

ロジスティック関数のパラメータ の推定量のモーメントについて

東京理科大学 工学部
平川 孝三 郎

1. まえがき x_i は指定変数であり, Y_i は確率変数であつて, x_i を固定したときの条件付期待値 $f(x_i)$ がロジスティック関数であるとする. すなわち

$$f(x_i) = E(Y_i | x_i) = \frac{1}{1 + e^{a - bx_i}} \quad (1)$$

とする. また x_i において大きさ n_i の無作為標本を抽出するとき, $n_i Y_i$ は 2 項分布 $b(\uparrow; n_i, f(x_i))$ に従うとする.

Y_1, Y_2, \dots, Y_m は x_1, x_2, \dots, x_m に対して上のように定義された確率変数であつて, 互に独立とする. Y の x 上への回帰関数 (1) のパラメータ a, b の最尤推定量を正確に求めることは困難であるので, 線型に変換し最小 2 乗法により得られる推定量の期待値, 分散, 歪度および尖度を評価し, 反復法によって得られる最尤推定量の期待値, 分散との関係についてふれる. a, b の推定量の分布に関する文献がほとんどないのでこれを補うのが本論の目的である.

2. a および b の最小 2 乗推定 1. において定義された $f(x)$ および Y_i に対して

$$Z = \log(1/f(x) - 1), \quad Z_i = \log(1/Y_i - 1) \quad (0 < Y_i < 1)$$

なる変換を行うと

$$Z = \log(1/E(Y|x) - 1) = a - bx, \quad Z_i = a - bx_i + \varepsilon_i$$

となる. ε_i は互に独立な確率変数であるが其期待値は一般に 0 でない.

実現値 $\{y_i\}$ のうち 0 または 1 に等しいものを除き

$$Z_i = \log(1/y_i - 1) \quad (i=1, 2, \dots, m)$$

とする。 m_i に依存する重みをつけることが考えられるが、どのような重みをつけるのが最適かは明白でない。重みを等しくするならば、最小2乗法による a および b の推定量が

$$\hat{b} = - \frac{\sum_{i=1}^m (x_i - \bar{x}) Z_i}{\sum_{i=1}^m (x_i - \bar{x})^2}, \quad \hat{a} = \bar{Z} + \hat{b} \bar{x} \quad (2)$$

によって与えられる。ただし $\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i$, $\bar{Z} = \frac{1}{m} \sum_{i=1}^m Z_i$

である。これら推定量の正確な分布を求めることは困難であるので、期待値、分散、歪度および尖度を評価する。まず Y を $b(r; n, f(x))$ に従う確率変数として $Z = \log(1/Y - 1)$ のモーメントを求めることから始める。そのために

$$Z = \log\left(\frac{1}{Y} - 1\right) = \log \frac{1 - (2Y-1)}{1 + (2Y-1)} = -2 \sum_{i=1}^{\infty} A_i (2Y-1)^{2i-1} \quad (3)$$

と表す。ただし $A_i = \frac{1}{2^i - 1}$ である。

3. Y の高次モーメント Z のモーメントを求めるためには、 $2Y-1$ の高次モーメント、したがってまた Y の高次モーメントが必要である。 nY が $b(r; n, p)$ に従うとき、 Y のモーメント母関数を $\phi(t)$ とすれば

$$\phi(t) = p^n \left(\frac{q}{p} + e^{\frac{t}{n}} \right)^n \quad (q=1-p)$$

であるから $k \leq n$ のとき

$$\phi^{(k)}(t) = \sum_{j=1}^k C(k, j) n P_j p^n \left(\frac{q}{p} + e^{\frac{t}{n}} \right)^{n-j} e^{\frac{jt}{n}} / n^k \quad (4)$$

となる。ただし $n P_j = n(n-1)\dots(n-j+1)$ であり、 $C(k, j)$ は漸化式

$$\begin{cases} C(k, j) = C(k-1, j-1) + j C(k-1, j) \\ C(k, 0) = 0, \quad C(k, k) = 1 \end{cases}$$

