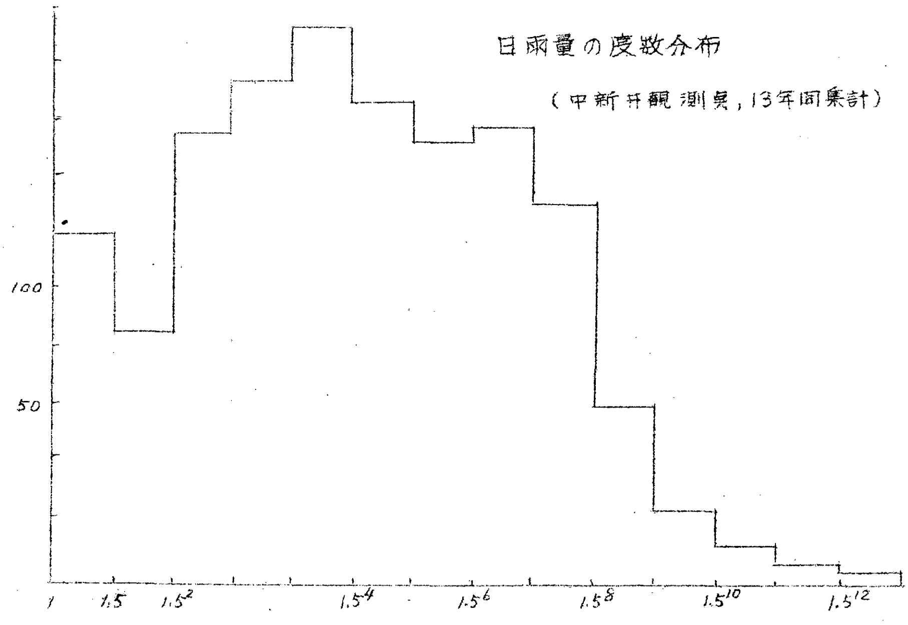
㉟ 観測値の一部が使えない場合の推定について

菅原　克己

雨量の統計をとつてみると，日雨量についても，一雨雨量（一つの低気圧，または一つの不連続線によつて降る一続きの雨の雨量）についても，1.5 mm程度以下の小雨が甚だ多い。

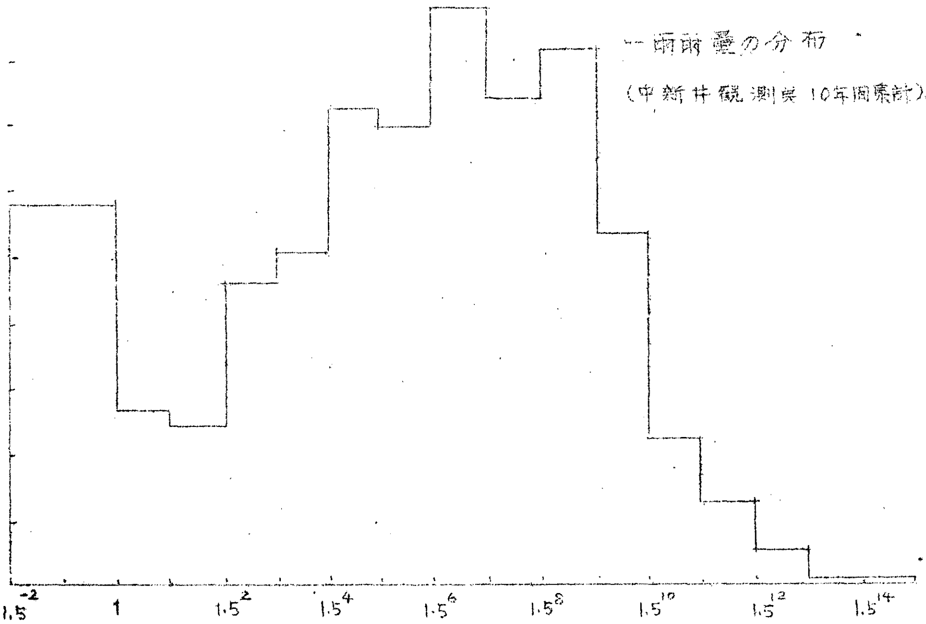
しかしそういう小雨を除いてみると，残りはかなりジブラ分布に近い。*



第一図（a）

---

* 近く発表の予定



第 一 図 (b)

