Folia Oeconomica
Acta Universitatis Lodziensis
ISSN 0208-6018 e-ISSN 2353-7663
www.czasopisma.uni.lodz.pl/foe/
3(335) 2018

DOI: http://dx.doi.org/10.18778/0208-6018.335.03

Michał Bernardelli
SGH Warsaw School of Economics, Institute of Econometrics, Probabilistic Methods Unit,
mbernard@sgh.waw.pl

Barbara Kowalczyk
SGH Warsaw School of Economics, Institute of Econometrics, Mathematical Statistics Unit,
bkowal@sgh.waw.pl

Optimal Allocation of the Sample in the Poisson Item Count Technique

Abstract: Indirect methods of questioning are of utmost importance when dealing with sensitive questions. This paper refers to the new indirect method introduced by Tian et al. (2014) and examines the optimal allocation of the sample to control and treatment groups. If determining the optimal allocation is based on the variance formula for the method of moments (difference in means) estimator of the sensitive proportion, the solution is quite straightforward and was given in Tian et al. (2014). However, maximum likelihood (ML) estimation is known from much better properties, therefore determining the optimal allocation based on ML estimators has more practical importance. This problem is nontrivial because in the Poisson item count technique the study sensitive variable is a latent one and is not directly observable. Thus ML estimation is carried out by using the expectation-maximisation (EM) algorithm and therefore an explicit analytical formula for the variance of the ML estimator of the sensitive proportion is not obtained. To determine the optimal allocation of the sample based on ML estimation, comprehensive Monte Carlo simulations and the EM algorithm have been employed.

Keywords: optimal allocation, latent variable, EM algorithm, sensitive question, indirect questioning, Poisson item count technique

JEL: C83, C60

1. Introduction

When dealing with sensitive attributes, indirect methods of questioning play a major role in statistical practice (Imai, 2011; Kuha, Jackson, 2014; Tourangeau, Yan, 2007; Wolter, Laier, 2014). They are designed to ensure the privacy of respondents so that it is impossible to know their answers to sensitive questions, i.e. questions about tax evasions, atypical sexual behaviours, bribes, etc. This paper refers to the new indirect method introduced by Tian et al. (2014) who propose to randomly assign each respondent in the sample to either control or treatment groups and apply the following procedure. In the control group, respondents are asked one neutral question, e.g.: How many times did you go to the cinema last month? Their answers may take values 0, 1, 2, …. In the treatment group, respondents are asked two questions, one being the same as in the control group and the other a sensitive one, e.g.: How many times did you go to the cinema last month? Did you buy any smuggled alcohol last month? Respondents are asked to assign 1 if their answer to the sensitive question is *yes* and 0 if *no*. Then they are asked to report only the sum of their answers to these two questions without revealing their answers to individual questions.

When designing an indirect survey with control and treatment groups, an important question arises how to allocate the sample size into proper groups. The optimal allocation based on the variance formula for the moment estimator of the sensitive proportion was analysed in Tian et al. (2014), which resulted in the conclusion that a balanced sample is a reasonable choice. In this paper, we analyse the optimal allocation of the sample size based on ML estimation. The justification for another approach is the fact that the maximum likelihood estimator has much better properties and is more desirable from the practical point of view. The optimal allocation based on ML estimation is at the same time more difficult due to the fact that in the proposed technique the study sensitive variable is a latent one and is not directly observable. Thus ML estimation has to be carried out by using the appropriate numerical algorithm, conventionally the EM algorithm, and an explicit variance formula for the ML estimator of the sensitive proportion is not available. Therefore, to determine the optimal allocation of the sample based on ML estimation, comprehensive Monte Carlo simulations have been employed. To facilitate the discussion, the presented paper focuses on the Poisson distribution of X.

In section 2, the mathematical background of the Poisson item count technique introduced in Tian et. al (2014) is briefly presented. Section 3 discusses the problem of optimal allocation of the sample size to control and treatment groups based on ML estimation and provides a detailed description of the numerical experiment regarding Monte Carlo simulations and technical aspects referring to the implementation of EM algorithm. In section 4, results of the comprehensive simulation study are presented. The article ends with the conclusion in Section 5.

