 1999 census and were provided by INSEE. G was taken 6000 because we noticed that the Monte Carlo approximation was good enough for this number of iterations. The Markov chains $alpha^g$ and μ_{isk}^g of parameters $alpha$ and μ_{isk} were obtained by estimating the model using Metropolis-Hastings algorithm within Gibbs sampling. The 22 regional estimations plus the estimation for the France mean together with their standard errors are given in Table 1. For the 5 regions with extra sample and the France mean there are also the INSEE estimations computed using INSEE methodology based on the sampling design. We also asked INSEE to compute and provide the other regions estimations but to the end of the research contract we didn't get them so they are not reported in Table 1.

It is clear that the model based methodology is a good alternative to the INSEE sampling design: for all the 5 regions and the France the standard errors computed using the hierarchical model we constructed for R04AM are less than those computed by INSEE, so our estimators are better than INSEE's. In Figure 4 we plotted the coefficients of variation versus regional sample size for our estimations represented by the solid black line as well as for INSEE estimations represented by the dotted blue line. We can see that our coefficients of variation are half those of INSEE.

Table 1: Regional estimations for R04AM based on the sample with extension (39900 observations)

Region * = region with extension	Estimation	Standard Error	Estimation INSEE	Standard Error INSEE
Ile de France*	2.30	0.0299	2.30	0.0540
Champagne-Ardenne*	2.06	0.0508	1.90	0.0800
Picardie*	1.98	0.0474	2.10	0.1440
Haute-Normandie	2.05	0.0968		
Centre	1.97	0.0725		
Basse-Normandie	1.82	0.0893		
Bourgogne	1.88	0.0828		
Nord Pas de Calais*	1.78	0.0378	1.80	0.0600
Lorraine	2.08	0.0728		
Alsace	2.35	0.0939		
Franche Comté	1.97	0.0926		
Pays de la Loire	2.00	0.0588		
Bretagne	2.10	0.0674		
Poitou Charente	1.89	0.0823		
Aquitaine	2.27	0.0672		
Midi-Pyrénées	2.24	0.0809		
Limousin	1.67	0.1030		
Rhône Alpes	2.04	0.0507		
Auvergne	2.00	0.1059		
Languedoc-Roussillon	2.09	0.0784		
PACA*	2.36	0.0446	2.40	0.0760
Corse	1.57	0.2254		
France Métropolitaine	2.10	0.0162	2.10	0.0280

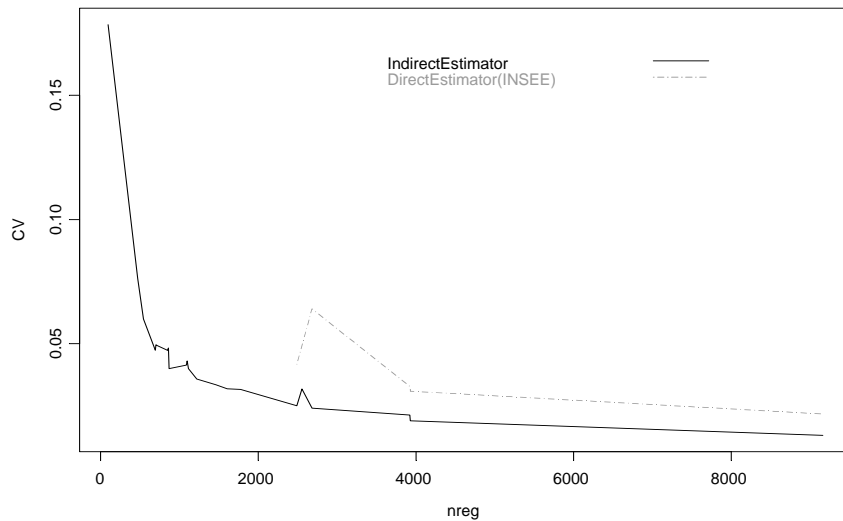


Figure 4. Coefficient of variation versus size of the regional sample size

The model we constructed with slight modifications can be used to obtain county estimations. The index i designating the region is replaced by the index c which will designate a county. Thus we introduce a county effect in our model which will allow in the same way as above the estimation of county means μ_c . The model for county estimation will be:

