㉘ 効率の一問題

寿 賀　保 志

買物をする場合には誰でも安くて質の良い品物を手に入れたいと思うが，余り多くの時間と足代を使っては引合はぬ事になるであろう。　そこで如何にすれば労少くして好結果が得られるか，と言う問題が考えられる。　市場にある所定の商品の全体を母集団と考えるが，便宜上価格 C_0 は全商品につき一定とし，その質を変量と考える。（これは大きな制限ではない。）　そして次の様な数学的模型を設ける。；

$F(x)$ なる分布を持つ母集団より n ヶの独立なランダム・サムプル

$$x_1,\ x_2,\ \cdots\cdots\cdots\cdots,\ x_n$$

を抽出し，これに要するコストを $C(n)$ とする。

$$u = \max\{x_1, x_2, \cdots\cdots, x_n\}$$

とし，母集団平均を m とすれば，サムプル中最良の物を手に入れた時の利得は，$(u-m)$ の函数 $g(u-m)$ で表はされるであろう。　従って単位コスト当りの利得の平均は，u の密度函数を，$h(u)$ とする時，

$$e(n) = \int_{-\infty}^{\infty} \frac{g(u-m)}{c(n)} \cdot h(u)\,du \qquad \cdots\cdots\cdots\cdots (1)$$

となり，これは単位コスト当りの利得即ち効率を表はすと考へら

れるから，この $e(n)$ を最大ならしめる n の値が，最有効な試行回数と言へるであろう。

一般に，

$$h(u)du = n\{F(u)\}^{n-1}f(u)du.$$

こゝに $f(u)=F'(u)$.

コストは一般に $C(n)=C_0+C_1n$ （C_1 は一試行当りのコスト）と表はせるから，(1) は次の様に書直せる：

$$e(n)=\frac{n}{C_0+C_1(n)}\int_{-\infty}^{\infty}g(u-m)\{F(u)\}^{n-1}f(u)du. \qquad (2)$$

さて，

$$g(u-m)=g(0)+g'(0)(u-m)+\frac{g''(0)}{2!}(u-m)^2+\cdots\cdots \qquad (3)$$

と展開して，第一近似として第一項迄取り，常識的に $g(0)=0$, $g'(0)=k>0$ と考えると，(2) は次の様になる．

$$e(n)=\frac{n}{C_0+C_1n}\int_{-\infty}^{\infty}k(u-m)\{F(u)\}^{n-1}f(u)du$$

$$=\frac{kn}{C_0+C_1n}\left[\int_{-\infty}^{\infty}u\{F(u)\}^{n-1}f(u)du-\frac{m}{n}\right] \qquad (4)$$

又は，

$$=\frac{k}{C_0+C_1n}(\bar{u}-m) \quad \text{-------------} \qquad (5)$$

但し $\bar{u}=E(u)=\int_{-\infty}^{\infty}u\cdot n\{F(u)\}^{n-1}f(u)du.$

［Ex. 1］　一様分布の場合

$$f(x)=\frac{1}{a} \qquad (0\leqq x\leqq a)$$

とすれば， $F(u) = \frac{u}{a}$ だから (4) より

$$\bar{u} = \int_0^a u \cdot n \left(\frac{u}{a}\right)^{n-1} \cdot \frac{1}{a}\, du$$

$$= n \int_0^a \left(\frac{u}{a}\right)^n du$$

$$= \frac{n}{n+1}\, a$$

従つて (5) に入れると，

$$e(n) = \frac{k}{C_0 + C_1 n} \left(\frac{n}{n+1} a - \frac{a}{2} \right)$$

$$= \frac{a k}{2(C_0 + C_1 n)} \cdot \frac{n-1}{n+1}$$

n について微分すると

$$e'(n) = \frac{-2k}{2(C_0 + C_1 n)^2} \left\{ C_1 n^2 - 2 C_1 n - (2C_0 + C_1) \right\}$$

$e'(n) = 0$ とおいて $e(n)$ を最大ならしめる n の値 n_0 を求めると，

$$n_0 = 1 + \sqrt{2\left(1 + \frac{C_0}{C_1}\right)} \quad \text{-----------} \quad (6)$$

即ち n_0 の値は C_0/C_1 なる比率によつて決まる。 これは，商品の価格とコストとが同じ単位で表現されてゐなければならぬことを示す。（つまり商品を探すに要する労力をも金額に換算する必要がある！）。又 (6) は a 及び $k = g'(0)$ の値に無関係であることも大切であろう。表にして見ると