2. Poisson item count technique

In this section, for the purpose of further discussion and analysis, we briefly present basic results for the Poisson item count technique (ICT) obtained in Tian et. al (2014). Let n_1 be the number of elements in the control group, n_2 the number of elements in the treatment group, and let $n_1 + n_2 = n$. In the Poison ICT, we have: $XPoisson(\lambda)$, $ZBernoulli(\pi)$ where X and Z are independent. In the control group, we observe $X_1, \ldots, X_{n1}$, whereas in the treatment group $Y_1, \ldots, Y_{n2}$, where $Y_j = X_{n1+j} + Z_j$ for $j = 1, 2, \ldots, n_2$. Z is a latent variable and is not directly observable in this model. The moments (difference in means) estimator of the unknown sensitive proportion π is a common difference in means estimator:

$$\hat{\pi}_M = \bar{Y} - \bar{X} = \frac{1}{n_2}\sum_{j=1}^{n_2} Y_j - \frac{1}{n_1}\sum_{j=1}^{n_1} X_j \tag{1}$$

with variance:

$$D^2(\hat{\pi}_M) = \frac{\lambda + \pi(1-\pi)}{n_2} + \frac{\lambda}{n_1} = \frac{\lambda + \pi(1-\pi)}{n_2} + \frac{\lambda}{n - n_2}. \tag{2}$$

ML estimators of model parameters π and λ are obtained via the iterative expectation-maximisation (EM) algorithm through classic E and M steps.

E step (iteration $t + 1$):

$$z_j^{(t)} = E\left(Z_j | Y_{obs}, \hat{\pi}^{(t)}, \hat{\lambda}^{(t)}\right) = \frac{y_j \hat{\pi}^{(t)}}{y_j \hat{\pi}^{(t)} + \hat{\lambda}^{(t)}\left(1 - \hat{\pi}^{(t)}\right)}, \; j = 1, \ldots, n_2. \tag{3}$$

M step (iteration $t + 1$):

$$\hat{\pi}^{(t+1)} = \frac{1}{n_2}\sum_{j=1}^{n_2} z_j^{(t)}, \tag{4}$$

$$\hat{\lambda}^{(t+1)} = \frac{1}{n_1 + n_2}\left(\sum_{i=1}^{n_1} x_i + \sum_{j=1}^{n_2}\left(y_j - z_j^{(t)}\right)\right). \tag{5}$$

When the variance formula for the method of moments estimator of the sensitive proportion is used, the problem of optimal allocation of the sample is quite simple and the optimal allocation n_2^{opt} can be obtained straightforwardly by minimising formula (2), which was done in Tian et al. (2014):

$$\frac{n_2}{n} = \frac{1}{1 + \sqrt{\frac{\lambda}{\lambda + \pi(1-\pi)}}}. \tag{6}$$

Of course, model parameters λ, π are unknown in advance, thus some compromise over possible various values of λ, π has to be implemented. Tian et. al (2014) proposed a balanced sample as a reasonable choice.

Although ML estimation via the EM algorithm (3)–(5) is strongly advocated in the work of Tian et al. (2014), no particular analysis for the optimal allocation of the sample size based on ML estimation via the EM algorithm is conducted in the original paper. This analysis will be provided in the next sections of the presented paper.

3. Problem of optimal allocation based on ML estimation

In this section, the problem of optimal allocation based on ML estimation is explained and a description of the proposed solution to this problem is presented. Having fixed a total sample size $n = n_1 + n_2$, one should seek the optimal allocation of the sample to treatment and control groups, i.e. such allocation n_2 and $n_1 = n - n_2$, at which the highest efficiency of the estimation is achieved.

For large sample sizes, it is common for practitioners to prefer a simple method of moments estimator given by formula (1), which does not need any iterative numerical algorithm, as opposite to ML estimator given by iterative formulas (3)–(5). For large sample sizes, variances of both estimators will be similar. Thus, deriving the optimal allocation based on variance formula (2) for moments estimator (1), which was done in Tian et al. (2014), is quite practical for large samples. This, however, does not apply to the non-asymptotic case. Let us keep in mind the fact that for moderate sample sizes probability that moments estimator (1) goes beyond interval [0, 1] is quite high, and additionally the variance of the method of moments estimator for the moderate sample size is visibly higher than the variance of ML estimator. Therefore, for practitioners, the most interesting question, which has not been answered till now, is how to allocate the sample size into treatment and control groups for ML estimators in the non-asymptotic case, i.e. for the moderate sample size. Of course, another question arises next: does this optimal allocation for ML estimators differ from that for moments estimators? These two questions will be answered in the presented paper.

