

最有効確率抽出法とその適用について

統計数理研究所 田口 時夫

はしがき

歪みの大きい母集団からの標本抽出を扱う場合、各個体或いは階層に対する抽出確率を変える事によって推定が可能、或いは有効となる事は容易に予想されるが、その根拠及び実現の方法についての理論は尚不充分であるといえる。

本稿は一般的不等確率抽出法を基礎にして、その適用の具体的な条件と有効限界の検討を通じて最有効確率抽出法を見出すと共に、更にそれを組織的に実現する為の諸方法を与えるものである。

§1. 最有効確率抽出法の理論的基礎

一般に不等確率抽出法とは、補助変量 Y が与えられた場合、その正值一価関数 $\varphi(Y)$ に比例する抽出確率をもつ抽出法と考える事が出来る。今此の方法によって得られた標本 $(X^{(i)}, Y^{(i)}); i=1, \dots, n$ を用いて主変量 X の平均 $E(X)$ を推定する

爲に、不偏推定量

$$(1) \quad \tilde{X}(\varphi) = \frac{1}{n} E\{\varphi(Y)\} \sum_{i=1}^n \frac{X_i^{(0)}}{\varphi(Y_i)}$$

を用いるものとするとき、その分散 $\sigma_{\tilde{X}(\varphi)}^2$ は

$$(2) \quad \sigma_{\tilde{X}(\varphi)}^2 = \frac{1}{n} [E\{\varphi(Y)\} E\{\frac{X^2}{\varphi(Y)}\} - E^2(X)]$$

となる。

すなわち、本稿で扱う最有効確率抽出法とは、次の定義によるものとする。

定義 1. 「凡ての Y の関数 φ の集合に於て、 $\sigma_{\tilde{X}(\varphi)}^2$ を最小にする $\varphi^{(0)}$ が存在するとき、 $\varphi^{(0)}(Y)$ に比例する確率抽出法を、 X の平均を推定する爲に $\tilde{X}(\varphi)$ を用いた場合の最有効確率抽出法という」

此の場合容易に定理 1 が得られるが、その説明に先立って次の補助定理を証明しよう。

補助定理 1. 「 $\varphi(y)$ が y の連続正值一価関数であれば、任意の y の連続一価関数 $\psi(y)$ に対して

$$(3) \quad [E\{e^{\psi(Y)}\}]^2 \leq E\{\varphi(Y)\} E\left[\varphi(Y) \left\{ \frac{e^{\psi(Y)}}{\varphi(Y)} \right\}^2\right]$$

が成立する。」

証明. 一般に $\phi(y)$ が連続な凸関数であれば、 $\phi(y) > 0$ のとき Jensen の不等式により

$$(4) \quad \phi\left(\int p \cdot f dy / \int p dy\right) \leq \int p \phi(f) dy / \int p dy$$

が成立する。従つて確率変数 Y に対して

$$(5) \phi \{ E(p \cdot f) / E(\phi) \} \leq E \{ p \cdot \phi(f) \} / E(\phi)$$

となる。こゝで

$$(6) \phi(f) = f^2, \phi(Y) = \phi(Y), f(Y) = \frac{e^{\psi(Y)}}{\phi(Y)}$$

とおくと (3) が得られる。(証明おわり)

補助定理 1 は ϕ, ψ が不連続の場合に容易に拡張することは出来る。

定理 1. 「主変数 X が正值で、且つ

$$(7) \log X = \psi(Y) + \log \zeta, \quad \zeta \perp Y$$

と表わされたならば、 $e^{\psi(Y)}$ に比例する確率抽出法が最も有効な確率抽出法を与える。又その場合 $\tilde{X}(\phi^{(0)})$ の変動係数 $C_{\tilde{X}}^{(0)}$ は

$$(8) C_{\tilde{X}}^{(0)} = \frac{1}{\sqrt{n}} \frac{\sigma_X}{E(\zeta)} = \frac{1}{\sqrt{n}} C_\zeta$$

である。こゝで C_ζ は ζ の変動係数を表わす。更に (7) の条件下では

$$(9) \psi(Y) = E(\log X | Y)$$

であり、従って $\log \zeta$ は $\log X$ の (9) による回帰残差を表わす。

証明. (7) の条件下で、任意の ϕ に対して $\sigma_{\tilde{X}(\phi)}^2$ は (2) により

$$\begin{aligned} (10) \quad \sigma_{\tilde{X}(\phi)}^2 &= \frac{1}{n} [E \{ \phi(Y) \} E \{ \frac{X^2}{\phi(Y)} \} - E^2(X)] \\ &= \frac{1}{n} [E \{ \phi(Y) \} E \{ \phi(Y) \{ \frac{e^{\psi(Y)}}{\phi(Y)} \}^2 \} - E^2(X)] \end{aligned}$$

となるから、補助定理により

$$\begin{aligned} (11) \quad \sigma_{\tilde{X}(\phi)}^2 &\geq \frac{1}{n} [E \{ e^{\psi(Y)} \} E(\zeta^2) - E^2(X)] \\ &= \frac{1}{n} [E \{ e^{\psi(Y)} \} E \{ e^{\psi(Y)} \zeta^2 \} - E^2(X)] \\ &= \frac{1}{n} [E \{ e^{\psi(Y)} \} E \{ \frac{X^2}{e^{\psi(Y)}} \} - E^2(X)] = \sigma_{\tilde{X}(e^\psi)}^2 \end{aligned}$$

が成立する。これは $e^{\psi(Y)}$ に比例した確率抽出法が最も有効であることを示す。又此の時

$$(12) \quad C_X^{(0)} = \frac{1}{\sqrt{n}} \frac{\sqrt{\bar{X}(e^{\psi})}}{E(X)} = \frac{1}{\sqrt{n}} \sqrt{\frac{E(Z^2)}{E(Z)}} = \frac{1}{\sqrt{n}} \frac{\sqrt{Z}}{E(Z)} = \frac{1}{\sqrt{n}} C_Z$$

となる。更に (7) 式の両辺に於いて Y における条件付期待値を考えると、

$$(13) \quad E(\log X | Y) = \psi(Y) + E(\log Z)$$

となり $E(\log Z)$ は Y に依存しない「常数」であるから、 $e^{\psi(Y)}$ に比例は $E(\log X | Y)$ に比例となる。(証明終わり)

上記の証明は補助定理1によらずに直接 Schwarz の不等式を適用しても得られる。又此の定理はベクトル値補助変数 $Y = (Y_1, \dots, Y_k)$ の場合に拡張され得る。

定義2. $\log X$ の $Y_1, \dots, Y_k$ に関する回帰関数を $R(Y)$ といったとき、 $e^{R(Y)}$ に比例する確率抽出法を回帰関数比例確率抽出法と定義する。特に回帰関数が回帰係数 $\beta_1', \dots, \beta_k'$ をもつ $\log X$ の $\log Y_1, \dots, \log Y_k$ に関する最小二乗線形回帰の場合は $\prod_{j=1}^k Y_j^{\beta_j'}$ に比例する確率抽出法となるが、これを特に最小二乗対数線形回帰比例確率抽出法と「し」