そこで，1.5 mm 以上の雨量について，度数分布図をかいてみると，山型分布の一方の裾を切り去つた形となる。

問題は，この母集団分布が正規であるとして，その母集団平均値，母集団標準偏差を求めることである。

この問題は，ある正規母集団に他の母集団が混入し，ある値以下（または以上）の標本が使えない場合（雨量統計の例），または観測寸法，測定器などの制限により，ある値以上（または以下）の標本が信用出來ない場合などに応用される途が広いと思う。

正規母集団の標本のうち，ある値以下のものを使用しないということは，ある値以下の標本は現れないことゝ考えてもよい。

その値を 0 にとれば，母集団分布をつぎのものと考えることである。

$$p(x)\begin{cases} = \dfrac{\dfrac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\dfrac{1}{\sqrt{2\pi}\,\sigma}\displaystyle\int_0^\infty e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx,} & (x \geqq 0) \\ = 0 & (x < 0) \end{cases}$$

この母集団から n 個の標本 x_1, x_2, ----, x_n ($x_i \geqq 0$) がとられたとき, μ, σ を推定することが問題である。

そのために Maximum liklihood estimate を計算する。

liklihood function は

$$L(\sigma,\mu) = \frac{\left(\dfrac{1}{\sqrt{2\pi}\,\sigma}\right)^n e^{-\frac{\sum(x_i-\mu)^2}{2\sigma^2}}}{\left(\dfrac{1}{\sqrt{2\pi}\,\sigma}\displaystyle\int_0^\infty e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx\right)^n}$$

$$= \frac{\left(\dfrac{1}{\sqrt{2\pi}\,\sigma}\right)^n e^{-\frac{\sum(x_i-\mu)^2}{2\sigma^2}}}{\left(\dfrac{1}{\sqrt{2\pi}}\right)^n\left(\displaystyle\int_{-\frac{\mu}{\sigma}}^\infty e^{-\frac{t^2}{2}}dt\right)^n}$$

である。

この対数をとり σ, μ につき微分すれば,

$$\frac{\partial}{\partial\sigma}\log L(\sigma,\mu)=-\frac{n}{\sigma}+\frac{\Sigma(x_i-\mu)^2}{\sigma^3}+\frac{e^{-\frac{\mu^2}{2\sigma^2}}\frac{\mu}{\sigma^2}}{\int_{-\frac{\mu}{\sigma}}^{\infty}e^{-\frac{t^2}{2}}dt}$$

$$\frac{\partial}{\partial\mu}\log L(\sigma,\mu)=\frac{\Sigma(x_i-\mu)}{\sigma^2}-n\frac{e^{-\frac{\mu^2}{2\sigma^2}}\frac{1}{\sigma}}{\int_{-\frac{\mu}{\sigma}}^{\infty}e^{-\frac{t^2}{2}}dt}$$

上の二式の右辺を 0 と置いて，この連立方程式から σ と μ を求めればよい。　上の二式の右辺を 0 と置いたものから，

$$\sigma^2-\frac{e^{-\frac{\mu^2}{2\sigma^2}}}{\int_{-\frac{\mu}{\sigma}}^{\infty}e^{-\frac{t^2}{2}}dt}\sigma\mu=\frac{\Sigma(x_i-\mu)^2}{n}, \quad \cdots\cdots (1)$$

$$\mu+\frac{e^{-\frac{\mu^2}{2\sigma^2}}}{\int_{-\frac{\mu}{\sigma}}^{\infty}e^{-\frac{t^2}{2}}dt}\sigma=\frac{\Sigma x_i}{n}. \quad \cdots\cdots (2)$$

(1) の右辺 $\frac{\Sigma(x_i-\mu)^2}{n}=\frac{\Sigma x_i^2}{n}-2\mu\frac{\Sigma x_i}{n}+\mu^2$

