Journal of Mathematics and Statistics 1 (4): 291-295, 2005
ISSN 1549-3644

Riemann Estimation for Replicated Environmental Sampling Designs

Lucio Barabesi and Marzia Marcheselli
Dipartimento di Metodi Quantitativi, Università di Siena
P.zza S.Francesco 17, 53100 Siena, Italy

Abstract: In many environmental surveys the population under study is made up of biological units scattered over a planar region. A variable is considered on each unit and the target parameter generally turns out to be the population total of the variable. In order to estimate the population total, field scientists commonly replicate a suitable design on the study region. Replicated environmental designs basically rely on the selection of a set of sample points, in such a way that each sample point corresponds to a single design replicate. Frequently, the sample points are located uniformly and independently over the planar region, even if more effective strategies are actually available. The population total is subsequently estimated by using the mean of the estimates obtained in each design replicate. However, this pooled estimator may be improved by considering a suitable weighted mean - rather than the simple mean - of the estimates. Thus, we propose a Riemann estimator of the population total which is actually borrowed from the Monte Carlo integration setting. The suggested estimator displays appealing performance from both theoretical and practical perspectives.

Key words: Replicated sampling design, continuous population, Riemann Monte Carlo estimator

INTRODUCTION

The aim of many quantitative environmental and agricultural studies is to estimate the total of a variable in the considered population. In order to collect the sampling information, some replicates of a suitable environmental design are carried out in the field[1]. For example, in the forestry setting, replicated line-intercept sampling is commonly adopted to estimate the canopy coverage in a delineated region[2,3], while replicated Bitterlich sampling is considered when the total basal area in a forest is the target parameter[4]. In ecological studies, replicated plot sampling is used to estimate species composition and density[4].

It is worth noting that the designs arising in environmental studies may be embedded in a unique theoretical framework, *i.e*. the “continuous population” paradigm[5,6]. Under this approach, the design is carried out by selecting a point on a *continuum* such as a portion of a straight line or a finite planar region. Actually, as pointed out by Barabesi and Pisani[7], practical environmental designs may be partitioned into two large families: the first family encompasses designs which are implemented by selecting a point on the baseline (the projection of the study region onto a line of arbitrary direction), while the second family includes designs which are carried out by selecting a point on the whole study region. In the present study we exclusively focus on the first family of designs, which in any case comprises many important practical designs such as line-intercept sampling[8], strip sampling[1], line-transect sampling under Burnham-Anderson's detection model[8] and line-transect sampling under Hayne's detection model[9], among others. Hence, once a suitable design is chosen, n replicates of the design are performed in the field, *i.e*. in this case n sample points are selected on the baseline. The usual population total estimate is simply obtained by averaging the n estimates obtained in the n design replicates. Therefore, in order to achieve accurate population total estimation, the focus boils down to the optimal placement of the n sample points.

Barabesi[5,10] has shown the equivalence of the strategies adopted for the placement of the sample points either in replicated designs or in Monte Carlo integration. Indeed, under the continuous population paradigm the population total may be represented as the integral of a certain function which depends on the chosen design. Thus, an optimal Monte Carlo integration strategy may be adopted in order to select the sample points in replicated designs. The modified Monte Carlo integration method introduced by Haber[11,12] is a highly suitable strategy. This Monte Carlo integration method involves partitioning the baseline into n equal segments and generating n independent random points in these segments. From an environmental sampling perspective the strategy is basically the so-called nonaligned systematic sampling of points suggested in the U.S. EPA QA/G-5S Guidance[13]. When this strategy is considered, Barabesi and Marcheselli[14,15] have shown that very accurate estimators - even displaying a $O(n^{-3})$ variance rate - may be achieved.