此の場合次の定理が得られる。

定理2. $\log X$ の Y に関する (9) 式の回帰関数を (10) 式の条件を除いて一般に

$$(14) \quad R(Y) = E(\log X | Y)$$

とすれば, $e^{K(Y)}$ 比例確率抽出法は, 最有効確率抽出法の一つの近似を与える。又最小二乗対数線形回帰比例確率抽出法は, 最有効確率抽出法他の近似となり得る。又其の降出等の有効性は, それぞれ $\log X$ の $Y_1, \dots, Y_k$ に関する重相関比 $\gamma'_{X(1 \dots k)}$ 及び $\log X$ の $\log Y_1, \dots, \log Y_k$ に関する重相関係数 $\rho'_{X(1 \dots k)}$ によって判定され得る。

証明. Y の任意の正值一価関数に対して

$$(15) \quad X = \zeta_\varphi \varphi(Y)$$

と表わすと, $\sigma_{X(\varphi)}^2$ は (2) により

$$(16) \quad \sigma_{X(\varphi)}^2 = \frac{1}{n} [E\{\varphi(Y)\}^2 E\{\zeta_\varphi^2(Y)\} - E^2\{\zeta_\varphi \varphi(Y)\}]$$

と表わされる。従つて ζ_φ が 1 に近く且つ変動が小さいであれば, つまり $\log \zeta_\varphi$ が 0 に近く且つその分散が小さいような φ が最有効条件に近いといえる。今 $\varphi(Y)$ として $e^{E(\log X|Y)}$ を用いた時の ζ_φ を特に ζ_e と表わすと

$$(17) \quad E(\log \zeta_e) = 0, \quad V(\log \zeta_e) = \min_{\varphi} V(\log \zeta_\varphi)$$

が成立するから, 上記の条件に近く従つて一つの最有効確率抽出法の近似を与える。又 $\log X$ の $\log Y_1, \dots, \log Y_k$ に関する最小二乗対数線形回帰の残差を $\log \zeta_a$ とすれば, 一般に ζ_a は Y_j ($j=1, \dots, k$) と無相関であり, これは $\zeta_a \perp Y_j$ の必要條件であるから, 定理 1 により此の抽出法は最有効確率抽出法他の一つの近似を与える。以上の証明に於て, ζ_e や ζ_a

従って $\log \zeta_e$ や $\log \zeta_a$ の分散がそれぞれ抽出法の有効性の基準を与えることが理解される。それは又 $\gamma'_{x(i-k)}$ や $\rho'_{x(i-k)}$ が判定基準となる事を意味し、定理が証明される。

以上の結果は、次の例題にみるように所得分布に対し特に有用である。

例1. 「 (X, Y) が $n+1$ 次元対数正規分布に従うものとする。此の場合最も有効確率抽出法は、最小二乗線形回帰比例確率抽出法である。又此の抽出法による $X^{(0)}$ の変動係数を $C_X^{(0)}$ とすれば

$$(18) C_X^{(0)} = \frac{1}{\sqrt{n}} \sqrt{e^{1 - \rho_{x(i-k)}^2} \sigma_x'^2 - 1}$$

となる。ここで $\sigma_x'^2$ は $\log X$ の分散を表わす。」

例2. 「 (X, Y) が2次元才2種 Pareto 分布 (Mardia [1] 参照) をもつ場合

$$(19) E\left\{\log\left(\frac{X}{a}\right) \mid Y\right\} = (1-\beta)p^{-1} + \beta p^{-1} \log\left(\frac{Y}{a}\right)$$

となるので、此の場合も最小二乗線形回帰比例確率抽出法が最も有効確率抽出法の近似を与える。」

2. 最も有効確率抽出法の実際的諸問題

最も有効確率抽出法と、その近似方式としての回帰関数比例確率抽出法 (前節参照) を具体的に適用する場合には次のような諸問題が生じる。第1は補助変数の選択方法であり、次に層別法の適用条件の考慮と標本の各層への割当の問題

が生じる。更に又、標本の抽出確率の近似計算法を挙げる事が出来る。

(1) 補助変量の選択について

最有効確率抽出法に於ては、母平均の推定精度の基準として、推定量の分散の代りに変動係数を用いる方が適当である。(前節参照)。その場合層別を無視すると、補助変量の選択は定理1及び2により、 $\log X$ を $R(Y) = E(\log X | Y)$ の同帰した残差を $\log r$ とした時、その変動係数を基準とすることが出来る。従って実際上は、 $E(\log r)$ は1に近いものと考えて、 $e^{\frac{X}{R(X)}}$ の分散又は $\log X$ と Y との相関比 $\gamma_{\log X, Y}$ を基準とすることが考えられる。故に $\log X$ と $\log Y$ とが線形に近い関係にあれば、 $\log X$ と $\log Y$ との相関係数が実用的である。

如く一般には、調査項目は $X_1, \dots, X_k$ と複数個に亘るので、正準相関等の利用が予想されるが、特に対数正規母集団に於ては、近似的に項目毎の変動係数を基準として、それを最小とする変量を補助変量とすることが出来る。(拙稿[2] 参照)。その場合最適抽出確率は、此の補助変量に関する各 X_i の対数線形同帰係数 β_i' の平均 $\bar{\beta}'$ を用いて $Y^{\bar{\beta}'}$ に比例する抽出確率となる。

(2) 層別法について

最有効確率抽出法は又独自の層別法を与えることが出来る。

つまり此の方法に於ける層別の本質的な意義は、層化を通じて各層内の残差を Y と独立に近い状態に近づけることである。此の事は層別法を定理1における最有効確率抽出法の存在条件を実現させる手段とする事を意味する。他方、此の場合の層別は、抽出作業の技術的簡約化と関連させて行なうことが出来る。例えば $\log X$ の Y に對する回帰関数 $R(Y)$ を折線で近似すれば、各折線の交点 y_j を各層の分点とする事により層化 e^{dY} 比例確率抽出法が得られる。又 $R(Y)$ を $\log Y$ との関係を捉えて折線で近似すれば、層化 Y^d 比例確率抽出法が得られ、 d が1に近い層は規模比例確率抽出法となる。又 $R(Y)$ を階段関数で近似すれば、跳びの生ずる点を各層の分点とする事により通常の層化無作為抽出法を与えることができる。

(3) サンプルの割当について

(2)で述べた層別法は、層の分点と共に各層内の抽出形態を同時に決定するが、サンプルの割当に關しては、改めて従来の最適割当の諸方式を適用する方が有効である。勿論この場合各層の抽出形態が一般的には不等確率抽出であることを考慮して σ_x^2 の代りに

$$(20) \quad \sigma_x'^2 = E\{\varphi(Y)\} E\left\{\frac{X^2}{\varphi(Y)}\right\} - E^2(X)$$