の $\frac{\Sigma x_i}{n}$ に（2）を代入して整理すれば

$$\sigma^2+\mu^2+\frac{e^{-\frac{\mu^2}{2\sigma^2}}}{\int_{-\frac{\mu}{\sigma}}^{\infty} e^{-\frac{t^2}{2}}dt}\,\sigma\mu=\frac{\sum x_i^2}{n} \cdots\cdots \quad (3)$$

（３）を（２）の平方で割り，

$$\frac{1+\left(\frac{\mu}{\sigma}\right)^2+F\left(\frac{\mu}{\sigma}\right)\frac{\mu}{\sigma}}{\left(\frac{\mu}{\sigma}+F\left(\frac{\mu}{\sigma}\right)\right)^2}=\frac{\frac{\sum x_i^2}{n}}{\left(\frac{\sum x_i}{n}\right)^2} \cdots\cdots \quad (4)$$

但し

$$F\left(\frac{\mu}{\sigma}\right)=\frac{e^{-\frac{\mu^2}{2\sigma^2}}}{\int_{-\frac{\mu}{\sigma}}^{\infty} e^{-\frac{t^2}{2}}dt}$$

（４）の左辺は $\frac{\mu}{\sigma}$ の函数であり，右辺は標本値の函数であるから，この式から $\frac{\mu}{\sigma}$ が定まり，これを（２）に代入して，μ, σ を求めることができる。

幸にして，（４）の左辺は $\frac{\mu}{\sigma}$ について，単調減少函数である。つぎにこの函数の数値表を掲げる（第一表）

実際に計算するときには，第一表を使うのは不便である。実用には第一表の逆函数の表を作つて置いた方がよい。また（２）を変形すれば

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

となるから，$1+F(\xi)\xi^{-1}$ の表も作つて置く。

第 一 表

k	$\frac{1+k^2+kF(k)}{(k+F(k))^2}$	k	$\frac{1+k^2+kF(k)}{(k+F(k))^2}$	k	$\frac{1+k^2+kF(k)}{(k+F(k))^2}$
-1.8	1,804	0.0	1,5708	2.0	1,2099
-1.6	1,787	0.1	1,5528	2.2	1,1839
-1.4	1,767	0.2	1,5344	2.4	1,1611
-1.2	1,757	0.3	1,5158	2.6	1,1412
-1.0	1,722	0.4	1,4967	2.8	1,1240
-0.9	1,709	0.5	1,4774	3.0	1,1093
-0.8	1,696	0.6	1,4579	4.0	1,0625
-0.7	1,682	0.7	1,4354	5.0	1,0400
-0.6	1,668	0.8	1,4188		
-0.5	1,666	0.9	1,3992		
-0.4	1,6378	1.0	1,3798		
-0.3	1,6219	1.2	1,3417		
-0.2	1,6053	1.4	1,3051		
-0.1	1,5883	1.6	1,2707		
		1.8	1,2358		

