㊽ 訂正と補遺

菅原　正己

講究録第5巻第10号にのせた「観測値の一部が使えない場合の推定について」の数値表の一部に計算の誤りがあることを発見したので，こゝに訂正したい。

上述の論文の目的は正規分布よりの標本中，ある数値（以下それを0とする）より小さい値をとる標本を用いないで，母集団パラメーターを求めることであった。

計算によれば，*Maximum liklihood estimate* も積率による *estimate* も，次の同一の結果を与える。

即ち，

$$F(x) = \frac{e^{-\frac{x^2}{2}}}{\int_{-x}^{\infty} e^{-\frac{t^2}{2}}\,dt}$$

とおけば，

$$\frac{1+\left(\frac{\mu}{\sigma}\right)^2 + F\left(\frac{\mu}{\sigma}\right)\frac{\mu}{\sigma}}{\left(\frac{\mu}{\sigma} + F\left(\frac{\mu}{\sigma}\right)\right)^2} = \frac{\frac{\sum x_i^2}{n}}{\left(\frac{\sum x_i}{n}\right)^2} \quad \cdots\cdots \quad (1)$$

$$\mu = \frac{1}{1+F\left(\frac{\mu}{\sigma}\right)\frac{\sigma}{\mu}} \cdot \frac{\sum x_i}{n} \quad \cdots\cdots \quad (2)$$

を解くことにより，σ 及び μ の推定値が得られる。

（２）の代りに

$$\sigma = \frac{1}{\frac{\mu}{\sigma} + F\left(\frac{\mu}{\sigma}\right)} \cdot \frac{\Sigma x_i}{n} \quad \cdots\cdots\cdots (2)'$$

を用いてもよい。

（１）より $\frac{\mu}{\sigma}$ を求め，これを（２）（２'）に代入して σ, μ を求めるためには，

$$\varphi(x) = \frac{1 + x^2 + F(x) \cdot x}{(x + F(x))^2}$$

の表や，$1 + x^{-1}F(x)$，$x + F(x)$ 等の表が必要である。

これらの表に少し誤があったので，以下，その表をかかげる。

第　一　表

x	$\varphi(x)$	x	$\varphi(x)$	x	$\varphi(x)$	x	$\varphi(x)$
-1.8	1.804	-0.4	1.6378	0.5	1.4774	2.0	1.2099
-1.6	1.787	-0.3	1.6219	0.6	1.4579	2.2	1.1839
-1.4	1.767	-0.2	1.6053	0.7	1.4384	2.4	1.1611
-1.2	1.746	-0.1	1.5883	0.8	1.4188	2.6	1.1412
-1.0	1.722			0.9	1.3992	2.8	1.1240
-0.9	1.709	0.0	1.5708	1.0	1.3798	3.0	1.1093
-0.8	1.696	0.1	1.5528	1.2	1.3417	4.0	1.0625
-0.7	1.682	0.2	1.5344	1.4	1.3051	5.0	1.0400
-0.6	1.668	0.3	1.5158	1.6	1.2707		
-0.5	1.653	0.4	1.4967	1.8	1.2388		

第 二 表

$\varphi(k)$	k	$k+F(k)$	$\varphi(k)$	k	$1+F(k)k^{-1}$	$k+F(k)$	$\varphi(k)$	k	$1+F(k)k^{-1}$
1.70	-0.831	0.561	1.50	0.383	2.493	0.954	1.30	1.430	1.110
1.69	-0.757	0.577	1.49	0.435	2.271	0.978	1.29	1.488	1.098
1.68	-0.686	0.594	1.48	0.487	2.070	1.003	1.28	1.546	1.085
1.67	-0.614	0.612	1.47	0.538	1.922	1.028	1.27	1.604	1.073
1.66	-0.547	0.629	1.46	0.589	1.793	1.053	1.26	1.667	1.064
1.65	-0.480	0.647	1.45	0.641	1.694	1.080	1.25	1.730	1.055
1.64	-0.415	0.665	1.44	0.692	1.603	1.107	1.24	1.793	1.047
1.63	-0.351	0.683	1.43	0.743	1.533	1.135	1.23	1.861	1.040
1.62	-0.289	0.702	1.42	0.794	1.467	1.164	1.22	1.930	1.034
1.61	-0.228	0.721	1.41	0.845	1.416	1.193	1.21	1.999	1.028
1.60	-0.169	0.740	1.40	0.896	1.366		1.20	2.076	1.023
1.59	-0.110	0.759	1.39	0.947	1.327		1.19	2.153	1.019
1.58	-0.053	0.779	1.38	0.999	1.288		1.18	2.234	1.015
1.57	0.004	0.800	1.37	1.051	1.261		1.17	2.322	1.012
1.56	0.060	0.820	1.36	1.104	1.233		1.16	2.411	1.009
1.55	0.115	0.841	1.35	1.156	1.206		1.15	2.512	1.007
1.54	0.170	0.863	1.34	1.209	1.180		1.14	2.614	1.005
1.53	0.224	0.885	1.33	1.264	1.162		1.13	2.730	1.004
1.52	0.277	0.908	1.32	1.319	1.143		1.12	2.854	1.002
1.51	0.330	0.931	1.31	1.373	1.125		1.11	2.991	1.002