を用いねばならぬ。猶量用條件等は最適割当を行なうこの段階で考慮することが出来る。

(4) 抽出確率の近似について

各層内の各個体の抽出確率は簡単な数値で示される事が望ましい。その為には2次層化を行なうことが考えられるが、その場合の層化基準は、各層内の Y の値のオーダーを揃えること、従って 10^5 を層の分点とすることであらう。つまりその時(2)式をもとにして、与えられた誤差の許容範囲での Y, β, Y^2 等の必要相数を与えうるからである。すなわち与えられた正数 ε に対して最有効抽出確率を与える $\varphi^{(0)}(y)$ の間に

$$(21) \quad (1-\varepsilon)\varphi^{(0)}(y) \leq \varphi_a^{(0)}(y) \leq (1+\varepsilon)\varphi^{(0)}(y)$$

の関数にある $\varphi_a^{(0)}(y)$ を用いて、 $\varphi_a^{(0)}(y)$ 比例確率抽出法を用いるものとしよう。その場合(2)式により

$$\begin{aligned} (22) \quad \sigma_{\bar{x}(\varphi_a^{(0)})}^2 &\leq \frac{1}{n} \left[\frac{1+\varepsilon}{1-\varepsilon} E\{\varphi^{(0)}(Y)\} E\left\{\frac{X^2}{\varphi^{(0)}(Y)}\right\} - E^2(X) \right] \\ &= \frac{1}{n} \left[(1+2\varepsilon) E\{\varphi^{(0)}(Y)\} E\left\{\frac{X^2}{\varphi^{(0)}(Y)}\right\} - E^2(X) + O(\varepsilon^2) \right] \\ &= \sigma_{\bar{x}(\varphi^{(0)})}^2 + 2\varepsilon \left\{ \sigma_{\bar{x}(\varphi^{(0)})}^2 + \frac{E^2(X)}{n} \right\} + O(\varepsilon^2) \\ &= \sigma_{\bar{x}(\varphi^{(0)})}^2 \left[1 + 2\varepsilon \left(1 + \frac{1}{C_x^{(2)}} \right) \right] + O(\varepsilon^2) \end{aligned}$$

但し (23) $C_x^{(2)} = \frac{\sigma_{\bar{x}}^2}{E(X)}$ である。

従って与えられた許容範囲規定量 η

$$(24) \quad 0 < \left\{ \sigma_{\bar{x}(\varphi_a^{(0)})}^2 - \sigma_{\bar{x}(\varphi^{(0)})}^2 \right\} / \sigma_{\bar{x}(\varphi^{(0)})}^2 < \eta$$

に対して ε は

$$(25) \quad \varepsilon < \frac{\eta}{2(1+C_x^{(2)})}$$

とすればよい。

13. 抽出法決定の爲のデータ解析と最有効確率抽出法のシミュレーション

(1) 抽出法決定の爲のデータ解析

最有効確率抽出法の適用例として、補助変量を資本金とした場合の法人企業統計調査を考えてみよう。才1図の各調査項目の資本金に対する対数相関図をみると、一般的に可成りの相関が認められる。従つて単に総額の推計のみを目的とする場合は、特に層別を加えずに、最小二乗対数線形回帰比例確率抽出法の単純な適用が可能と思われるが、才1図をより詳細にみると、全体として曲帰曲線は $\log Y$ 軸に凸になる傾向が窺える。更に回帰曲線による残差に着目すると、各資料(才1図)を通じて明らかに $\log Y$ との相関が認められる。つまり対数残差 $\log \epsilon$ の分散は、一般に $\log Y$ が小なる程大である。従つて上記の回帰関数比例確率抽出法を、最有効確率抽出法に接近させる爲には、層別により各層内の残差と $\log Y$ との相関を小さくすることが望ましい。此の観点から層別を考えると、資本金額14万〜5千万階層と5千万〜1億階層とは特に区別する必要が認められながら、他の階層は略安当であると思われる。但し10億以上層は更に100億以内を階層に分することもある。

他方、才1表は各階層別に各抽出法(等確率、規模比例及

び最小二乗対数線形同帰確率比例の諸抽出法)による標本分散規定量の $\frac{1}{x_n}$ を計算した結果を示すものであるが、同帰関数比例確率抽出法が此の面でも最も有効であり、特に等確率抽出法に対しては、時に1桁小さい分散を得るのである。

(2) 最も有効確率抽出法のシミュレーション

表2表は、資本金10億、100億の法人階層を対象として上記諸抽出法のシミュレーションを行った結果を示すものである。

この際、同帰関数比例確率^(抽出法)は各項目 x_n の資本金 y に関する最小二乗対数線形同帰係数 β_n の項目平均 $\bar{\beta}_n$ を用いて $y^{\bar{\beta}_n}$ に比例する抽出確率を用いる抽出法を意味する。各表の全体を通じて同帰関数比例確率抽出法が相対的に安定したよい結果を与えることが理解されよう。然し乍ら此の程度の精度の向上は、箱期待を下廻るものであり、その原因は此の階層における各項目の資本金に対する対数相関が比較的低いことにあると推測される。此の意味で、層内の対数相関をたかめる層別法についての研究が今後必要となるであらう。

むすび

概、標本設計に於て有効性をたかめることは、如何なる意義をもつであらうか。勿論それは統計数値の信頼性をたかめる反面、標本数の削減による調査の効率化等の直接的効用をもつものであるが、それを経費の節減に直結させることにつ

いは大きな疑惑を抱かざるを得ない。標本精度の有効性の向上によって得られる経費の余剰は、更に統計調査を質的・内容的に充実させる為に標本設計や集計・分析及び調査諸手当の単価の増大に振向けるべきであらう。そして此の方向が真に統計を発達させ、社会に寄与することは云ふまでもないことであらう。

参考文献

- [1] Mardia, K.V, Multivariate Pareto distribution,
Ann. Math. Statist. Vol. 33 (1962)
- [2] 田口 時夫, 最有効確率抽出法と同帰関数比例確率抽出法, 「統計数理研究所彙報」, 第24巻, 第1号 (予定)
- [3] 田口 時夫, 同帰抽出法の提唱について, 「理論・計量経済学会 1975 年度大会」の口頭発表・報告要旨 理論・計量経済学会
- [4] 田口 時夫, ある種の最適確率抽出法の実用化とそれに直した層別法等について, 「サンプリングの数理的研究予稿集」, 京都大学経済研究所