Corresponding Author: Lucio Barabesi, Dipartimento di Metodi Quantitativi, Università di Siena, P.zza S.Francesco 17, 53100 Siena, Italy

Unfortunately, the data are often collected in the field by means of independent sample points uniformly placed over the baseline[8]. This sampling strategy is equivalent to the crude Monte Carlo integration method, which solely produces a population total estimator with a $O(n^{-1})$ variance rate. However, it is again possible to achieve accurate estimation by adopting the Riemann Monte Carlo estimators[16]. The suggested Riemann Monte Carlo estimator of the population total is based on the weighted mean - rather than the simple mean - of the estimates obtained in the n replicates. Even if the Riemann Monte Carlo estimator is biased, it displays a $O(n^{-2})$ mean square error and hence it improves over the simple mean.

PRELIMINARIES

Let us consider a well-defined planar study region and a population of N units scattered over this region at fixed locations. Furthermore, let $(y_1, y_2, \dots, y_N)$ be the values of the target variable on the N units, in such a way that

$$T_y = \sum_{l=1}^{N} y_l$$

represents the population total. Moreover, let us assume that the estimation of T_y is performed using the replication of a design which is implemented by selecting a sample point on the baseline. Without loss of generality and for the sake of simplicity, let us suppose that the baseline is given by the interval $(0,1)$. Moreover, let u be the position of a point selected on the baseline.

Once a suitable design is chosen, the inclusion set P_l of the l-th unit is a suitable interval contained in the baseline[10] and the l-th unit is selected - and y_l is measured - if $u \in P_l$. As an example, let us consider a population of plants and let y_l be the biomass of the l-th plant, in such a way that the target parameter is the total biomass in the forest. If line intercept sampling is adopted, the inclusion sets are the projections of the plant crowns onto the baseline. Indeed, a plant is selected if the corresponding crown is intercepted by a line perpendicular to the baseline at location u.

In order to obtain a suitable representation for T_y, it is worth noting that, if solely the l-th unit is considered, the intensity of the target variable over the l-th inclusion set is y_l / π_l, where $\pi_l = \int_0^1 I_{\{u \in P_l\}} du$ is the length of P_l and I_A is the usual indicator of a set A. Hence, the intensity of the variable at location u is given by

$$y(u) = \sum_{l=1}^{N} \frac{y_l}{\pi_l} I_{\{u \in P_l\}} \,. \qquad (1)$$

Incidentally, it is at once apparent that $y(u)$ is simply the Horvitz-Thompson estimate of T_y when location u is selected. The total intensity of the variable of interest over the study region turns out to be

$$\int_0^1 y(u)\, du = \int_0^1 \sum_{l=1}^{N} \frac{y_l}{\pi_l} I_{\{u \in P_l\}}\, du = T_y \,,$$

i.e. an integral representation is achieved and hence the estimation of T_y reduces to an integration problem. Hence, the strategies adopted for the Monte Carlo quadrature of an integral can be used in order to choose n sample points $(u_1, u_2, \dots, u_n)$ on the baseline. When the n sample points are independently and randomly chosen, the crude Monte Carlo integration strategy is actually adopted. In this case, $(u_1, u_2, \dots, u_n)$ are the realization of n independent random variables $(U_1, U_2, \dots, U_n)$ uniformly distributed over the baseline. Hence, the Monte Carlo estimator is given by

$$\overline{T}_y = \overline{T}_y(U_1, U_2, \dots, U_n) = \frac{1}{n} \sum_{i=1}^{n} y(U_i) \,. \qquad (2)$$

It is at once apparent that the pooled estimator (2) is the mean of the n Horvitz-Thompson estimators corresponding to the n design replicates. Actually, this is the usual estimation procedure adopted in replicated environmental sampling designs. Indeed, once the sample points are positioned and the data are collected, n Horvitz-Thompson estimates are obtained for each design replicate and they are subsequently averaged in order to achieve an overall estimate for T_y. Obviously, in this section the same procedure has been described from a Monte Carlo integration perspective.