によって与えられる定数であり, 標本比率 Y の k 次モーメントは (5)

$$E(Y^k) = \sum_{j=1}^k C(k, j) n P_j p^j / n^k$$

となる。表1は $k=1, 2, \dots, 25; j=1, 2, \dots, k$ に対する $C(k, j)$ の値を示している。

表1 $C(k, j)$ の値 ($j=1, 2, \dots, k$)

k-1	246730	1096190550	1	96416888184100
1	179487	2734926558	524287	998969857983405
k 2	63987	3281882604	580606446	4382641999117305
1	11880	2141764053	45232115901	9741955019900400
1	1155	820784250	749206090500	12320068811796900
k 3	55	193754990	4306078895384	9593401297313460
1	1	28936908	11143554045652	4864251308951100
3	**k=12**	2757118	15170932662679	1672162773483930
1	1	165620	12011282644725	401282560341390
k 4	2047	6020	5917584964655	68629175807115
1	86526	120	1900842429486	8479404429331
7	611501	1	411016633391	762361127264
6	1379400	***k 17***	61068660380	49916988803
1	1323652	1	6302524580	2364885369
k 5	627396	65535	452329200	79781779
1	159027	21457825	22350954	1859550
15	22275	694337290	741285	28336
25	1705	5652751651	15675	253
10	66	17505749898	190	1
1	1	25708104786	1	***k 24***
k 6	**k=13**	20415995028	***k 21***	1
1	1	9528822303	1	8388607
31	4095	2758334150	1048575	47063200806
90	261625	512060978	1742343625	11681056634501
65	2532530	62022324	181509070050	485000783495250
15	7508501	4910178	3791262568401	6090236036084530
1	9321312	249900	26585679462804	31677463851804540
k 7	5715424	7820	82310957214948	82318282158320505
1	1899612	136	132511015347084	120622574326072500
63	359502	1	123272476465204	108254081784931500
301	39325	***k 18***	71187132291275	63100165695775560
350	2431	1	26826851689001	24930204590758260
140	78	131071	6833042030178	6888836057922000
1	1	64439010	1204909218331	1362091021641000
k 8	**k 14**	2798860985	149304004500	195820242247080
1	1	28958095545	13087462558	20677182465555
1	8191	110687251039	809944464	1610949936915
127	788970	197462483400	34952799	92484925445
966	10391745	189036065010	1023435	3880739170
1701	40075035	106175395755	19285	116972779
1050	63436373	37112163803	210	2454606
266	49329280	8391004908	1	33902
28	20912320	1256328866	***k 22***	276
1	5135130	125854638	1	1
k 9	752752	8408778	2097151	***k 25***
1	66066	367200	5228079450	1
255	3367	9996	727778623825	16777215
3025	91	153	19137821912055	141197991025
770	1	1	163305339345225	46771289738810
6951	**k 15**	***k 19***	602762379967440	2436684974110751
2646	1	1	1142399079991620	37026417000002430
462	16383	262143	1241963303533920	227832482998716310
36	2375101	193448101	835143799377954	690223721118368580
1	42355950	11259666950	366282500870286	1167921451092973005
k 10	210766920	147589284710	108823356051137	1203163392175387500
1	420693273	693081601779	22496861868481	802355904438462660
511	408741333	1492924634839	3295165281331	362262620784874680
9330	216627840	1709751003480	345615943200	11448507334374260
34105	67128490	1144614626805	26046574004	25958110360896000
42525	12662650	477297033785	1404142047	4299394655347200
22827	1479478	129413217791	53374629	526655161695960
5880	106470	23466951300	1389850	48063331393110
750	4550	2892439160	23485	3275678594925
45	105	243577530	231	166218969675
1	1	13916778	***k 23***	6220194750
k 11	**k 16**	527136	1	168519505
1	1	12597	1	3200450
1023	32767	171	4194303	40250
28501	7141686	1	15686335501	300
145750	171798901	***k 20***	2916342574750	1

4. Z のモーメント 級数 (3) における係数 $\{A_i\}$ から

$$B_i = \sum_{j=1}^i A_j A_{i-j+1}, \quad C_i = \sum_{j=1}^i A_j B_{i-j+1}$$

$$D_i = \sum_{j=1}^i A_j C_{i-j+1} = \sum_{j=1}^i B_j B_{i-j+1}$$

を定義すると

$$Z^2 = 4 \sum_{i=1}^{\infty} B_i (2Y-1)^{2i}, \quad Z^3 = -8 \sum_{i=1}^{\infty} C_i (2Y-1)^{2i+1}$$

$$Z^4 = 16 \sum_{i=1}^{\infty} D_i (2Y-1)^{2i+2}$$

を得る。適当に大きい K を定め、 $k=1, 2, \dots, K$ に対して

$E[(2Y-1)^k]$ を計算しておくことにより、近似値

$$E(Z^k) \approx (-2)^k \sum_{i=1}^K G_{ki} E[(2Y-1)^{2i-2+k}] \quad (6)$$

$$(G_{1i} = A_i, G_{2i} = B_i, G_{3i} = C_i, G_{4i} = D_i)$$

を得る。ここでは $K=50$ としたが、 $E(Y)=p$ が 0 または 1 に近いとき、たとえば $p=0.05$ のときは収束は非常におそい。表 2 は $n=100, 150, 200$; $p=.10(.05).50$ の場合における Z の平均、分散、歪度、尖度の近似値を示している。