Due to the fact that the ML estimator in the Poisson ICT can be obtained only via some iterative algorithm and its variance is not known, we base our analysis

on a Monte Carlo simulation study. Let us notice that for the sample size n we have $n-1$ different possible allocations, assuming non-empty control and treatment groups. In particular: 1st allocation: $n_2 = 1$, $n_1 = n-1$, 2nd allocation: $n_2 = 2$, $n_1 = n-2$, 3rd allocation: $n_2 = 3$, $n_1 = n-3$,, etc. The last possible allocation is $n_2 = n-1$, $n_1 = 1$. Therefore, the problem of determination of optimal allocation is equivalent to the problem of comparison of $n-1$ different ML estimators $\hat{\pi}_{ML}(n_2)$, $n_2 = 1, 2, \ldots, n-1$ based on n element overall sample size with different allocations. Having obtained the best estimator $\hat{\pi}_{ML}^{opt} = \hat{\pi}_{ML}(n_2^{opt})$, i.e. the estimator with the smallest MSE, the optimal allocation n_2^{opt} is determined straightforwardly.

Due to the fact that we do not know model parameters π, λ in advance, some compromise among different π, λ should be made. Let us denote by *RMSE* the root mean square error of the estimator of the sensitive proportion π, and let $RMSE^{opt}$ denote its minimum value taken over all estimators based on the same overall sample size n with different allocations to control n_1 and treatment groups n_2, $n_1 + n_2 = n$. For practitioners, it is of main interest to obtain an allocation under which, for all possible model parameters, the absolute maximum acceptable distance Δ from $RMSE^{opt}$ is not exceeded. In a simulation experiment, this can be accomplished by first obtaining, under given π, λ, interval $\left[n_{2,min,\lambda,\pi}^{\Delta}, n_{2,max,\lambda,\pi}^{\Delta}\right]$ such that if $n_2 \in \left[n_{2,min,\lambda,\pi}^{\Delta}, n_{2,max,\lambda,\pi}^{\Delta}\right]$, then $|RMSE(n_2, \lambda, \pi) - RMSE^{opt}(\lambda, \pi)| < \Delta$. The final interval can be attained by taking common parts of obtained intervals over all considered parameters from the possible set of values $\pi \in (0, 1)$ and $\lambda > 0$:

$$\left[n_{2,\min}^{\Delta}, n_{2,\max}^{\Delta}\right] = \bigcap_{\pi,\lambda} \left[n_{2,\min,\lambda,\pi}^{\Delta}, n_{2,\max,\lambda,\pi}^{\Delta}\right]. \tag{7}$$

The same can be done by determining the relative maximum acceptable distance δ. Analogously, first we obtain, under given π, λ, interval $\left[n_{2,min,\lambda,\pi}^{\delta}, n_{2,max,\lambda,\pi}^{\delta}\right]$ such that if $n_2 \in \left[n_{2,min,\lambda,\pi}^{\delta}, n_{2,max,\lambda,\pi}^{\delta}\right]$, then $\left|\frac{RMSE(n_2, \lambda, \pi) - RMSE^{opt}(\lambda, \pi)}{RMSE^{opt}(\lambda, \pi)}\right| < \delta$.

Next common parts are taken of obtained intervals over all considered parameters from the possible set of values $\pi \in (0, 1)$ and $\lambda > 0$:

$$\left[n_{2,\min}^{\delta}, n_{2,\max}^{\delta}\right] = \bigcap_{\pi,\lambda} \left[n_{2,\min,\lambda,\pi}^{\delta}, n_{2,\max,\lambda,\pi}^{\delta}\right]. \tag{8}$$

For the Monte Carlo (MC) simulation experiment associated with the iterative EM algorithm conducted to determine the optimal allocation, the following simulation parameters were assumed:

1) model parameters:
   $\lambda \in \{1.5; 2; 2.5\}$, $\pi \in \{0.05; 0.10; 0.15; 0.20; 0.25; 0.30; 0.35; 0.40\}$,
2) sample size $n = 200$,

3) number of replications in the MC experiment: 35 000–60 000,
4) maximal number of iteration in the EM algorithm: 50 000,
5) accuracy of calculations: 1e–10.