It is straightforward to show that $\overline{T}_y$ is unbiased with a $O(n^{-1})$ variance rate. In addition, $\text{Var}[\overline{T}_y]$ may be unbiasedly estimated by means of

$$\widehat{\text{Var}}[\overline{T}_y] = \frac{1}{n(n-1)} \sum_{i=1}^{n} [y(U_i) - \overline{T}_y]^2 \,.$$

However, the crude Monte Carlo strategy precludes the small variance rates for the pooled estimator which can be achieved when more refined Monte Carlo strategies are adopted[14,15]. Accordingly, the aim of the following section is to introduce an estimator with elevated performance, even if the sample points are collected using the crude Monte Carlo strategy.

THE RIEMANN MONTE CARLO ESTIMATOR

In order to improve on estimator (2), it is worthwhile to consider a weighted estimator of type

$$\begin{aligned}\overline{\overline{T}}_y &= \overline{\overline{T}}(U_1,U_2,\dots,U_n)\\ &= \sum_{i=1}^{n} w_i(U_1,U_2,\dots,U_n)\,y(U_i),\end{aligned} \qquad (3)$$

where the $w_i(U_1,U_2,\dots,U_n)$ s are positive weights such that $\sum_{i=1}^{n} w_i(U_1,U_2,\dots,U_n)=1$. If the w_i s are non-random, *i.e.* $w_i(U_1,U_2,\dots,U_n)=w_i$, it is straightforward to prove that estimator (3) is unbiased and the corresponding variance is minimum when $w_i = 1/n$. In this case, estimator (3) actually reduces to estimator (2). Accordingly, the weights must be chosen as random functions in order to improve on estimator (2). First, it is at once apparent that estimator (3) may expressed as

$$\overline{\overline{T}}_y = \sum_{i=1}^{n} w_i(U_{(1)},U_{(2)},\dots,U_{(n)})\,y(U_{(i)}),$$

where $(U_{(1)},U_{(2)},\dots,U_{(n)})$ represents the order statistic corresponding to $(U_1,U_2,\dots,U_n)$. Hence, a reasonable choice of the weights is given by $w_i(U_{(1)},U_{(2)},\dots,U_{(n)}) = U_{(i+1)} - U_{(i)}$, assuming that $U_{(0)} = 0$ and $U_{(n+1)} = 1$. In this case, estimator (3) reduces to the usual Riemann Monte Carlo estimator[16] given by

$$\overline{\overline{T}}_y = \sum_{i=0}^{n} (U_{(i+1)} - U_{(i)})\,y(U_{(i)}). \qquad (4)$$

Robert and Casella[16] have proven that (4) is biased with a $O(n^{-2})$ mean square error when y has a bounded derivative. However, in the present setting y does not achieve this regularity condition. Indeed, it is at once apparent that the function y defined in (1) is an elementary function. In any case, it can be proven that $\overline{\overline{T}}_y$ is biased with a $O(n^{-2})$ mean square error even if y is solely an elementary function (Result 1 in the Appendix). Hence, estimator (4) is preferable to estimator (2), at least in a large-sample setting. Moreover, $\overline{\overline{T}}_y$ generally displays a $O(n^{-1})$ bias (see the Remark in the Appendix). However, on the basis of the same Remark, if the sampling design is slightly modified to ensure that y is null in the narrow interval $(1-\varepsilon,1]$ (where ε is a small positive constant), then $\overline{\overline{T}}_y$ achieves a $o(n^{-1})$ bias. This requirement may be easily achieved in practice by suitably modifying the inclusion probability of the right-border units.

As to the variance estimation of $\overline{\overline{T}}_y$, it can be shown that

$$\hat{V} = \frac{1}{n^2}\sum_{i=0}^{n} [y(U_{(i+1)}) - y(U_{(i)})]^2 \qquad (5)$$

is a consistent estimator for $\mathrm{E}[(\overline{\overline{T}}_y - T_y)^2]$ (see Result 2 in the Appendix). Obviously, when y is null in $(1-\varepsilon,1]$, the estimator $\hat{V}$ turns out to be consistent for $\mathrm{Var}[\overline{\overline{T}}_y]$.