上述の推定法を用いて実際に計算した例を次にかゝげる。

以下は，Deming の "Statistical adjustment of data" の附錄の表よりとつた例である。

これは E.L. Dodd が Tipett の乱数表を変換して作つた表で，平均値 0，標準偏差 1 なる正規母集団よりの random sample を 800 個のせてある。

これより sample size が夫々 80，160，320，800 の標本をとり，この標本全部を用いた推定平均値 $\bar{x}=\frac{\sum x_i}{n}$，推定標準偏差* $s=\sqrt{\frac{\sum(x_i-\bar{x})^2}{n}}$ の計算値，および標本の一部，例えば標本値 −1 以下のものを切り捨てたものから平均値，標準偏差を推定した値を，次の表にかかげる。

註．$s=\sqrt{\frac{\sum(x_i-\bar{x})^2}{n}}$ を用いたのは，標本の一部が切り捨てられているときの推定法と相応させるためである。標本の大きさが 80，160，…… のときには $\left(\frac{\sum(x_i-\bar{x})^2}{n}\right)^{\frac{1}{2}}$ も $\left(\frac{\sum(x_i-\bar{x})^2}{n-1}\right)^{\frac{1}{2}}$ も実質的にはあまり差異がない。

第　三　表

n は標本数，m，s は平均値，標準偏差の推定値を示す

標本番号	標本全部を用いたとき			−1以下を切り捨てたとき			0以下を切り捨てたとき			0以上を切り捨てたとき		
	n	m	s	n	m	s	n	m	s	n	m	s
No.1	80	0.038	0.853	72	-0.18	1.01						
No.2	80	-0.088	1.015	66	-0.14	1.04						
No.3	80	-0.098	1.014	65	-0.54	1.29						
No.4	80	-0.017	1.067	62	0.09	1.06						
No.5	80	-0.096	1.032	64	-0.15	1.09						
No.6	160	0.073	0.816	146	-0.05	0.92	87	-0.74	1.12	73	-0.51	0.48
No.7	160	-0.110	1.054	134	-0.34	1.14	70	0.59	0.72	90	0.38	1.235
No.8	160	-0.083	1.012	131	-0.22	1.11	67	0.39	0.88	93	0.15	1.01
No.9	160	-0.014	1.060	127	0.02	1.10	76	0.24	1.01	84	-0.47	0.79
No.10	160	0.032	1.007	137	0.01	1.02	77	0.25	0.96	83	0.74	1.21
No.11	320	-0.019	0.947	280	-0.17	1.02	157	0.10	0.90	163	0.00	0.97
No.12	320	-0.048	1.037	258	-0.11	1.10	143	0.32	0.96	177	-0.23	0.89
No.13	800	-0.020	0.996	674	-0.11	1.05	376	0.20	0.95	424	0.00	0.98

標本番号	−2以下を切り捨てたとき			1以上を切り捨てたとき		
	n	m	s	n	m	s
No.11	313	0.01	0.90	269	-0.12	0.88
No.12	315	-0.12	1.10	270	-0.15	0.95
No.13				675	-0.10	0.93

以上の計算の結果を見てわかることは，特に平均値に於いて，推定値が真の価とかなり離れていることである。標準偏差の方がかえつて安定していることは，通常平均値の方が，標準偏差よりも信頼できることが多いのと異つている。

殊に平均値以下，または以上を切り捨てたときは，相当の大標本でないと平均値の推定値を信頼することができないのである。

しかしこれは推定法が悪いからではないと考えられる。

正規分布 $\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ の分布曲線を描いてみればわかるように，$x=\mu$ に於ける確率密度の値と，$x=\mu\pm 0.5\sigma$ に於ける確率密度の値とは 10％ 程しか異らないので，相当多数の標本をとつてヒストグラムを描いても，山の頂はかなり左右に移動することが多い。

ヒストグラムの山の頂が左右に偏つても，平均値や標準偏差に大きな影響を及ぼすのは左右の裾の方に現れる標本値であつて，中央附近にある標本値は効いて来ない。

ところが現在我々が取り扱う問題では一方の裾を切り捨てるため，左右に動き易い山の頂附近の値が効いて，平均値の推定値が大きくゆれるのである。事実ヒストグラムの一方の裾を切り捨てたものに，我々の推定法できめた m，s を用いて正規曲線を描くと，裾を切り捨てられたヒストグラムと曲線とはかなりよく一致する。

即ち，推定法が悪いのではなく，ある値以下の標本を切り捨てたときには，どうしても多数の標本を用いなければ，μ や σ を精密には推定できないのである。

以上の計算に於いて，吉田禮子，大澤政子二氏に負う所が甚だ多い。こゝに，感謝の意を表する。