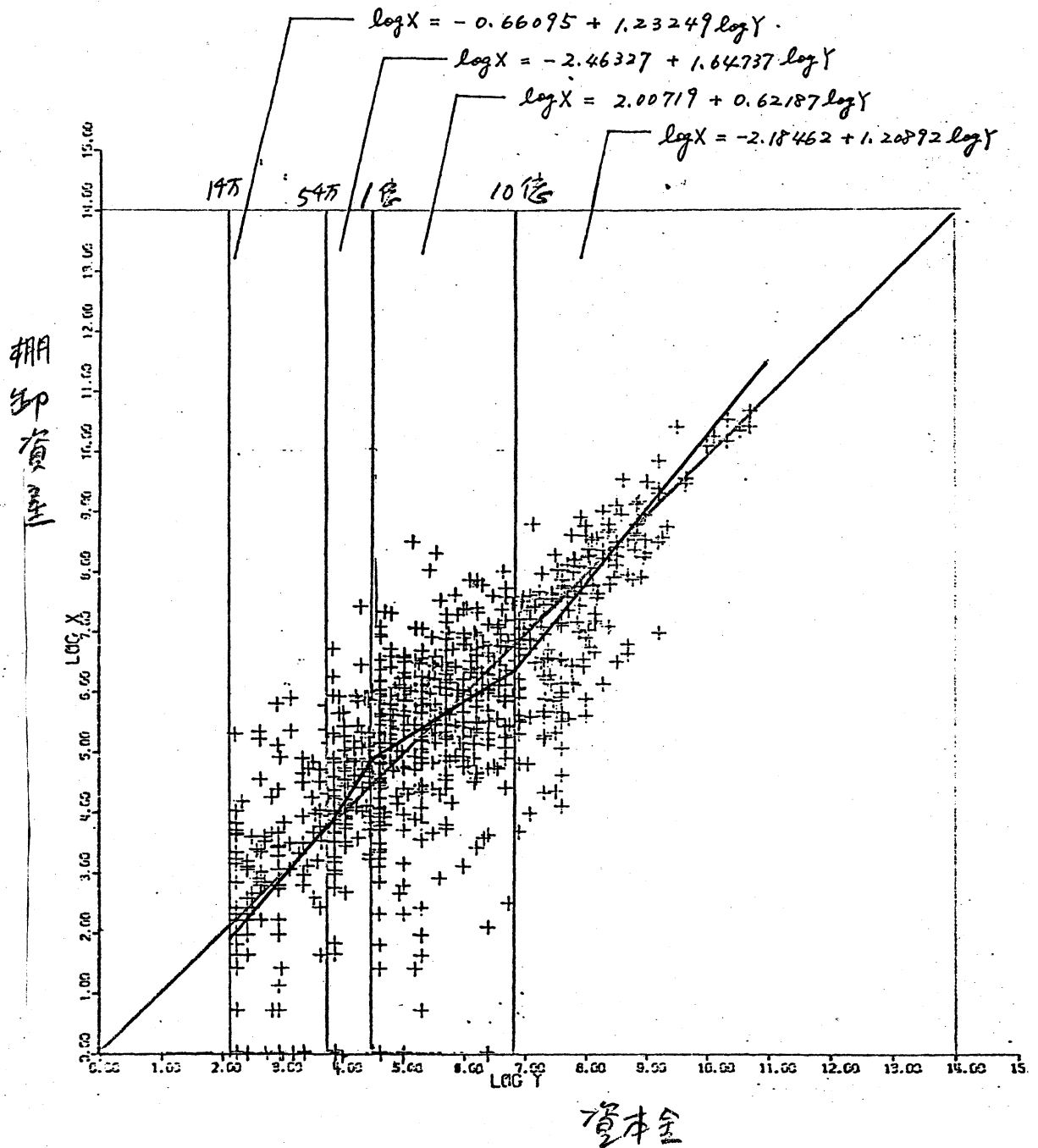
資料

大蔵省法人企業統計調査, 昭和48年4~6月期季報用調査, 行及承統 267号, 発証第1446号.

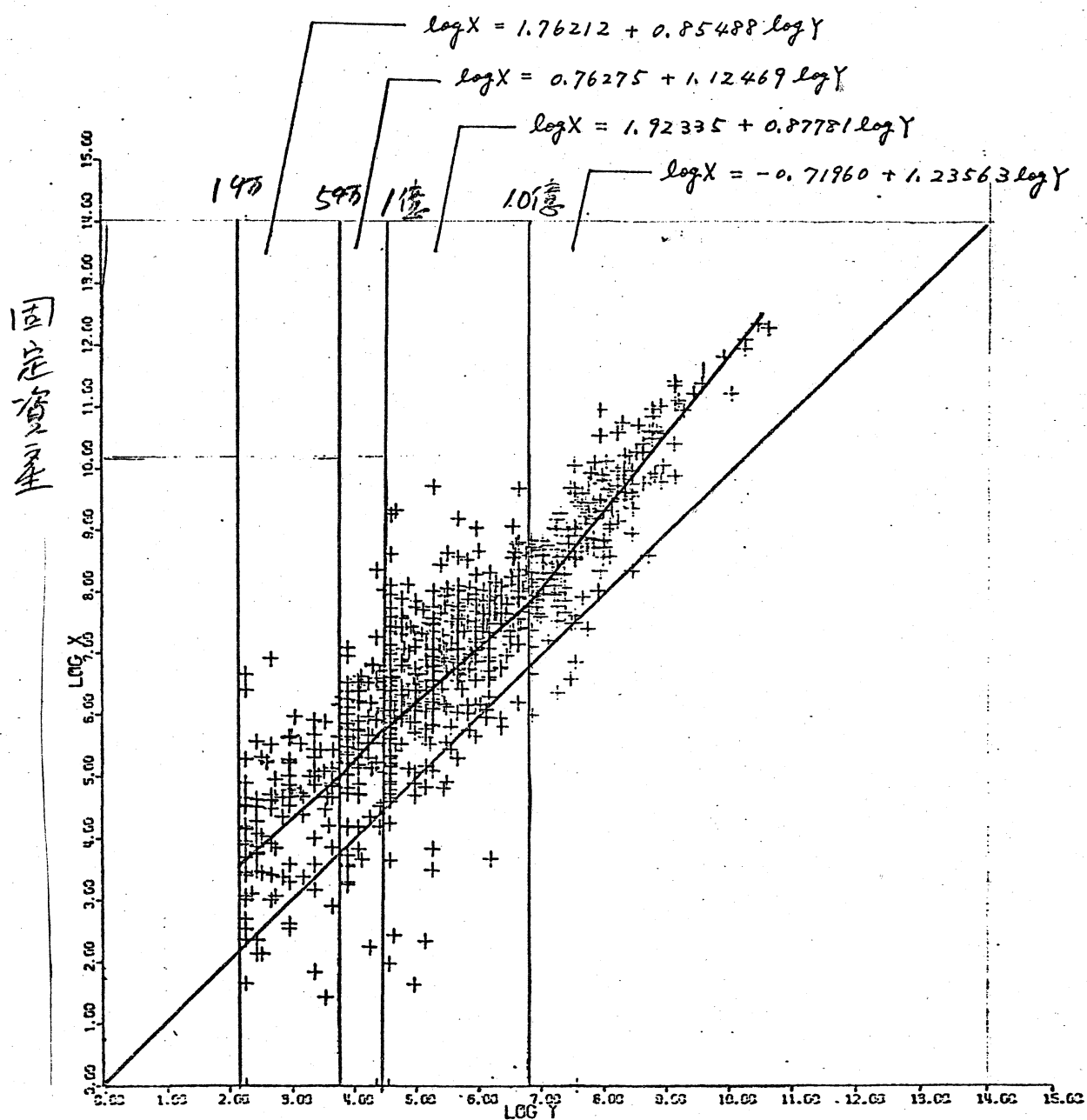
第1回 資本金に対する相関図

(産業・化学、資本金14万円以上の全標準企業、)
(資料・昭和48年4月~6月期分法人企業統計による)