$\frac{\mu}{\sigma}$ が 0 に近いときは (2′) は不便であるから

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

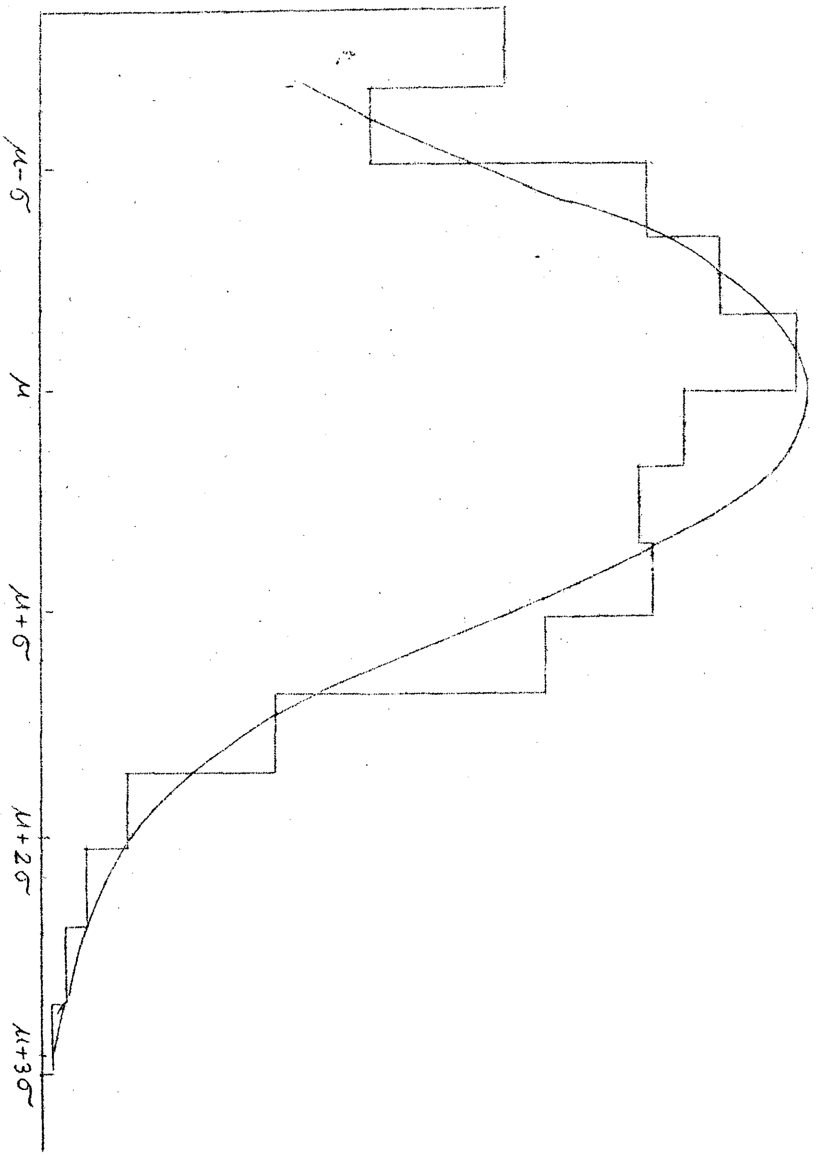
により，σ を求めた方がよい，そこで k の小さな値に対しては，$k+F(k)$ の値を掲げる．