A SIMULATION STUDY

In order to assess the small-sample properties of estimator (4) with respect to estimator (2), a simulated experiment dealing with line intercept sampling has been considered. In this setting, it is worth noting that an interesting use of line-intercept design is described by Thompson[8]: if the study region is snowed and the total of a certain animal species (such as wolverines or arctic wolves) is the target parameter, the selected transects are flown under appropriate weather conditions with observers in the aircraft looking for animal tracks in the snow. Once a track is encountered, it is followed in each direction and mapped. Hence, the animal total is estimated on the basis of the estimated track total[17].

The previous survey setting was mimicked by simulating three populations of twenty tracks on the unit square. Thus, the target parameter was given by the population total, *i.e.* $T_y = N = 20$. The three populations were settled in such a way that the first population consisted of lines randomly located, the second population consisted of lines positioned with a slight trend, while the third population consisted of lines positioned with a marked trend. The three simulated populations are displayed in Fig. 1. These populations of lines may be considered quite representative of real situations[8].

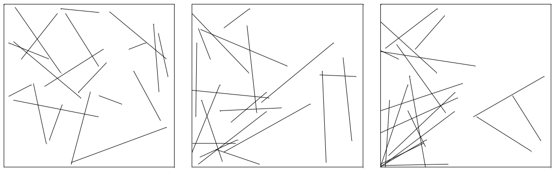

Fig. 1: The three simulated populations of lines

For the sake of simplicity, it was assumed that the baseline corresponded to the base of the unit square. Since the line-intercept design was assumed, the inclusion sets of the lines were obviously given by their projections onto the baseline. The three functions y corresponding to the simulated populations are reported in Fig. 2.

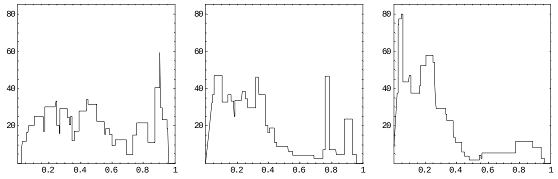

Fig. 2: The three functions y corresponding to the simulated populations of lines

The sample sizes *n* = 20,30,40 were considered in the simulation. For each sample size and for each population, $B = 2{,}000$ simulations of replicated line intercept designs were carried out. For each simulation, the realizations of estimators (2) and (4) were computed. On the basis of the *B* estimates, the simulated bias, the simulated mean square error (MSE) and the simulated relative efficiency (RE) - *i.e*. the ratio of the simulated mean square errors - were computed for the estimators (2) and (4). The corresponding results were reported in Table 1.

Table 1: Simulated performance indexes of estimators $\overline{T}_y$ and $\overline{\overline{T}}_y$

Population	*n*	Bias(MSE) of $\overline{T}_y$	Bias(MSE) of $\overline{\overline{T}}_y$	RE
Random	20	0.01 (5.24)	-0.50 (4.09)	1.28
	30	0.02 (3.43)	-0.21 (2.25)	1.52
	40	0.01 (2.57)	-0.10 (1.41)	1.81
Slight trend	20	-0.04 (17.88)	-0.26 (19.76)	0.91
	30	-0.01 (11.80)	-0.22 (10.70)	1.10
	40	-0.01 (8.75)	-0.08 (7.73)	1.13
Marked trend	20	0.02 (20.06)	-0.12 (10.40)	1.93
	30	0.00 (13.23)	-0.03 (5.25)	2.52
	40	0.01 (9.52)	-0.05 (3.15)	3.02

From Table 1, it is at once apparent that estimator (4) always outperforms estimator (2), except for the second population and $N = 20$. The performance of the Riemann estimator obviously increases as *n* increases. The best performance is achieved for the third population of lines, *i.e*. for the most irregular *y* function. Further simulations (not reported here) seem to confirm that this behavior generally occurs, even if the superiority of (4) over (2) is not marked for very small *n* values (say *n* less than 10). Finally, it should be emphasized that the bias of (4) is always negative in the simulation, a result which is consistent with the findings in the Remark of the Appendix.