(i) 棚卸資産

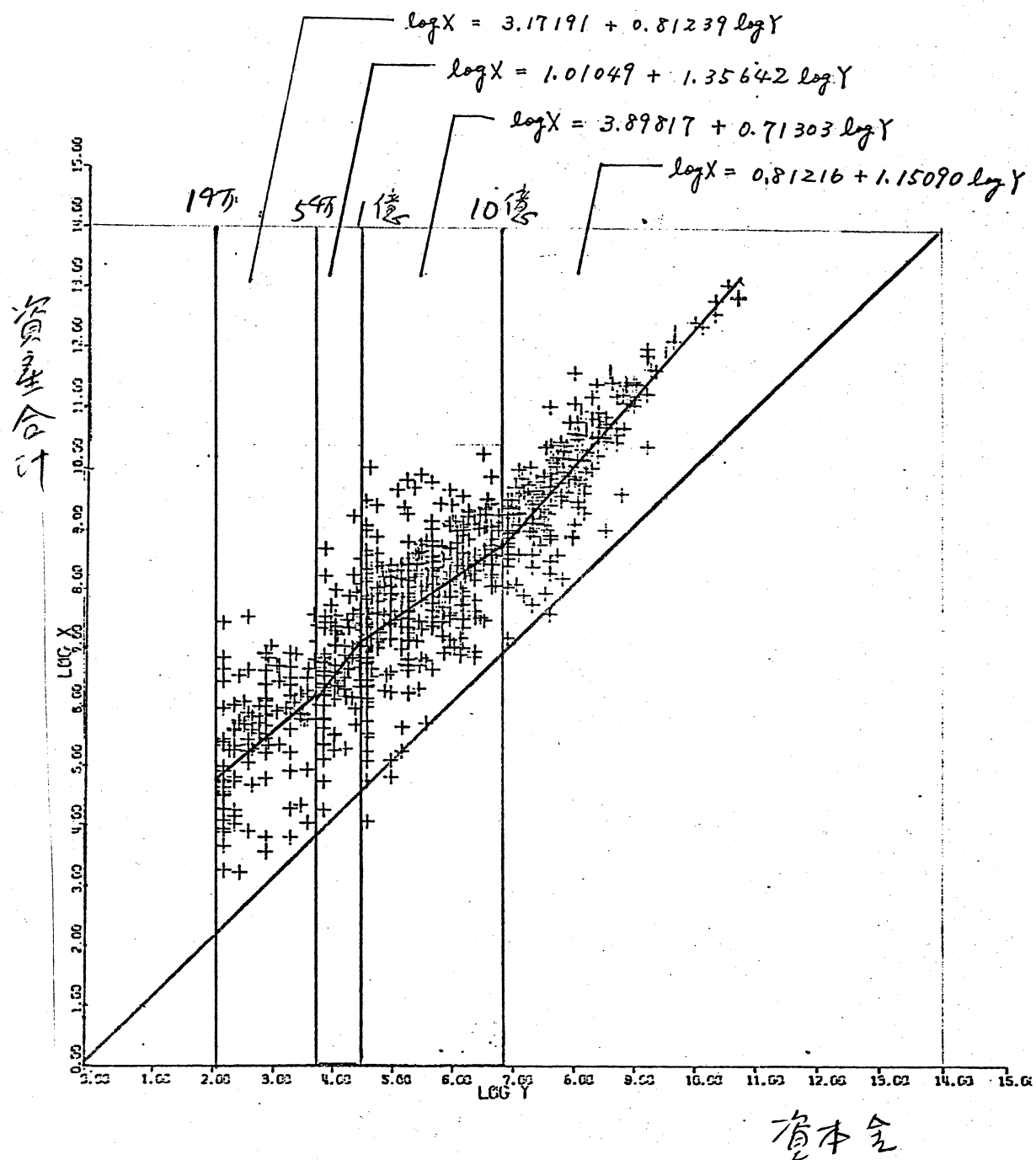


(ii) 固定資產

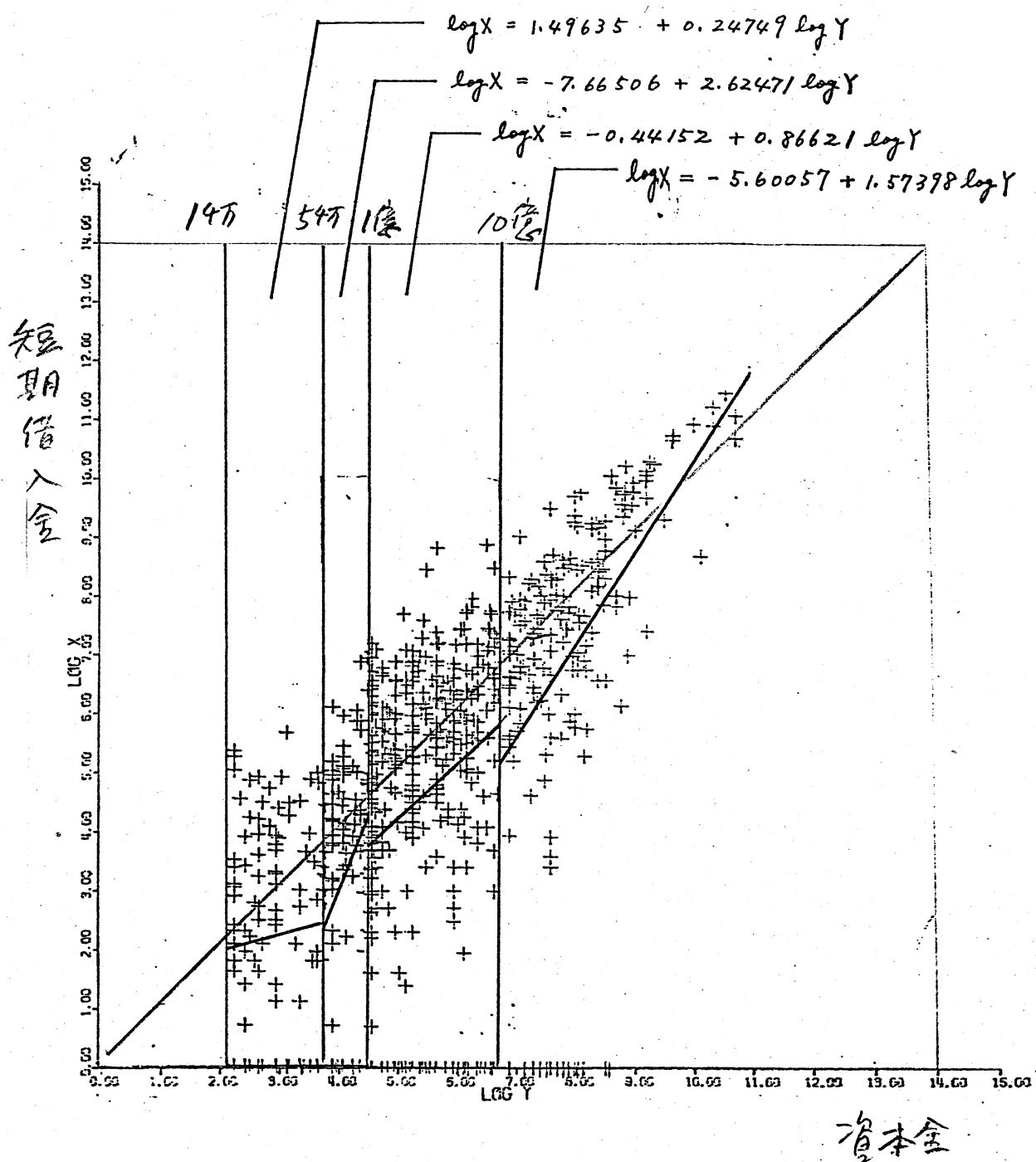


資本金

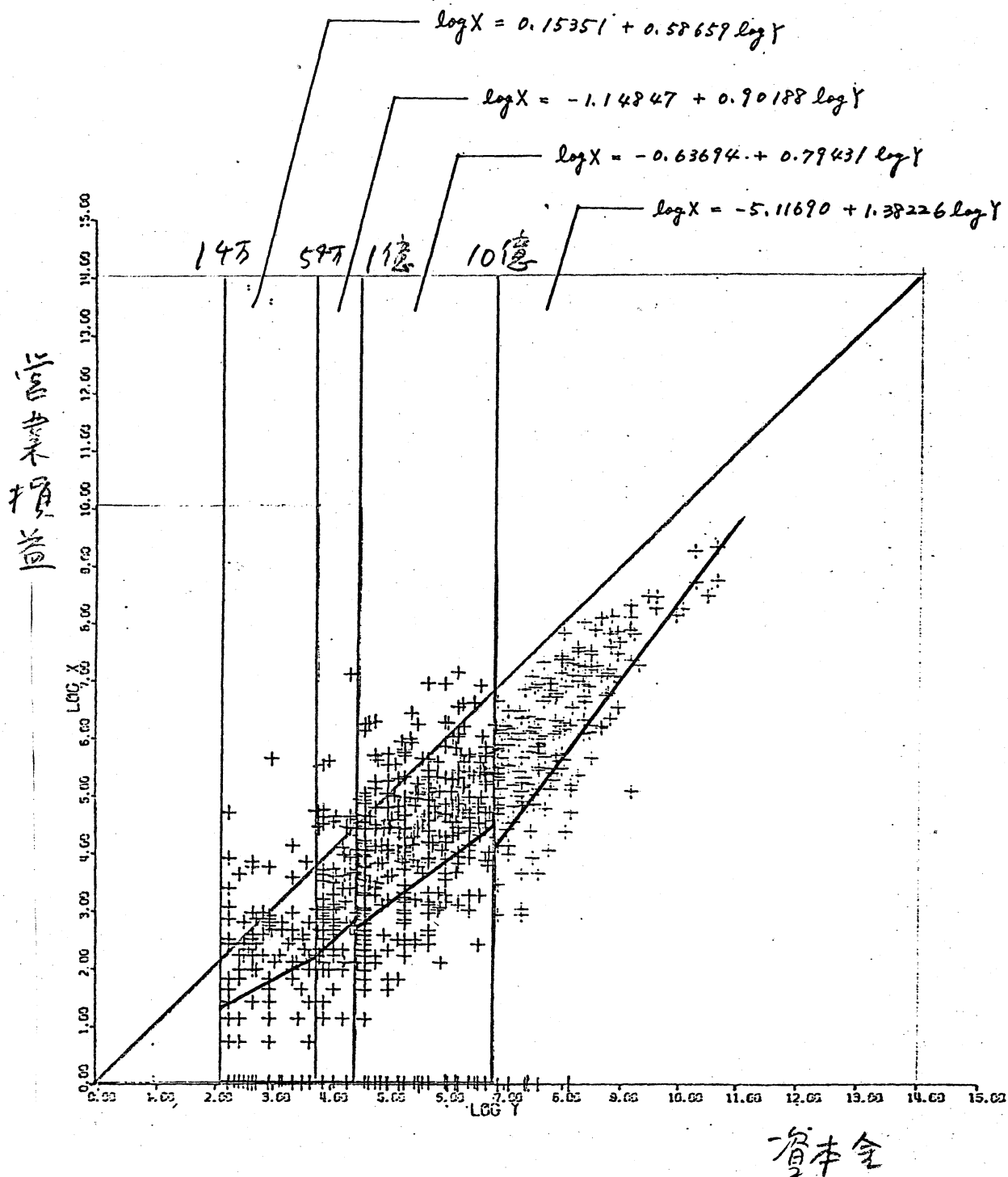
(iii) 資產合計



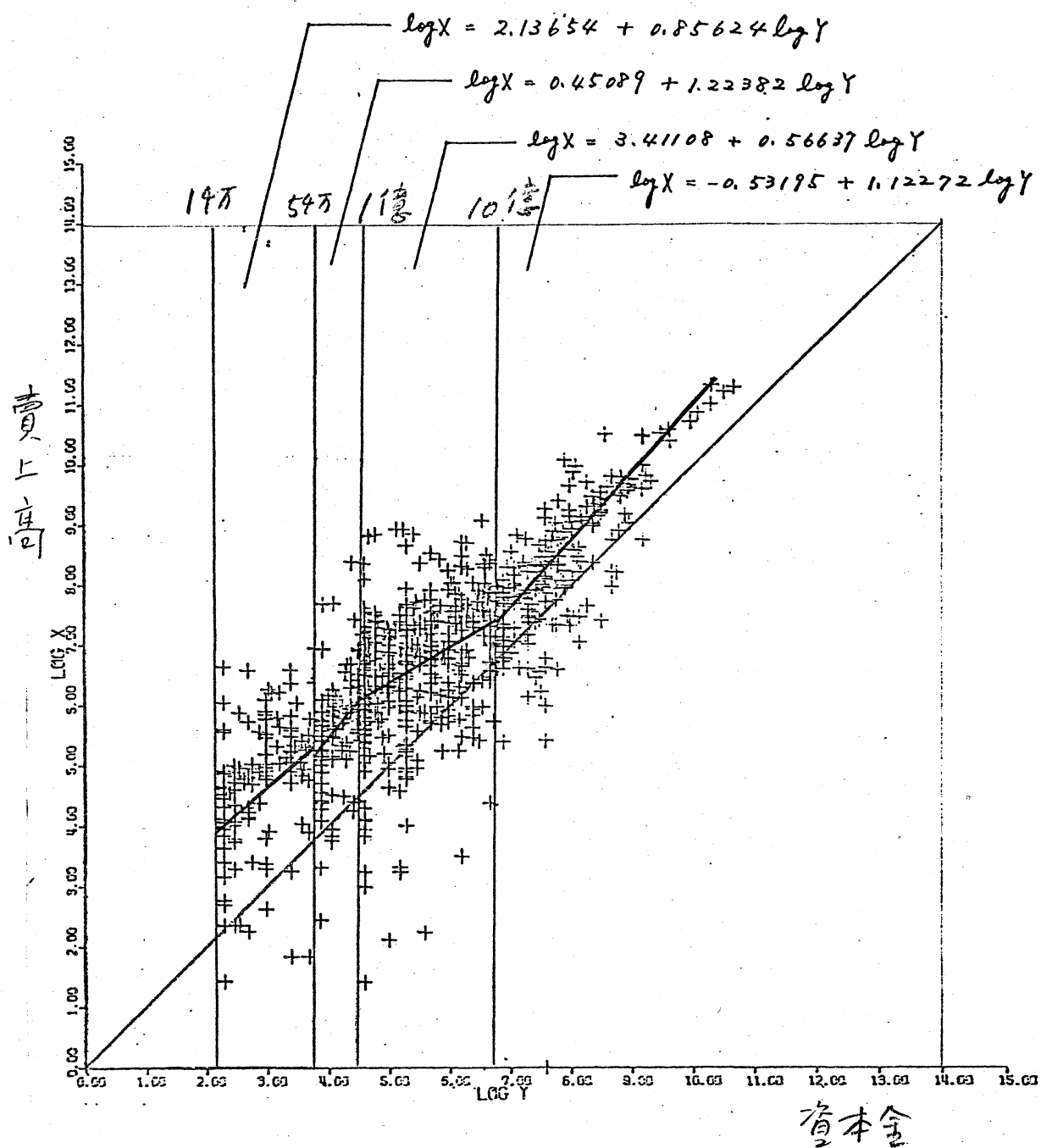
(IV) 短期借入金



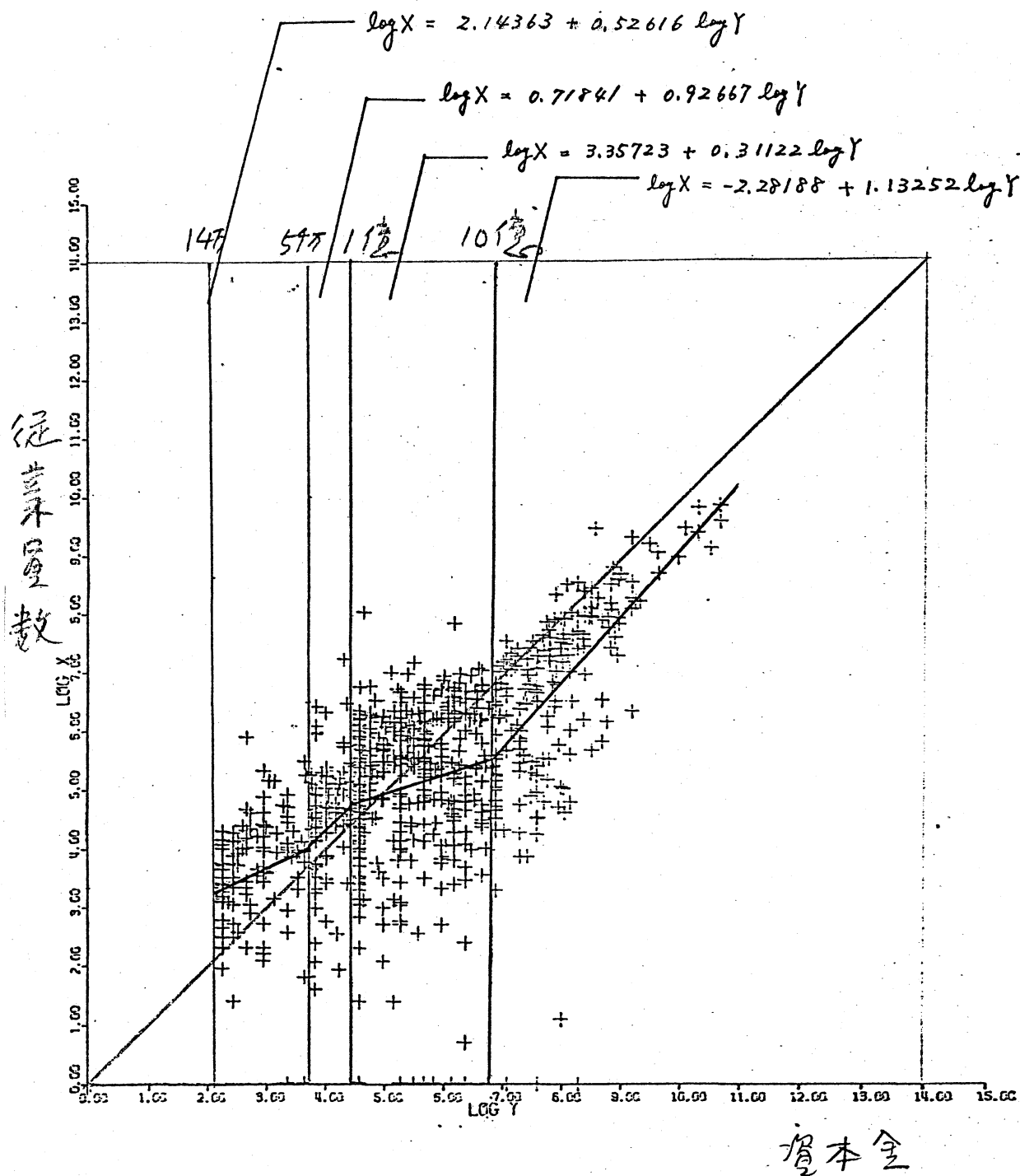
(V) 營業損益



(VI) 賣上高



(VII) 従業員数



(11) 永南助交量。元基置救

$z - 1$	$E(x)$	$V(x)$	$C(x)$	$V_0(x)$	$T_1(x)$	$V_2(x)$	$C_2(x)$	β_0'	β_1'	ρ_{xy}'
1.	7269.	330714880.	2.50168	5534416.	6271536.	5679328.	0.32783	0.03901	0.96989	0.91417