The justification for the model parameters choice is the following. As the problem deals with sensitive features only, π is commonly assumed to be less than 0.5 in the literature (see Imai, 2011; Tian et al., 2014; Kowalczyk, Wieczorkowski, 2017). To protect respondents' privacy properly and at the same time to ensure the not too high estimation error, Tian et al. (2014) state that a good choice of λ is 2 (see Tian et al., 2014). In practice, we cannot predict respondents' answers in advance, thus values of λ around 2 are also taken into account in simulations. Therefore, in summary, 24 series of simulations are performed for various sets of model parameters. In each series, 199 approximations, due to the given tolerance, of ML estimators are obtained separately. More precisely, 199 approximations of ML estimators based on overall $n = 200$ sample size with $n_2 = 1, 2, \ldots, 199$ treatment group sizes. All estimators are computed using the presented in section 2 iterative formulas for the EM algorithm, introduced in Tian et al. (2014). The following stopping criteria are used – reaching the maximum number of iterations (50 000) or the lower bound in the change in the value of the model parameters in two consecutive iterations (tolerance = 1e–10). In each case, between 35 000 and 60 000 replications (Monte Carlo iterations) are used, depending on the stability of the computations.

4. Simulation results

The simulation results obtained for all sets of model parameters are consistent. They are presented below in three parts. Firstly, some basic comparison between the method of moments and maximum likelihood estimators is provided. In the second part, a more detailed discussion concerning numerical results for the ML estimation is presented. The third part consists of the summary of results for all the considered sets of model parameters.

In all the comparisons given below, results for ML estimators are obtained based on the conducted series of comprehensive simulation studies. And results for the method of moments estimators are obtained by using formula (2).

To illustrate the evident difference between properties of moments and ML estimators of the sensitive proportion π, the ranking of the estimators based on their $RMSE(n_2; \lambda, \pi)$ was computed and presented graphically. The ranking position is on the x-axis and the size of the corresponding treatment group n_2 on the y-axis. The exemplary graph for $\lambda = 2.5$ and $\pi = 0.3$ is presented in Figure 1. Additional graphs are presented in the appendix.