Appendix: Let *y* be an elementary function, *i.e*.

$$y(u) = \sum_{k=1}^{m} a_k I_{\{b_k \le u < b_{k+1}\}} , \tag{6}$$

where $(a_1, a_2, \ldots, a_m)$ are given constants. In turn, $(b_1, b_2, \ldots, b_{m+1})$ are constants such that $b_1 \le b_2 \le \ldots \le b_{m+1}$ and $b_1 = 0$ and $b_{m+1} = 1$. It is straightforward to prove that (6) may be expressed as

$$y(u) = \sum_{k=1}^{m} (a_k - a_{k-1}) I_{\{b_k \le u \le 1\}} , \tag{7}$$

where $a_0 = 0$.

For each $x \in (0,1)$, let us assume that

$$\alpha_n(x) = \sum_{i=0}^{n} \mathrm{E}[(U_{(i+1)} - x)^2 I_{\{U_{(i)} < x \le U_{(i+1)}\}}] ,$$

while for each $x, z \in (0,1)$ such that $x < z$, let us assume that

$$\beta_n(x,z) = \sum_{i=1}^{n}\sum_{j=0}^{i-1} \mathrm{E}[(U_{(i+1)} - z) \times \\ \times (U_{(j+1)} - x) I_{\{U_{(i)} < z \le U_{(i+1)},\, U_{(j)} < x \le U_{(j+1)}\}}] \ .$$

By using expression (7) in estimator (4), it follows that

$$\overline{\overline{T}}_y - T_y = -\sum_{k=1}^{m}\sum_{i=0}^{n} (a_k - a_{k-1})(U_{(i+1)} - b_k) I_{\{U_{(i)} < b_k \le U_{(i+1)}\}}$$

and hence the mean square error of $\overline{\overline{T}}_y$ may be expressed as

$$\mathrm{E}[(\overline{\overline{T}}_y - T_y)^2] = \sum_{k=1}^{m} (a_k - a_{k-1})^2 \alpha_n(b_k) \\ + 2\sum_{1 \le h < k}^{m} (a_h - a_{h-1})(a_k - a_{k-1}) \beta_n(b_h, b_k) \,. \tag{8}$$

Therefore, in order to obtain the large-sample properties of $\mathrm{E}[(\overline{\overline{T}}_y - T_y)^2]$ - and consequently of its estimator (5) - it suffices to analyze the asymptotic properties of $\alpha_n(x)$ and $\beta_n(x,z)$ as $n \to \infty$.

Result 1: *For each* $x, z \in (0,1)$ *such that* $x < z$, *it follows that*

$$\alpha_n(x) \sim 2n^{-2},\ \beta_n(x,z) \sim n^{-2}. \tag{9}$$

M*oreover*, $\overline{\overline{T}}_y$ *has a* $O(n^{-2})$ *mean square error such that*

$$\mathrm{E}[(\overline{\overline{T}}_y - T_y)^2] \sim \frac{1}{n^2} \sum_{k=1}^{m} (a_k - a_{k-1})^2 + \frac{a_m^2}{n^2} \,. \tag{10}$$

