

動く調査対象集団に対する標本調査-VIII*

—RST 法とベイズ推定—

林 知己夫・駒 沢 勉

(1973年1月受付)

Estimation of the Size of Mobile Population-VIII*

—Bayesian estimation in RST-method based on
the observed number of randomly selected points—

Chikio Hayashi and Tsutomu Komazawa

This is the continuation of the paper "Estimation of the Size of Mobile Population-V — Estimation of trace length of a hare in one night by RST method and computer simulation — (Proc. Inst. Statist. Math. Vol. 19, No. 1 1971, 15-27). The authors find the Bayesian estimation of $E(x)/E(l)$ from the observed number of randomly selected points in [Model III] on the previous paper where they assume several types of prior distribution of \bar{k} (the mean value of intersection points). The results are shown in Table 1, 2. The differences are not so remarkable. The estimation based on the uniform prior distribution of \bar{k} will be adopted as practically better. This research was carried out as a part of JIBP Project (Contributions from JIBP-PT No. 163).

The Institute of Statistical Mathematics

§1 はじめに

これは動く調査対象集団に対する標本調査-V (統数研彙報第19巻第1号, 1971, 15頁-27頁) の続篇である。基本方針, そのモデルは一向に変わりはないのであるが, 調査から出た交点数 k から $E(x)/E(l)$ を推定するところの改良である。前論文25頁脚注に少しく触れてあるところであるが, 交点の平均値 \bar{k} の事前分布を考慮に入れ, ベイズ推定を行ってみることにする。モデルとしては, **MIII** を用いることにする。つまり, 従前の通り一羽の野兎の足跡長の分布は $N(\bar{l}, \sigma_l^2)$, 交点数の分布は $N(\bar{k}, \sigma_k^2)$ としておく。

§2 \bar{k} に対する事前分布

事前分布として次のものを採用する。

A (6) : これは \bar{k} が1から6までの範囲をうごくものとする。

A₁ (6) : 等確率分布

A₂ (6) : 二等辺三角形分布

A₃ (6) : 右上り直角三角形分布

A₄ (6) : 左上り直角三角形分布

の4種類を考え, この事前分布にもとづいて実現交点数 k から $E(x)/E(l)$ の推定を行う。これについて平均と分散, 及び分布を求めておく。

この行き方は, 実現値 k からまづ \bar{k} に帰因する確率 $P(\bar{k}|k)$ を求める。このため, まず \bar{k} をさだめ, 前論に示したモンテカルロ法によるシミュレーション (1000回試行) により k の分布をきめる。これを $\varphi(k|\bar{k})$ とする。次に \bar{k} の事前分布に従い \bar{k} の分布 $q(\bar{k})$ を定める。これから

* この研究は IBP の PT 部門の一環たることを意識して行われたものである。JIBP-PT No.163 である。

$$P(\bar{k}|k) = \frac{\varphi(k|\bar{k})q(\bar{k})}{\sum_k \varphi(k|\bar{k})q(\bar{k})}$$

により、 k の値に対して、それが \bar{k} にもとづく確率を求めることになる。 \bar{k} がきまれば、前論文に式したシミュレーションを用いモンテカルロ法によって $E(x) / E(l) = r(\bar{k})$ が求まるし、(1000 回試行) その分散 $\sigma(r(\bar{k}))^2$ も求まる。

$$\Sigma r(\bar{k})P(\bar{k}|k) = R(k) \text{ という平均値が求まる。}$$

一方分散は

$$\Sigma \sigma(r(\bar{k}))^2 P(\bar{k}|k) = \sigma_w(k)^2$$

という内分散が求まるので

$$\sigma^2(k) = \sigma_w(k)^2 + \sum_k (r(\bar{k}) - R(k))^2 P(\bar{k}|k) = \sigma_w(k)^2 + \sigma_b(k)^2$$