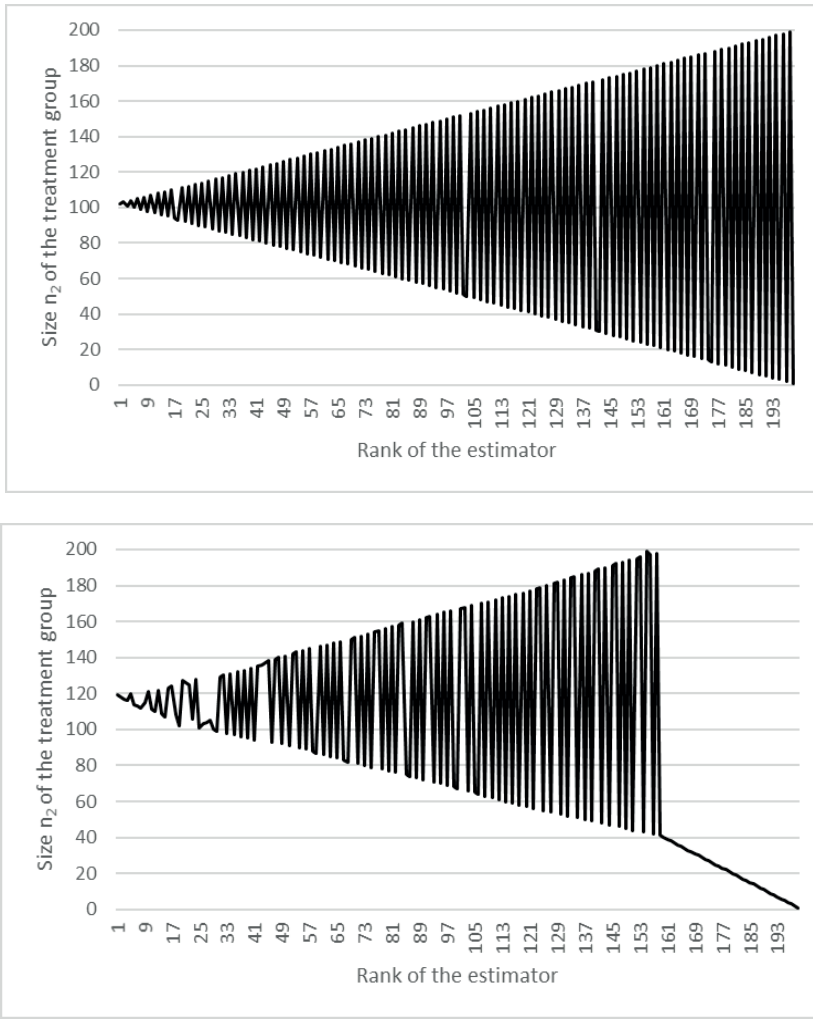


Figure 1. Relationship between size n_2 and the rank of the estimator of the parameter π regarding its RMSE for moments (upper) and ML (lower) estimators, λ = 2.5 and π = 0.3. Figures for other sets of model parameters λ and π are presented in the appendix

Source: own calculations

Based on the presented graphs, at least two key conclusions should be stated. First, the optimum value of n_2 computed based on the variance formula for the method of moments estimator is visibly lower (the size of the treatment group around 100) than the optimum value of n_2 computed for the ML estimator (the size of the treatment group visibly above 100). Secondly, there is a clear symmetry for the moments estimators ranking, in contrast to the ML estimators. It means that choosing a too small or too large size of the treatment group when using moments estimation is approximately equally bad in terms of the obtained RMSE. Where-

as in the case of ML estimators, too small treatment groups are more dangerous in terms of efficiency of the estimation as compared to too small control groups (specific ends of all graphs for ML estimators). This can be of particular interest for surveys with a different unit cost in control and treatment groups.

Next, apart from the ranking itself, also exact values of the obtained RMSE for the two types of estimators of the sensitive proportion are presented. For definiteness, results for RMSE are illustrated graphically in Figure 2 for one exemplary set of model parameters.

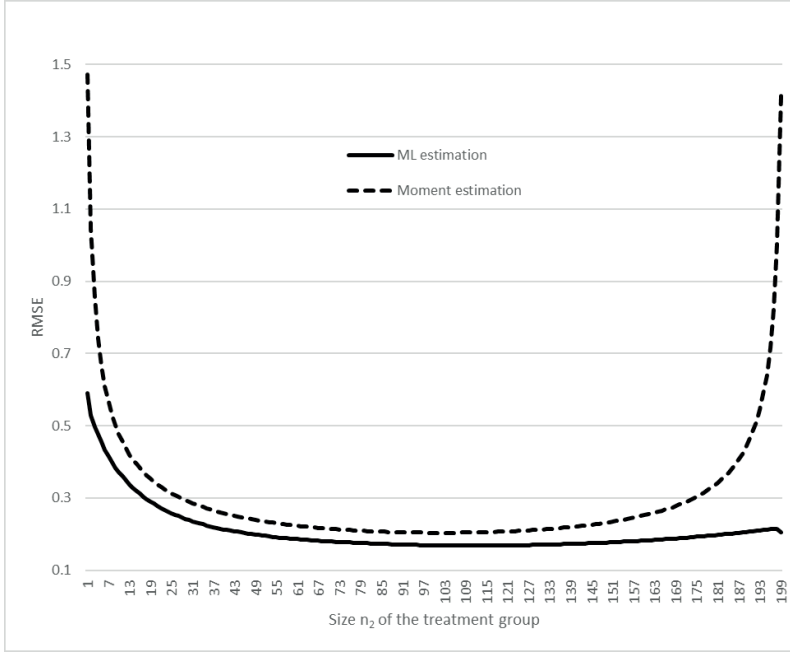


Figure 2. Root mean square error (RMSE) of the MM (dotted line) and ML (solid line) estimators of the sensitive proportion π for $\lambda = 2$, $\pi = 0.2$

Source: own calculations

Let us notice that enormously large values of the RMSE of the estimators of the parameter π are obtained in cases of extremally small values of n_2 for the ML estimator and extremely small values of n_2 and $n - n_2$ for the MM estimator. What is further important, in the classic graph presented above, it looks like both of the lines in Figure 2 are smooth. The magnification of the RMSE line around its optimum (the lowest value) for the ML estimators obtained in the simulation study is presented in Figure 3.