表 2 $Z = \log(1/Y - 1)$ の平均、分散、歪度、尖度

n	E(Y) Z	.10	.15	.20	.25	.30	.35	.40	.45	.50
		100	平均	2.2462	1.7638	1.4059	1.1124	.8571	.6258	.4097
	分散	.1298	.0856	.0663	.0556	.0492	.0452	.0426	.0413	.0408
	歪度	.70	.47	.34	.25	.19	.14	.09	.04	.00
	尖度	4.28	3.62	3.35	3.22	3.15	3.11	3.08	3.07	3.06
150	平均	2.2287	1.7536	1.3992	1.1077	.8538	.6235	.4083	.2020	.0000
	分散	.0815	.0553	.0433	.0366	.0324	.0298	.0282	.0273	.0270
	歪度	.52	.36	.26	.20	.15	.11	.07	.03	.00
	尖度	3.69	3.35	3.21	3.14	3.09	3.07	3.05	3.04	3.04
200	平均	2.2205	1.7487	1.3959	1.1054	.8521	.6224	.4076	.2017	.0000
	分散	.0595	.0409	.0321	.0272	.0242	.0223	.0211	.0204	.0202
	歪度	.42	.30	.22	.17	.13	.09	.06	.03	.00
	尖度	3.46	3.24	3.15	3.10	3.07	3.05	3.04	3.03	3.03

5. \hat{a}, \hat{b} のモーメント 平均値および平均値の周りの
 各次モーメントをそれぞれ M_1 および M_k ($k \geq 2$) で表し, 歪度,
 尖度を λ_3, λ_4 とし, $S_x = \sum_{i=1}^m (x_i - \bar{x})^2$ とすると

$$M_k(\hat{b}) = \frac{1}{S_x^k} \sum_{i=1}^m (x_i - \bar{x})^k M_k(Z_i) \quad (k=1, 2, 3) \quad (7)$$

$$M_4(\hat{b}) = \frac{1}{S_x^4} \left\{ \sum_{i=1}^m (x_i - \bar{x})^4 M_4(Z_i) + 6 \sum_{i < j} (x_i - \bar{x})^2 (x_j - \bar{x})^2 M_2(Z_i) M_2(Z_j) \right\} \quad (8)$$

$$\lambda_3(\hat{b}) = M_3(\hat{b}) / \{M_2(\hat{b})\}^{\frac{3}{2}} \quad (9)$$

$$\lambda_4(\hat{b}) = M_4(\hat{b}) / \{M_2(\hat{b})\}^2 \quad (10)$$

である。また一つの $0 < p_0 < 1$ に対して

$$\hat{Z}_{p_0} = \hat{a} - \hat{b}x_0 \quad \text{ただし} \quad \frac{1}{1 + e^{a - bx_0}} = p_0$$

とすると \hat{Z}_{p_0} のモーメントは

$$M_k(\hat{Z}_{p_0}) = \sum_{i=1}^m C_i^k M_k(Z_i) \quad (k=1, 2, 3) \quad (11)$$

$$M_4(\hat{Z}_{p_0}) = \sum_{i=1}^m C_i^4 M_4(Z_i) + 6 \sum_{i < j} C_i^2 C_j^2 M_2(Z_i) M_2(Z_j) \quad (12)$$

$$\text{ただし} \quad C_i = \frac{1}{m} - \frac{(x_i - \bar{x})(\bar{x} - x_0)}{S_x} \quad (13)$$