Proof: Since the joint probability density function of $(U_{(i)}, U_{(i+1)})$ is given by

$$g(t,u) = n(n-1)\binom{n-2}{i-1} t^{i-1}(1-u)^{n-i-1} I_{\{t \le u\}} ,$$

for $i = 1, 2, \ldots, n-1$, it turns out that

$$\alpha_n(x) = c_n + \frac{2}{(n+1)(n+2)} \sum_{i=1}^{n-1} \binom{n+2}{i} x^i (1-x)^{n-i+2} ,$$

where

$$c_n = \mathrm{E}[(U_{(1)} - x)_+^2] + (1-x)^2 P(U_{(n)} \le x) \,.$$

Hence, on the basis of Newton's formula it turns out that

$$\alpha_n(x) = c_n + \frac{2}{(n+2)(n+1)} (1 - d_n) ,$$

where

$$d_n = (1-x)^{n+2} + \sum_{i=n}^{n+2} \binom{n+2}{i} x^i (1-x)^{n-i+2} \ .$$

Since $c_n + d_n = o(n^{-2})$, the first part of (9) follows. In addition, since the joint probability density function of $(U_{(1)}, U_{(2)}, \ldots, U_{(n)})$ is given by

$$g(u_1, u_2, \ldots, u_n) = n!\, I_{\{u_1 \le u_2 \le \ldots \le u_n\}} ,$$

it turns out that

$$\mathrm{E}[(U_{(i+1)}-z)(U_{(j+1)}-x)I_{\{U_{(i)}<z\le U_{(i+1)},\,U_{(j)}<x\le U_{(j+1)}\}}]=$$
$$=\frac{n!x^{j}(z-x)^{i-j+1}(1-z)^{n-i+1}}{j!(i-j+1)!(n-i+1)!}$$

for each $j \le i-1$ and $i = 2,3,\dots,n-1$. Thus, since

$$\beta_n(x,z) \sim \sum_{i=2}^{n-1} \frac{n!(1-z)^{n-i+1}}{(n-i+1)!} \sum_{i=2}^{n-1} \frac{x^{j}(z-x)^{i-j+1}}{j!(i-j+1)!}$$
$$\sim \sum_{i=2}^{n-1} \frac{n!z^{i+1}(1-z)^{n-i+1}}{(n-i+1)!(i+1)!} \sim \frac{1}{(n+1)(n+2)},$$

the second part of (9) follows. Moreover, on the basis of (8) and (9), it turns out that

$$\mathrm{E}[(\overline{\overline{T}}_y-T_y)^2]\sim$$
$$\sim\frac{2}{n^2}\sum_{k=1}^{m}(a_k-a_{k-1})^2+\frac{2}{n^2}\sum_{1\le h<k}^{m}(a_h-a_{h-1})(a_k-a_{k-1})$$
$$=\frac{1}{n^2}\sum_{k=1}^{m}(a_k-a_{k-1})^2+\frac{a_m^2}{n^2},$$

since $\sum_{k=1}^{m}(a_k-a_{k-1})=a_m$.

Remark: By similar argumentation, it follows that

$$\mathrm{Bias}[\overline{\overline{T}}_y]=\mathrm{E}[\overline{\overline{T}}_y-T_y]$$
$$=-\sum_{k=1}^{m}\sum_{i=0}^{n}(a_k-a_{k-1})\,\mathrm{E}[(U_{(i+1)}-b_k)I_{\{U_{(i)}<b_k\le U_{(i+1)}\}}]$$
$$\sim-\frac{a_m}{n}.$$

Hence, $\overline{\overline{T}}_y$ displays a negative large-sample bias. In addition, if $a_m=0$, by means of reasoning similar to Result 1, it can be easily proven that a $o(n^{-1})$ bias is achieved (with in turn a negative leading term). In this case, since $\mathrm{Bias}[\overline{\overline{T}}_y]^2=o(n^{-2})$, it follows that

$$\mathrm{Var}[\overline{\overline{T}}_y]\sim\frac{1}{n^2}\sum_{k=1}^{m}(a_k-a_{k-1})^2.$$

Result 2: *The r.v.* $\hat{V}/\mathrm{E}[(\overline{\overline{T}}_y-T_y)^2]$ *converges almost surely to* 1 *as* $n\to\infty$.