Results of iterative algorithms are only approximations of the optimal solutions. Depending on the algorithm, the obtained solution could be only a local, not global extremum. Therefore, based on the graph from Figure 3, we cannot

give the exact solution. However, given the values on the *y*-axis, differences between the values of the obtained approximation and the actual solution are extremely small and do not have practical importance. All values in the neighbourhood of the solution pointed by the EM algorithm differ in the fourth decimal place. From the practical point of view then, it is approximately equally good if we choose the size of the treatment group equal to the computed solution or close to it. It is also possible to use some smoothing of the results given by the Monte Carlo simulations and the EM algorithm. Using different kinds of interpolation methods, one may get various results. For example, using the quadratic interpolation, which seems to be adequate taking into consideration the presumed monotonicity of the function (function convex up), we can choose the approximation of the solution as the middle node of the interpolation and equidistant two other points. An example of such a 32 width interval interpolation performed on the data from Figure 3 is presented in Figure 4. In this particular example, the solution given by the MC simulations and the EM algorithm was 112, but after interpolation, the result changed to 110.

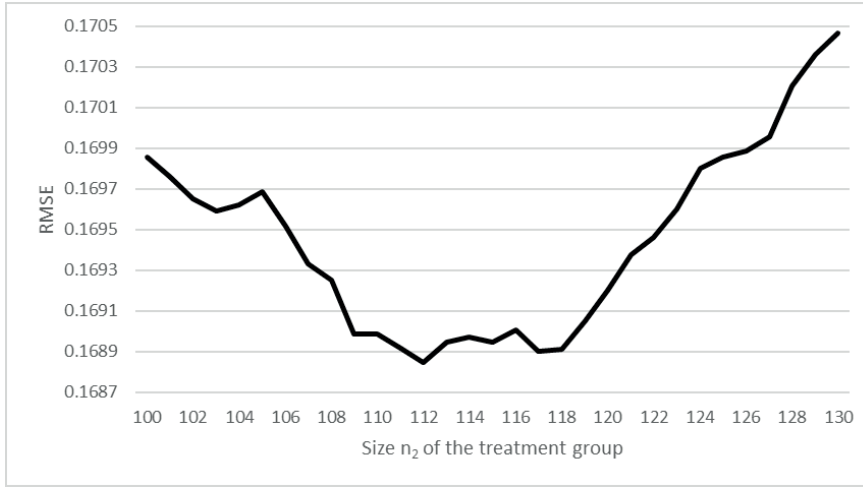


Figure 3. The fragment of the root mean square error of ML estimators for $\lambda = 2$ and $\pi = 0.2$

Source: own calculations

Monte Carlo simulations results for all the considered sets of model parameters are given in Table 1. For more detailed comparisons, values of optimal sizes computed based on the variance formula for moment estimators as well as the results of the quadratic interpolation (based on points in optimum given by MC simulations, 16 lower and 16 greater) are also presented. It is clear that the optimal allocation for ML estimators (before and after interpolation) gives, in general, larger sizes of treatment group sizes than the optimal allocation for the moment estimators.

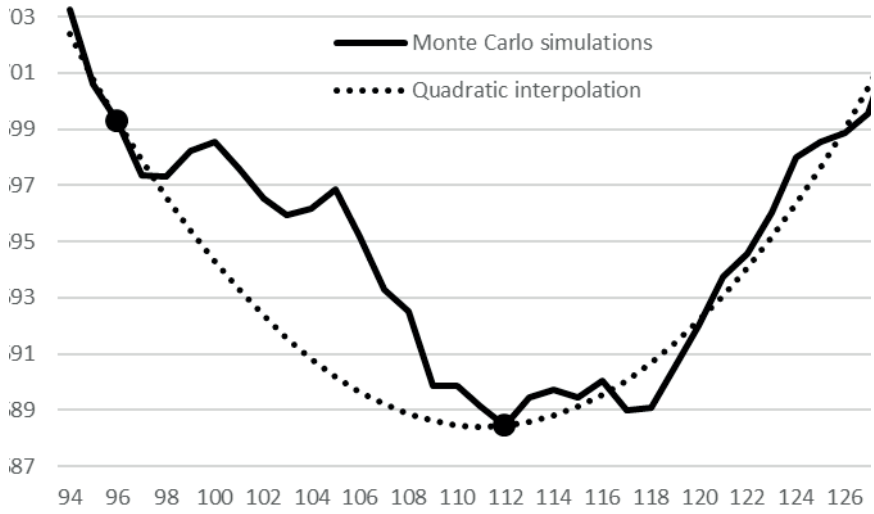


Figure 4. Exemplary quadratic interpolation (dotted line) on the fragment of the root mean square error of ML estimators (solid line) for $\lambda = 2$ and $\pi = 0.2$ based on the results corresponding to the treatment group sizes: 96, 112 and 128 (black dots)