であって, $x_0 = 0$ とおけば S_x のモーメントとなる。

ここで $\{P_i\}$ ($i=1, 2, \dots, m; 0 < P_i < 1$) を定め

$$f(x_i) = \frac{1}{1 + e^{a - bx_i}} = P_i$$

をみたす各 x_i において一定の大きさ n の標本を抽出して
 得られる標本比率 Y_1, Y_2, \dots, Y_m から \hat{a}, \hat{b} を求めるとき
 $\frac{\hat{b}}{b}$ のモーメント および $Z_{0.5}$ のモーメントの数値計算を行う。

$$x_i - \bar{x} = \frac{1}{b} \left\{ -\log\left(\frac{1}{P_i} - 1\right) + \frac{1}{m} \sum_{i=1}^m \log\left(\frac{1}{P_i} - 1\right) \right\}$$

であるから $\frac{\hat{b}}{b}$ のモーメントは a, b に依存しない。また

\hat{Z}_b のモーメントも a, b に依存しない。

$$f(x_j) = p_j = 0.05 \quad (j=1, 2, \dots, 19)$$

をみたす 19 個の x_j から m 個の x_j ($j=j_1, j_2, \dots, j_m$) を選んで
 新しく x_1, x_2, \dots, x_m とし, Y_i の観測値 y_1, y_2, \dots, y_m を得て
 (x_i, y_i) ($i=1, 2, \dots, m$) にロジスティック曲線のあてはめを行って
 パラメータの推定を行うときの \hat{b}/b および $\hat{Z}_{0.5}$ の平均, 分散,
 歪度および尖度をいくつかの場合について計算した結果を表 3 に
 与える。 p_1, p_{19} に対しては Z のモーメントの値を精度よく求めることが
 困難であるのでそれらを含まない場合を示しておく。

表 3 \hat{b}/b および $\hat{Z}_{0.5}$ の平均, 分散, 歪度, 尖度

n	No.	m	(j ₁ , j ₂ , ..., j _m)	\hat{b}/b				$Z_{0.5}$			
				平均	分散	歪度	尖度	平均	分散	歪度	尖度
100	1	17	(2, 3, 4, ..., 18)	1.018	.0038	.25	3.21	0	.0037	0	3.05
	2	9	(2, 4, 6, ..., 18)	1.019	.0068	.36	3.40	0	.0076	0	3.12
	3	5	(2, 6, 10, 14, 18)	1.021	.0108	.46	3.57	0	.0160	0	3.28
	4	9	(10, 11, 12, ..., 18)	1.021	.0192	.34	3.45	.0046	.0162	.06	3.07
	5	6	(8, 10, 12, ..., 18)	1.020	.0189	.44	3.65	.0014	.0128	.02	3.03
	6	7	(6, 8, 10, ..., 18)	1.018	.0126	.40	3.56	-.0006	.0083	.04	3.03
150	1	17	(2, 3, 4, ..., 18)	1.011	.0024	.18	3.11	0	.0024	0	3.03
	2	9	(2, 4, 6, ..., 18)	1.012	.0043	.26	3.21	0	.0049	0	3.07
	3	5	(2, 6, 10, 14, 18)	1.013	.0068	.34	3.31	0	.0102	0	3.15
	4	9	(10, 11, 12, ..., 18)	1.013	.0123	.25	3.24	.0029	.0106	.04	3.04
	5	6	(8, 10, 12, ..., 18)	1.013	.0121	.31	3.35	.0009	.0084	.02	3.02
	6	7	(6, 8, 10, ..., 18)	1.012	.0081	.29	3.24	-.0004	.0055	.03	3.02
200	1	17	(2, 3, 4, ..., 18)	1.008	.0018	.15	3.07	0	.0017	0	3.02
	2	9	(2, 4, 6, ..., 18)	1.009	.0032	.22	3.14	0	.0036	0	3.04
	3	5	(2, 6, 10, 14, 18)	1.010	.0050	.28	3.21	0	.0075	0	3.10
	4	9	(10, 11, 12, ..., 18)	1.010	.0090	.20	3.16	.0021	.0079	.03	3.03
	5	6	(8, 10, 12, ..., 18)	1.009	.0089	.26	3.23	.0006	.0063	.01	3.01
	6	7	(6, 8, 10, ..., 18)	1.009	.0059	.23	3.17	-.0003	.0041	.03	3.01