但し $\sigma_b(k)^2$ は第2項

によって求められる。

A (10): これは \bar{k} が 1 から 10 までの範囲を動くものとする。

- A₁ (10)
 - A₂ (10)
 - A₃ (10)
 - A₄ (10)
- 前述の A (6) の場合と同様。

この A (10) についても A (6) と同じ計算を実行する。

§ 3 計算結果

以上のモデルについて計算した結果観測交点 k からの推定は次の様になった (Table 1)。また、分散も Table 2 に示しておいた。

Table 1 $E(x)/E(l)$

事前分布の モデル 実測による 実現値 k	A ₁ (6)	A ₂ (6)	A ₃ (6)	A ₄ (6)	A ₁ (10)	A ₂ (10)	A ₃ (10)	A ₄ (10)
1	1.035	1.047	1.076	1.018	1.046	1.082	1.117	1.028
2	1.080	1.080	1.121	1.049	1.115	1.138	1.196	1.075
3	1.125	1.113	1.158	1.085	1.176	1.180	1.246	1.124
4	1.161	1.143	1.182	1.125	1.222	1.210	1.276	1.171
5	1.186	1.163	1.201	1.150	1.265	1.240	1.309	1.209
6	1.200	1.180	1.210	1.174	1.294	1.256	1.333	1.235
7	1.208	1.192	1.215	1.186	1.315	1.271	1.351	1.253
8	1.210	1.192	1.217	1.187	1.322	1.278	1.356	1.261
9	1.219	1.206	1.222	1.205	1.349	1.292	1.379	1.282
10	1.221	1.210	1.223	1.209	1.357	1.300	1.385	1.291
15	1.222	1.214	1.224	1.213	1.394	1.335	1.415	1.329
20	1.228	1.231	1.228	1.231	1.427	1.365	1.442	1.364

以上の表からみると、事前分布によりいくらかの差異がみとめられ、とくに A (6) と A (10) とではそのくいちがいが一応認められる。これは当然のことと思われる。いまかりに、見透しをよくするために事前分布による $E(x)/E(l)$ の算術平均 (事前分布の事前分布がこのかぎりにおいて一様の仮説) を出して比較し、前記論文のものも参考に示しておいた。Table 3 参照こうした大局的な見方をすれば 6 以下では A₁(6, A₁(10) ととても著しいものではない。

Table 2 Variance of $x/E(l)$

事前分布 k	$A_1(6)$	$A_2(6)$	$A_3(6)$	$A_4(6)$	$A_1(10)$	$A_2(10)$	$A_3(10)$	$A_4(10)$
1	0.194	0.236	0.299	0.151	0.222	0.314	0.395	0.177
2	0.310	0.314	0.400	0.238	0.389	0.437	0.566	0.297
3	0.405	0.379	0.472	0.320	0.518	0.519	0.663	0.406
4	0.473	0.434	0.515	0.401	0.607	0.574	0.718	0.496
5	0.526	0.472	0.558	0.450	0.694	0.636	0.783	0.576
6	0.552	0.502	0.573	0.493	0.752	0.669	0.831	0.626
7	0.565	0.521	0.582	0.512	0.791	0.696	0.865	0.660
8	0.574	0.526	0.591	0.517	0.808	0.713	0.875	0.680
9	0.594	0.554	0.605	0.552	0.863	0.745	0.923	0.723
10	0.598	0.561	0.607	0.559	0.878	0.760	0.934	0.741
15	0.609	0.571	0.616	0.570	0.950	0.831	0.991	0.819
20	0.634	0.631	0.634	0.631	1.016	0.892	1.045	0.891

Table 3 Arithmetic mean of $E(x)/E(l)$ on prior distributions

	A (6)	A (10)	前論文の推定
1	1.044	1.068	1.00
2	1.083	1.131	1.06
3	1.120	1.182	1.12
4	1.153	1.220	1.18
5	1.175	1.256	1.23
6	1.191	1.280	1.28
7	1.200	1.298	1.32
8	1.202	1.304	1.36
9	1.213	1.326	1.40
10	1.216	1.333	1.43
15	1.218	1.368	1.56
20	1.230	1.400	1.63