第　二　表

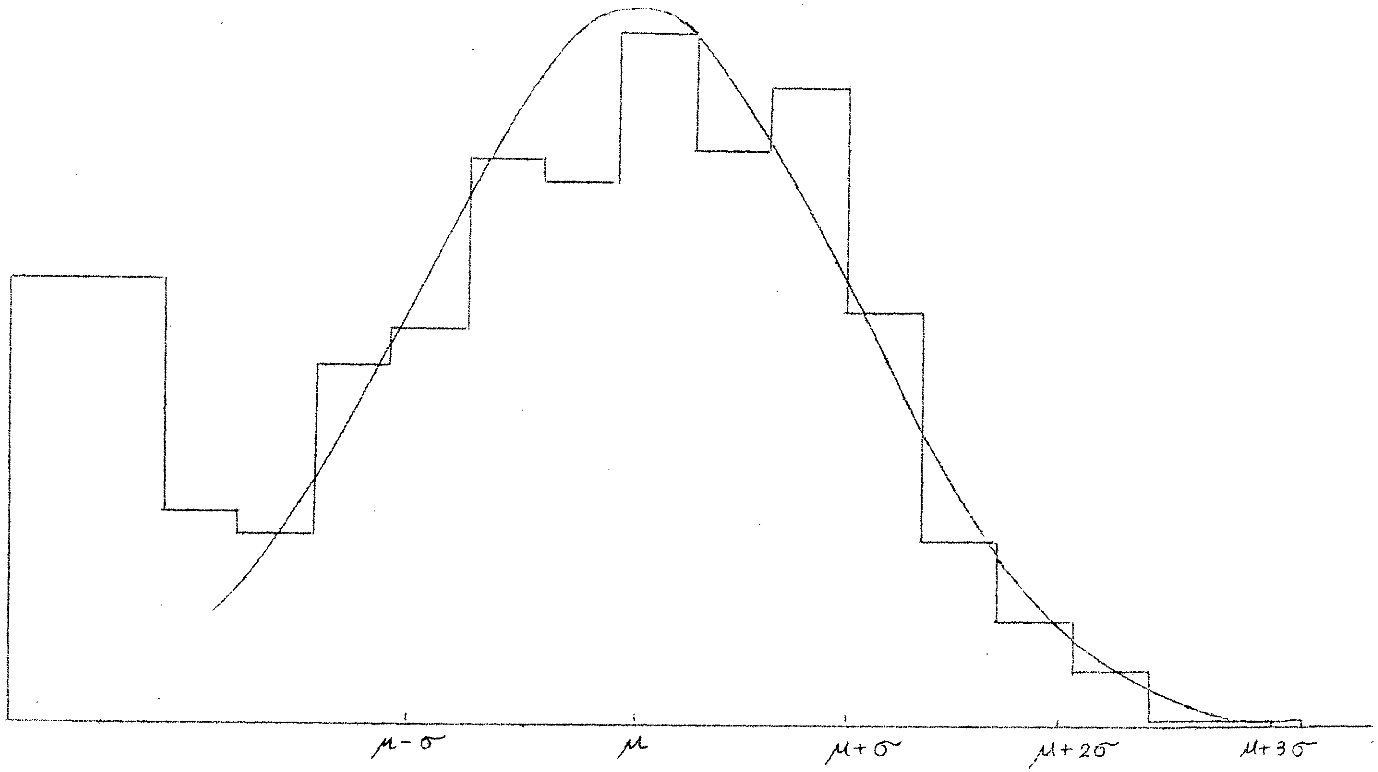
$$\varphi(k)=\frac{1+k^2+kF(k)}{(k+F(k))^2}$$ より k を求める表

$\varphi(k)$	k	$1+F(k)k^{-1}$	$\varphi(k)$	k	$1+F(k)k^{-1}$	$\varphi(k)$	k	$k+F(k)$
1.10	3.012	1.001	1.30	1.429	1.108	1.50	0.383	0.954
1.11	2.957	1.002	1.31	1.373	1.124	1.51	0.330	0.930
1.12	2.853	1.002	1.32	1.319	1.140	1.52	0.277	0.907
1.13	2.746	1.003	1.33	1.264	1.157	1.53	0.224	0.885
1.14	2.613	1.005	1.34	1.209	1.178	1.54	0.169	0.862
1.15	2.511	1.007	1.35	1.157	1.201	1.55	0.115	0.841
1.16	2.411	1.009	1.36	1.104	1.227	1.56	0.060	0.820
1.17	2.321	1.012	1.37	1.052	1.255	1.57	0.004	0.799
1.18	2.234	1.015	1.38	0.999	1.288			
1.19	2.153	1.019	1.39	0.947	1.324			
1.20	2.076	1.023	1.40	0.896	1.367			
1.21	1.999	1.028	1.41	0.845	1.412			
1.22	1.930	1.033	1.42	0.794	1.466			
1.23	1.861	1.039	1.43	0.743	1.528			
1.24	1.792	1.046	1.44	0.692	1.600			
1.25	1.729	1.054	1.45	0.640	1.688			
1.26	1.667	1.062	1.46	0.589	1.790			
1.27	1.605	1.072	1.47	0.538	1.909			
1.28	1.546	1.083	1.48	0.488	2.055			
1.29	1.487	1.095	1.49	0.435	2.248			

第二表を用いて，第一図の分布に対する，平均値と標準偏差を求め，それに対応する正規分布の曲線を画くと第二図になる。



第二図（a）



第二図（b）

パラメーターを推定するもう一つの方法は，標本値能率を母集団能率に等しいと置いてパラメーターを求める方法である。
今の場合は，この方法によつても，maximum likelihood によるのと同じ推定値が得られる。

（1950，3，22，本研究は文部省科学研究費による）