Model for counties

$$\begin{aligned}
 y_{cskl} | v_{cskl} &\sim \text{Poisson}(v_{cskl}), \\
 v_{cskl} | \mu_{csk}, \alpha &\sim \text{Gamma}\left(\frac{\mu_{csk}}{\alpha}, \frac{1}{\alpha}\right), \\
 \log(\mu_{csk}) &= \beta_{1c} + \beta_{3s} + \beta_{4k}, \\
 \alpha &\sim \text{Unif}(0,100), \\
 \beta_{1c} &\sim \text{Unif}(-10,10), \beta_{3s} \sim \text{Unif}(-10,10), \beta_{4k} \sim \text{Unif}(-10,10)
 \end{aligned}$$

We tested the fit of the new model concluded that it is as good as the model for regions. Then we used it in a similar manner to obtain estimations for county means together with their precisions. Due to the large number of French counties we don't present the results here. They can be found in Stefan[4]. In Figure 5 we plotted the coefficient of variation of the county means estimations versus size of the county sample. Just as we estimated earlier the R04AM France mean with the model in section 2, we will also be able with the new model to estimate the regional means and the France mean. We did it and we obtained very close values to those in Table 1.

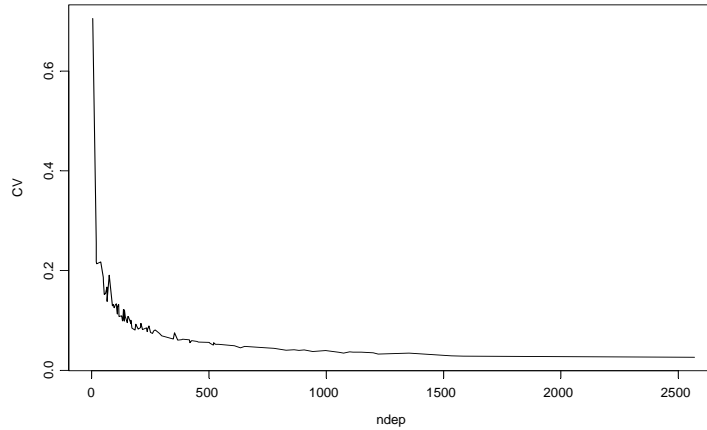


Figure 5. Coefficient of variation versus size of the county sample size

5. CONCLUSION

The objective of this paper was to construct a model on which to base the inference aimed at obtaining regional estimations for the number of visits to the generalist in the last year. We compared our estimations with INSEE design-based estimations and found that when model-based the methodology for small areas produces coefficients of variations half those of INSEE.

We considered one variable R04AM which is a count variable like the variable considered in Stefan[5]. Their methodologies were similar. In the future we will construct models for other variables in FHS (binary or continuous). For other types of variables the Poisson distribution will be replaced by Bernoulli or Binomial distributions. Over or under dispersion will have to be checked and accounted for properly.

In deriving the formulas (1) and (2) we supposed that the cell sample sizes n_{isk} are non-random. In practice this is generally not true. In the classical survey sampling theory computations using random n_{isk} are not feasible, that's why under such circumstances analyses are conditional on the realized sample sizes. In a full hierarchical Bayesian context Oleson and al.[2] proposed a model accounting for random sample sizes and also random population sample sizes. Based on their paper we will extend our present work.

Survey sampling are generally characterized by nonresponse and FHS is no exception. If not properly accounted for the nonresponse can lead to biased estimation. In our paper we supposed that there is

complete response. In fact we removed the approximately 1000 individuals that didn't provide any value for R04AM and realized our analysis on the remaining ones. Nandram and al.[1] and the references therein constitute a large literature to see how the nonresponse in FHS can be properly dealt with in a full hierarchical Bayesian context.

REFERENCES

1. NANDRAM, B., COX, L., CHOI, W.J., *Bayesian analysis of nonignorable missing categorical data: An application to bone mineral density and family income*, Survey Methodology, Vol. 31, 2, pp. 213-225, 2005
2. OLESON, J., HE, C., SUN, D., SHERIFF, S., *Bayesian estimation in small areas when sampling design strata differ from the study domains*, Survey Methodology, Vol. 33, 2, pp. 173-185, 2007
3. RAO, J.N.K., *Small area estimation*, John-Wiley & Sons, Hoboken-New Jersey, 313 pp, 2003
4. STEFAN, M., *Enquête décennale santé 2002/2003: Estimation petits domaines*, Rapport final de recherche, DREES, Paris, 121 pp., 2006
5. STEFAN, M., *Hierarchical Bayesian estimation of the number of visits to the generalist in 2002/2003 French Health Survey*, Romanian Journal of Economic Forecasting, 2, pp. 67-91, 2008

Received December 4, 2008