Source: own calculations

Table 1. The optimal allocation of the treatment group size n_2 for ML estimators based on MC simulation results and interpolation (quadratic, 32 width interval) juxtaposed with the optimal allocation for moment estimators

$\lambda = 1.5$								
π	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
ML simulation	110	109	114	111	114	112	116	119
quadratic interpolation	105	108	111	114	115	116	119	119
moment estimation	101	101	102	103	103	103	104	104
$\lambda = 2.0$								
π	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
ML simulation	108	108	106	112	111	113	113	122
quadratic interpolation	107	108	108	110	112	114	114	118
moment estimation	101	101	102	102	102	102	103	103
$\lambda = 2.5$								
π	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
ML simulation	110	107	105	114	117	119	112	117
quadratic interpolation	105	105	117	115	115	113	115	116
moment estimation	100	101	101	102	102	102	102	102

Source: own calculations

The optimal allocation intervals obtained according to formulas (7) and (8) are calculated for both considered classes of estimators. Results are given in Table 2. It needs to be emphasised that intervals obtained based on the variance formula

for the method of moments estimators have in general smaller values than those obtained for maximum likelihood estimators. For the smallest value of absolute tolerance $\Delta = 0.001$, calculated intervals for different types of estimators are disjoint events. Additionally, for small values of tolerance, the balanced sample with $n_2 = 100$ is not even within the obtained range for ML estimators. The resulting conclusion for practitioners is that it is better to allocate a slightly larger part of the sample to the treatment group, and not to use the perfectly balanced sample, as far as maximum likelihood estimation is taken into account.

Table 2. The optimal allocation of the treatment group size n_2 with respect to the given absolute Δ and percentage δ tolerance over possible sets of model parameters

Tolerance Δ	**Moment estimation** $[n^{\Delta}_{2,min}, n^{\Delta}_{2,max}]$	**ML estimation** $[n^{\Delta}_{2,min}, n^{\Delta}_{2,max}]$	**ML estimation after interpolation** $[n^{\Delta}_{2,min}, n^{\Delta}_{2,max}]$
0.001	**[94, 109]**	**[110, 115]**	**[108, 115]**
0.0025	[88, 115]	[103, 123]	[102, 121]
0.005	[81, 121]	[93, 131]	[102, 122]
0.01	[72, 129]	[83, 142]	[102, 122]
Tolerance δ	**Moment estimation** $[n^{\delta}_{2,min}, n^{\delta}_{2,max}]$	**ML estimation** $[n^{\delta}_{2,min}, n^{\delta}_{2,max}]$	**ML estimation after interpolation** $[n^{\delta}_{2,min}, n^{\delta}_{2,max}]$
1%	[90, 114]	[107, 119]	[106, 118]
2.5%	[82, 122]	[98, 128]	[102, 122]
5%	[73, 130]	[87, 138]	[102, 122]
10%	[62, 142]	[72, 150]	[102, 122]

Source: own calculations

5. Conclusions

For a moderate sample size, the optimal allocation of the sample in the Poison item count technique based on the minimisation of the variance formula for the method of moments estimator differs quite visibly from the optimal allocation obtained for the ML estimator. The balanced sample proposed in Tian et al. (2014) is not the best choice in terms of maximum likelihood estimation. As far as the ML estimator is concerned, it is better to allocate a slightly larger part of the sample to the treatment group. In the future research, it is reasonable to develop an asymptotic theoretical optimal allocation for ML estimators and analyse the rate of convergence to the asymptotic result.

References

Imai K. (2011), *Multivariate regression analysis for the item count technique*, "Journal of the American Statistical Association", vol. 106, no. 494, pp. 407–416.