上の結果は x_i に対応する p_i が既知の場合について, 最小 2 乗推
 定量 \hat{a}, \hat{b} に対し \hat{b}/b および $\hat{Z}_{0.5}$ のモーメントを評価したものであって,
 実際問題としては指定変数 x_i に対する p_i が不明であるから, 精密
 な結論は得られないのであるが, 表 3 から $E(\hat{b})$ は $\{x_i\}$ の配置
 には余り関係することなく $n=100$ のとき約 2%, $n=150$ のとき 1.1~1.3%,
 $n=200$ のとき約 1% 過大になることがわかる No. 4, No. 5 のように
 $\{x_i\}$ の配置が著しく偏った場合は $\text{Var}(\hat{b}/b), \text{Var}(\hat{Z}_{0.5})$ はかなり大きくなるが
 期待値に対する大きな影響はみられない。また $\lambda_3(\hat{b})$ の最大のものが 0.46 で
 あることから, かなり正規分布に近いものと推察される。

6. a, b の近似最尤推定量のモメント

a, b の初期近似値を \hat{a}_0, \hat{b}_0 として

$$\hat{P}_{0i} = 1 / (1 + e^{\hat{a}_0 - \hat{b}_0 x_i}), \quad \hat{Q}_{0i} = 1 - \hat{P}_{0i}, \quad w_{0i} = \hat{P}_{0i} \hat{Q}_{0i}$$

$$\Delta \hat{b}_0 = \frac{(\sum n_i w_i) (\sum n_i x_i y_i - \sum n_i \hat{P}_{0i} x_i) - (\sum n_i w_i x_i) (\sum n_i y_i - \sum n_i \hat{P}_{0i})}{(\sum n_i w_i) (\sum n_i w_i x_i^2) - (\sum n_i w_i x_i)^2} \quad (14)$$

$$\Delta \hat{a}_0 = \{ \sum n_i (\hat{P}_{0i} - y_i) + (\sum n_i w_i x_i) \Delta \hat{b}_0 \} / (\sum n_i w_i) \quad (15)$$

として, $\hat{b}_0 + \Delta \hat{b}_0, \hat{a}_0 + \Delta \hat{a}_0$ によって得られた値を再び \hat{b}_0, \hat{a}_0 において数回反復を行うことによって最尤推定値の近似値 \hat{b}^*, \hat{a}^* を求めることができる [3].

ここでは \hat{b}_0, \hat{a}_0 として前述の最小2乗推定値 \hat{b}, \hat{a} を用い反復によって得られた最終値 \hat{b}^*, \hat{a}^* との差を $\Delta \hat{b}, \Delta \hat{a}$ とし, モンテカルロ法によって推定値 $\hat{Var}(\Delta \hat{b}), \hat{Cov}(\hat{b}, \Delta \hat{b}), \hat{Var}(\Delta \hat{a}), \hat{Cov}(\hat{a}, \Delta \hat{a})$ を求め

$$\hat{Var}(\hat{b}^*) = \hat{Var}(\hat{b}) + \hat{Var}(\Delta \hat{b}) + 2\hat{Cov}(\hat{b}, \Delta \hat{b}) \quad (16)$$

$$\hat{Var}(\hat{a}^*) = \hat{Var}(\hat{a}) + \hat{Var}(\Delta \hat{a}) + 2\hat{Cov}(\hat{a}, \Delta \hat{a}) \quad (17)$$

によって $\hat{Var}(\hat{b}^*), \hat{Var}(\hat{a}^*)$ を求めることを提案する. $\hat{Var}(\hat{b}),$

$\hat{Var}(\hat{a})$ は §5 までにおいて求めた値を用いる. $b=1, a=0$ の場合についてモンテカルロ法を用い (16), (17) を求めれば, 一般の場合についての $\hat{Var}(\hat{b}^*/b), \hat{Var}(\hat{a}^*/a)$ が得られる. $n_1 = n_2 = \dots = n_m = 100$ として推定値を 100 回求めた場合の計算例および $(\hat{b}/b, \hat{b}^*/b), (\hat{a}/a, \hat{a}^*/a)$ の同時分布の例を前頁に示す. ○ は最小2乗推定値より, 最尤推定値の方が真の値に近くなったものを示し, □ は逆に真の値から遠くなったものを示している. (上記 1~6 頁の大部分は下記 [2] からの抜粋である)

[1] Hirakawa, K. The comparison of power of distribution-free two-sample tests. TRU Mathematics, Vol. 10 (1974).

[2] 平川孝三郎 ロジスティック関数のパラメータの推定. 東京理科大学紀要(教養篇) 第7号 (1975).

[3] Berkson, J. Tables for the maximum likelihood estimate of the logistic function. Biometrics. - 8 -