C_0/C_1	1	2	5	10	20	30	40	50	60
n_0	3	3.4	4.5	5.7	7.5	8.9	10.1	11.1	12.0

となり，多くの場合 10 以上サムプルを取るのは無用と言へる。

［$\mathcal{E}x.2$］　指数分布の場合

$$f(x) = ae^{-ax} \quad \left(\begin{array}{l} a > 0 \\ 0 \leqq x < \infty \end{array} \right)$$

この時は (4) 式を使うと簡単に結果が得られぬから，n が比較的大なる時近似的に

$$\int_0^{\bar{u}} f(x)dx = \frac{n}{n+1} \quad \cdots\cdots\cdots\cdots \quad (7)$$

と考へると，

$$1 - e^{-a\bar{u}} = \frac{n}{n+1}$$

$$\bar{u} = \frac{1}{a} \log (n+1) \quad \cdots\cdots\cdots\cdots \quad (8)$$

これを (5) に入れると

$$e(n) = \frac{k}{c_0 + c_1 n} \left(\frac{1}{a} \log (n+1) - \frac{1}{a} \right) \cdots\cdots\cdots (9)$$

$$e'(n) = \frac{k}{a(c_0 + c_1 n)} \left\{ \frac{c_0 + c_1 n}{n+1} - c_1 (\log (n+1) - 1) \right\}$$

$e'(n) = 0$　とおけば

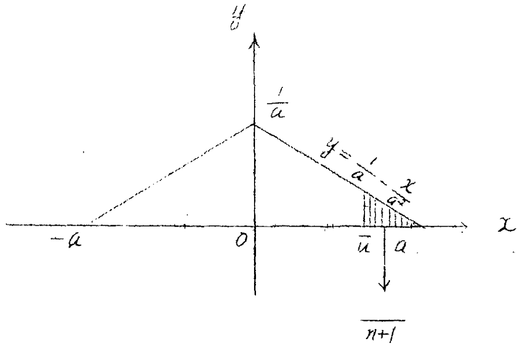
$$(n+1) \log (n+1) = 2n + \left(\frac{c_0}{c_1} + 1 \right) \cdots\cdots\cdots (10)$$

この根は，次の2つのグラフの交点として求められる。［附図1参照］

$$\left\{ \begin{array}{ll} y = 2x + \left(\dfrac{c_0}{c_1} + 1 \right) \cdots\cdots\cdots\cdots & (11) \\ y = (x+1) \log (x+1) \cdots\cdots\cdots\cdots & (12) \end{array} \right.$$

これは随分能率の悪い場合であるが，それでもサイズは高々20前後で十分と言へる。

〔Ex. 3〕　三角分布の場合



$$
f(x) = \begin{cases} \dfrac{1}{a} + \dfrac{x}{a^2} & (-a \leqq x < 0) \\ \dfrac{1}{a} - \dfrac{x}{a^2} & (0 \leqq x \leqq a) \end{cases}
$$

$$
\int_{\bar{u}}^{a} \left(\frac{1}{a} - \frac{x}{a^2} \right) dx = \frac{1}{n+1}
$$

$$
\frac{1}{2} - \left(\frac{\bar{u}}{a} - \frac{\bar{u}^2}{2a^2} \right) = \frac{1}{n+1}
$$

これをとくと，

$$
\bar{u} = a \left(1 - \sqrt{\frac{2}{n+1}} \right)
$$

又明かに $m = 0$ だから

$$
e(n) = \frac{ak}{C_0 + C_1 n} \left(1 - \sqrt{\frac{2}{n+1}} \right)
$$

$e'(n) = 0$　とおくと

$$
3n + \left(\frac{C_0}{C_1} + 2 \right) = \sqrt{2}\,(n+1)^{\frac{3}{2}} \qquad \cdots\cdots \qquad (13)
$$

となる。　これをとくには次の２つのグラフを描きその交点を求めれば良い。（附図２参照）

$$
\begin{cases} y = 3x + \left(\dfrac{C_0}{C_1} + 2 \right) & \text{-----------} \qquad (14) \\ y = (x+1)\sqrt{2(x+1)} & \text{-----------} \qquad (15) \end{cases}
$$

これは最もありふれた場合であろうが，n_0 は高々10前後で良い様である。 以上を通観すれば，

多くの場合最有効試行回数は，そんなに大きくはない。（余程高価な或は多量の買附は別にして，大体サムプルを10前後，高々15個を取れば十分でそれ以上は無駄骨折）と言う結論が出る。（正規分布の場合は適当な三角分布で代用出来る）。