Kowalczyk B., Wieczorkowski R. (2017), *Comparing proportions of sensitive items in two populations when using Poisson and negative binomial item count techniques*, "Quantitative Methods in Economics", vol. 18, no. 1, pp. 68–77.

Kuha J., Jackson J. (2014), *The item count method for sensitive survey questions: modeling criminal behavior*, "Journal of the Royal Statistical Society: Series C", vol. 63, no. 2, pp. 321–341.

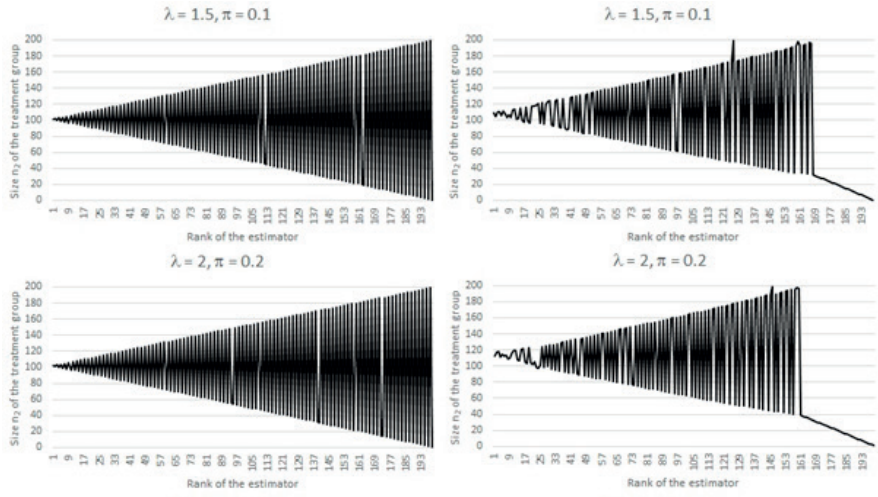
Tian G-L., Tang M-L., Wu Q., Liu Y. (2014), *Poisson and negative binomial item count techniques for surveys with sensitive question*, "Statistical Methods in Medical Research", Pre-published online on December 16, 2014, http://dx.doi.org/10.1177/0962280214563345.

Tourangeau R., Yan T. (2007), *Sensitive questions in surveys*, "Psychological Bulletin", vol. 133, no. 5, pp. 859–883.

Wolter F., Laier B. (2014), *The Effectiveness of the Item Count Technique in Eliciting Valid Answers to Sensitive Questions. An Evaluation in the Context of Self-Reported Delinquency*, "Survey Research Methods", vol. 8, no. 3, pp. 153–168.

Appendix

Relationship between size n_2 and rank of the estimator regarding its RMSE for the MM (left) and ML (right) estimators for different sets of model parameters.



Optymalna alokacja próby w badaniu cechy drażliwej

Streszczenie: Pośrednie metody ankietowania stanowią podstawowe narzędzie stosowane w przypadku pytań drażliwych. Artykuł nawiązuje do nowej, pośredniej metody zaproponowanej w pracy Tiana i wsp. (2014) i dotyczy optymalnej alokacji próby między grupę badaną i kontrolną. W przypadku gdy alokacji dokonuje się w oparciu o estymatory metodą momentów, rozwiązanie optymalne nie nastręcza trudności i zostało podane w pracy Tiana i wsp. (2014). Jednak to estymacja metodą największej wiarogodności ma lepsze własności, w związku z czym wyznaczenie alokacji optymalnej na jej podstawie jest zadaniem, którego rozwiązanie wydaje się mieć większe znaczenie praktyczne. Zadanie to nie jest trywialne, gdyż w przypadku omawianej metody pośredniej drażliwa zmienna badana ma charakter ukryty i jest zmienną nieobserwowalną. Wzór *explicite* na wariancję estymatora największej wiarogodności nieznanej frakcji cechy drażliwej nie jest dostępny, a sam estymator wyznaczyć można, używając odpowiednich algorytmów numerycznych. Do określenia optymalnej alokacji próby w oparciu o estymatory NW wykorzystane zostały symulacje Monte Carlo oraz iteracyjny algorytm EM.

Słowa kluczowe: alokacja optymalna, zmienna ukryta, algorytm EM, cecha drażliwa, pytania pośrednie, eksperyment z listą

JEL: C83, C60

Received: 2016-12-17; verified: 2017-11-29. Accepted: 2018-01-29