2.	19249.	2066684900.	2.36174	45229568.	42857216.	45924352.	0.35206	0.50805	1.03141	0.91896
3.	14541.	1265996200.	2.44692	21099840.	17784672.	21695712.	0.32032	0.02447	1.05123	0.92488
4.	21088.	2456137900.	2.35014	56949760.	48956672.	58083072.	0.36140	0.02594	1.10108	0.94464
5.	72257.	29409419000.	2.37336	334155770.	24730090.	348057600.	0.25819	1.57969	1.06169	0.96356
6.	13883.	734957050.	1.95278	46013440.	43839888.	45604432.	0.48644	1.01768	0.94768	0.88095
7.	11116.	973282300.	2.80649	43870736.	136540080.	44260432.	0.58848	2.75737	0.59097	0.23033
8.	5706.	226376990.	2.63679	18124192.	19081536.	18163440.	0.74689	0.52515	0.92340	0.40091
9.	1707.	20333200.	2.64224	1362067.	1932659.	1360428.	0.68345	4.11135	1.22531	0.57651
10.	19533.	1910721200.	2.23779	68896512.	68428288.	69188096.	0.42583	0.64307	1.02411	0.92874
11.	7234.	214922410.	2.02656	0.	0.	3984.	0.00873	0.00000	1.00000	1.00000
$\overline{P} \pm \overline{\beta}$									0.99334	0.78942