残された問題は，

(i) 商品価格を一定としないで変量と考へるべきこと。

(ii) 利得の表現として $g(u-m)$ が適当であるか否か。

（これは Wald の考へ方を取り入れたものであるが，それ以後の立場は違つて来てゐるから，その比較検討も必要となる。）

(iii) 指数分布，三角分布の場合に近似式を使って解いたが，実際の正しい値との違ひ如何．

然し焦点を我々の行動の大体の目安を定める所に置けば，iii) は大して問題にならぬであろうし，i) も亦大体満足されたと考へて良い。 従って ii) が宿題として残された事になる。

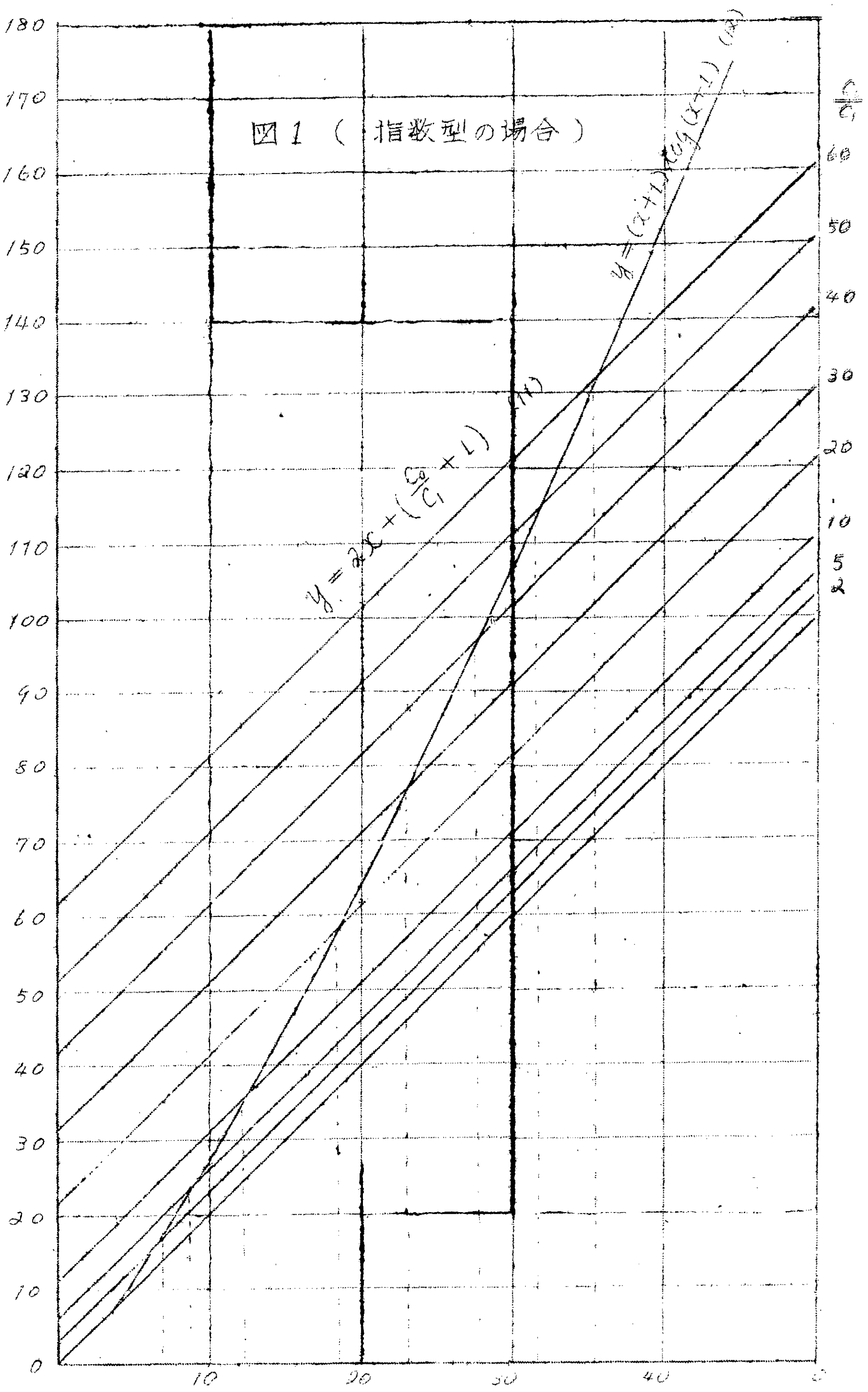


図１（ 指数型の場合 ）
$y = 2x + \left(\frac{C_2}{C_1} + 1\right)$ (1)
$y = (x+1)\log(x+1)$ (2)
$\frac{C_2}{C_1}$
60
50
40
30
20
10
5
2
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
10
20
30
40