それでは、実際何を用いるべきかが問題になるが、私としては一応 $A_1(6)$ か $A_1(10)$ を採用してよいのではないかと考えている。なお、交点数が少いとき (3~4 以下位) は $A_1(6)$ が望ましく、多いとき (4~5 以上) は $A_1(10)$ を用いるのが好ましいのではないと思う。交点が 3~4 以下ではこの両者の差位は、誤差として見込んでおいて考えるべきものと思っている。交点数が多くなると $A_1(U)$ ($U > 10$) を考えなくてはなるまいが、前記 $A_1(6)$ の場合との差が大きくないところからこの場合もさして差は大ではないと考えられる。

前記 Table 1 の比較を見やすくするために第 1 図、第 2 図を示しておく。

なお参考のため $P(\bar{k}|k)$ の分布を示すと Table 4 の様になる。すべての例は面倒なので $k=$

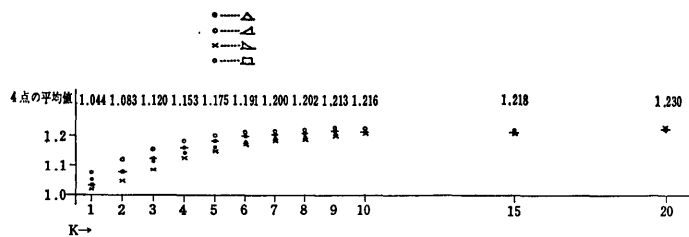


Fig. 1

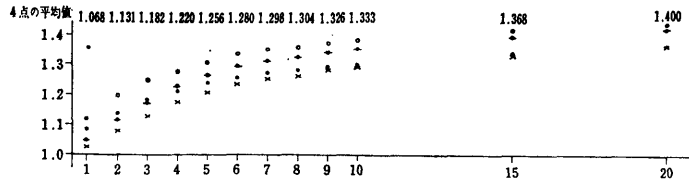


Fig. 2

4, $k=5, k=8$ について $A_1(6), A_1(10)$ の場合を示そう.

この表からみても前記した様な $A_1(6), A_1(10)$ のつかいわけが無理なものでないことが了解せられよう.

Table 4

観測 k	事前分布	\bar{k}											計
		0	1	2	3	4	5	6	7	8	9	10	
3	$A_1(6)$	0.07	0.19	0.29	0.21	0.11	0.08	0.05	—	—	—	—	1
	$A_1(10)$	0.01	0.07	0.17	0.22	0.16	0.11	0.09	0.06	0.04	0.04	0.03	1
4	$A_1(6)$	0.00	0.03	0.11	0.25	0.29	0.18	0.14	—	—	—	—	1
	$A_1(10)$	0.00	0.02	0.08	0.20	0.22	0.14	0.11	0.08	0.07	0.04	0.04	1
5	$A_1(6)$	0.00	0.00	0.08	0.17	0.25	0.25	0.25	—	—	—	—	1
	$A_1(10)$	0.00	0.00	0.05	0.11	0.17	0.17	0.17	0.11	0.10	0.07	0.05	1
8	$A_1(6)$	0.00	0.00	0.02	0.09	0.23	0.31	0.35	—	—	—	—	1
	$A_1(10)$	0.00	0.00	0.01	0.05	0.11	0.15	0.17	0.16	0.15	0.10	0.10	1

参考のため、この論文シリーズ VII における RST 法のデータをここで示した値で修正したものを示しておくとして Table 5 の様になる。全体としては前論文の便宜的推定とくらべいちじるしい差がないことが認められる。

Table 5

	データ		推定	前回の推定
	水平距離	観測交点数		
1	1573	0	1573	1573
2	458	2	424	432
3	682	0	682	682
4	475	0	475	475
5	1499	2	1388	1414
6	540	1	522	540
7	575	1	556	575
8	1122	0	1122	1122
9	458	0	458	458
10	1973	7	1500	1495
平均	—	—	870	877