Proof: By using (10), it suffices to prove that $n^2\hat{V}$ converges almost surely to $\sum_{k=1}^{m}(a_k-a_{k-1})^2+a_m^2$. Moreover, it should be noticed that

$$n^2\hat{V}=\sum_{k=1}^{m}(a_k-a_{k-1})^2+a_m^2+2\sum_{1\le h<k}^{m}(a_h-a_{h-1})(a_k-a_{k-1})X_{hk},$$

where $X_{hk}=\sum_{i=0}^{n}I_{\{U_{(i)}<b_h\le U_{(i+1)},\,U_{(i)}<b_k\le U_{(i+1)}\}}$. Hence, the result follows since

$$\{X_{hk}\ne0\}\subset\{\max_{i=0,1,\dots,n}(U_{(i+1)}-U_{(i)})\ge u_*\},$$

where $u_*=\min_k(u_{k+1}-u_k)$. Indeed, $u_*>0$ and $\max_{i=0,1,\dots,n}(U_{(i+1)}-U_{(i)})$ converges almost surely to 0 on the basis of the Dvoretzky-Kiefer-Wolfowitz inequality[18].

REFERENCES

1. Barabesi, L. and L. Fattorini, 1998. The use of replicated plot, line and point sampling for estimating species abundancies and ecological diversity. Environ. and Ecolog. Stat., 5: 353-370.
2. Bonham, C.D., 1989. Measurements for Terrestrial Vegetation. Wiley, New York.
3. Husch, B., C.I. Miller and T.W. Beers, 1982. Forest Mensuration. Wiley, New York.
4. Schreuder, H.T., T.G. Gregoire and G.B. Wood, 1993. Sampling Methods for Multiresource Forest Inventories. Wiley, New York.
5. Barabesi, L., 2004. Replicated environmental sampling designs and Monte Carlo integration methods: two sides of the same coin, invited paper in the Proceedings of XLII Meeting of the Italian Statistical Society, June 9-11, Bari, Italy.
6. Williams, M.S. and M. Eriksson, 2002. Comparing the two paradigms for fixed area sampling in large-scale inventories. Forest Ecol. and Manage., 168: 135-148.
7. Barabesi, L. and C. Pisani, 2002. Ranked set sampling for replicated sampling designs. Biometrics, 58: 586-592.
8. Thompson, S.K., 2002. Sampling. Wiley, New York.
9. Overton, W.S., 1969. Estimating the Number of Animals in Wildlife Populations. In Wildlife Management Techniques. R.H. Giles (Ed.), The Wildlife Society, Washington DC, pp: 405-455.
10. Barabesi, L., 2003. A Monte Carlo integration approach to Horvitz-Thompson estimation in replicated environmental designs, Metron, LXI: 355-374.
11. Haber, S., 1966. A modified Monte-Carlo quadrature. Mathematics of Computation, 20: 361-368.
12. Haber, S., 1967. A modified Monte-Carlo quadrature. II, Mathematics of Computation, 21: 388-397.
13. U.S. Environmental Protection Agency, 2002. Guidance on choosing a sampling design for environmental data collection. EPA QA/G-5S, Washington DC, pp: 1-166.
14. Barabesi, L. and M. Marcheselli, 2003. A modified Monte Carlo integration. Intl. Mathl. J., 3: 555-565.
15. Barabesi, L. and M. Marcheselli, 2005. Some large-sample results on a modified Monte Carlo integration method. J. Stat. Planning and Inference, 135: 420-432.
16. Robert, C.P. and G. Casella, 2002. Monte Carlo Statistical Methods. Springer, New York.
17. Fattorini, L. and M. Marcheselli, 2002. Empirical investigation about statistical properties of abundance estimates based on line-intercept and network sampling of tracks. Stat. Methods and Applications 11: 217-226.
18. Massart, P., 1990. The tight constant in Dvoretzky-Kiefer-Wolfowitz inequality. Annals of Probability, 18: 1269-1283.