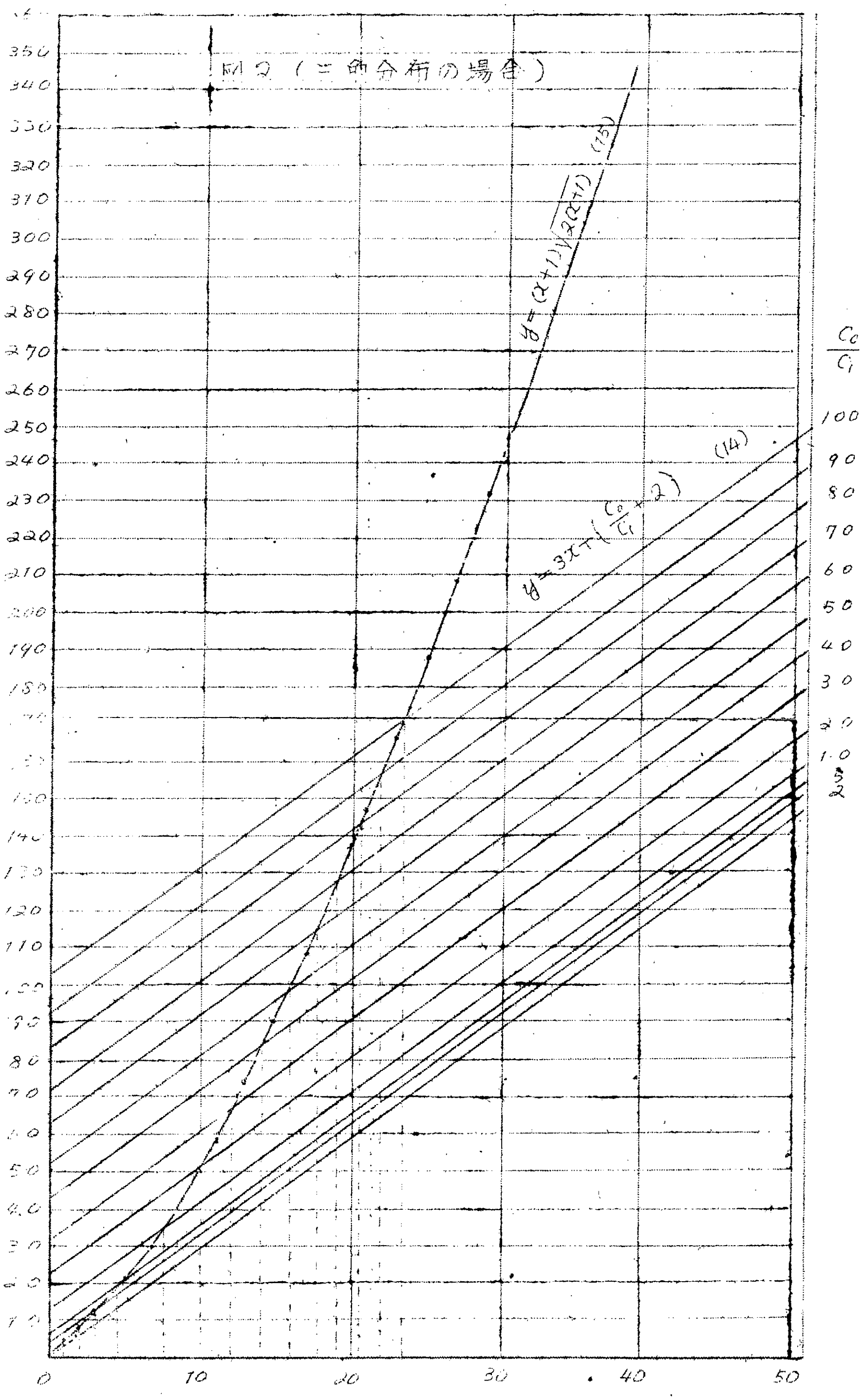


図2（三角分布の場合）
$y=(x+1)\sqrt{2(x+1)}$ (15)
$y=3x+\left(\frac{C_0}{C_1}+2\right)$ (14)
$\frac{C_0}{C_1}$
100
90
80
70
60
50
40
30
20
10
5
2
350
340
330
320
310
300
290
280
270
260
250
240
230
220
210
110
120
130
140
100
90
80
70
60
50
40
30
20
10
0
10
20
30
40
50