(三) 在单位变量上高

$2-f$	$E(x)$	$V(x)$	$C(x)$	$V_0(x)$	$V_1(x)$	$V_2(x)$	$C_2(x)$	β_0'	β_1'	β_2'
1.	7269.	330714880.	2.50168	3430656.	4121008.	3745312.	0.26623	-1.43846	0.91691	0.95224
2.	19249.	2066684900.	2.36174	20721664.	19978752.	20760576.	0.23671	-1.08075	0.97680	0.95892
3.	14541.	1265996200.	2.44692	16436096.	16440304.	17692496.	0.28927	-1.44271	0.98064	0.95063
4.	21088.	2456137900.	2.35014	22469632.	24030976.	25392384.	0.23896	-1.64503	1.03522	0.97858
5.	72257.	29409419000.	2.37336	0.	0.	41246720.	0.08888	0.00000	1.00000	1.00000
6.	13883.	734957050.	1.95278	40572544.	28997792.	32594896.	0.41124	-0.38477	0.89187	0.91350
7.	11116.	973282300.	2.80649	43651696.	95357776.	44039468.	0.59699	1.08483	0.63447	0.27247
8.	5706.	226376990.	2.63679	16790912.	16322752.	16304784.	0.70765	-2.69435	0.94779	0.45340
9.	1707.	20353200.	2.64224	1083394.	1485802.	1016964.	0.59091	-5.67276	1.12843	0.58499
10.	19533.	1910721200.	2.23779	42425600.	38072576.	37577728.	0.31382	-0.76949	0.95370	0.95295
11.	72334.	214922410.	2.02656	4734944.	2294176.	3013632.	0.23997	-0.80116	0.87451	0.96356
F_{11}	$\bar{P}_1, \bar{P}_2$								0.94003	0.81648

(iv) 補助要量、資金合計

$z - \frac{1}{2}$	$E(x)$	$\cdot V(x)$	$C(x)$	$V_0(x)$	$V(x)$	$V(x)$	$C_2(x)$	β'_0	β'_1	β'_{24}
1.	7269.	330714880.	2.50168	13033376.	15868560.	14562960.	0.52496	0.33991	0.84537	0.87864
2.	19249.	2066684900.	2.36174	53668352.	55811232.	65149184.	0.41932	0.41366	0.94530	0.92873
3.	14541.	1265996200.	2.44692	65124032.	68398784.	70458304.	0.57726	0.47652	0.90220	0.87528
4.	21088.	2456137900.	2.35014	50970368.	51261952.	65867520.	0.38486	0.08052	0.98600	0.95278
5.	72257.	29409419000.	2.37336	568602620.	608305150.	767270910.	0.38335	1.66911	0.95220	0.95295
6.	13883.	734957050.	1.95278	42890816.	35524720.	35352688.	0.42829	1.03912	0.85647	0.87794
7.	11116.	973282300.	2.80649	87991040.	303475450.	89385328.	0.85050	4.33750	0.35902	0.15430
8.	5706.	220376990.	2.63679	33032752.	36650784.	32360576.	0.99694	0.56834	0.71468	0.34216
9.	1707.	20353200.	2.64224	700279.	1447883.	815873.	0.52927	5.78492	1.29749	0.67316
10.	19533.	1910721200.	2.23779	0.	0.	11145216.	0.17091	0.00000	1.00000	1.00000
11.	7234.	214922410.	2.02656	9801264.	8168256.	7942000.	0.38957	0.57319	0.84224	0.92874
β_1, β_{24}								0.88191		0.77679

表2 最少二乗対数線形回帰比例確率
抽出法のシミュレーション結果(1)
(但レ 実現率 = $\frac{\text{標本推定平均}}{\text{母平均}}$ とする)

補助変量 機械 10億 100億 1/2 抽出 $\bar{\beta}_1 = 0.82350$	Simulation No	等確率		規模比例		回帰関数($\bar{\beta}_1$)比例	
		平均実現率	実現率の 二乗偏差	平均実現率	実現率の 二乗偏差	平均実現率	実現率の 二乗偏差
	①	1.16015	0.00822	0.95220	0.00367	1.01280	0.00782
	②	1.04883	0.00229	1.07275	0.01285	0.99996	0.00662
	③	0.96259	0.00214	0.84075	0.00685	0.86358	0.00489
	④	1.08197	0.00523	0.94861	0.00228	1.13707	0.01609
	⑤	0.78653	0.00431	0.92178	0.00714	0.86541	0.00479
	⑥	0.91285	0.00312	1.02883	0.00300	0.99397	0.00334
	⑦	0.94394	0.00911	1.10323	0.01390	1.07417	0.00671
	⑧	0.85639	0.01018	0.98102	0.00435	0.95815	0.00653
	⑨	0.98826	0.00390	0.82320	0.00862	0.96251	0.00606
	⑩	0.91731	0.00460	0.88659	0.01336	0.89630	0.01611
	平均	0.96588	0.00531	0.95590	0.00760	0.97639	0.00790

補助変量 電気 10億 100億 1/2 抽出 $\bar{\beta}_1 = 0.85603$	Simulation No	等確率		規模比例		回帰関数($\bar{\beta}_1$)比例	
		平均実現率	実現率の 二乗偏差	平均実現率	実現率の 二乗偏差	平均実現率	実現率の 二乗偏差
	①	0.74754	0.00414	0.90305	0.00294	0.98108	0.00985
	②	0.99339	0.03117	1.18193	0.00676	1.07986	0.00389
	③	0.88875	0.00152	1.09394	0.00474	0.93313	0.00206
	④	0.81477	0.00431	1.01317	0.00435	0.99618	0.00338
	⑤	1.21271	0.01641	1.08137	0.00198	0.89082	0.01047
	⑥	0.83859	0.00484	0.91318	0.00228	0.93770	0.00168
	⑦	0.83362	0.00729	0.89277	0.00450	0.89048	0.00613
	⑧	0.78460	0.00664	1.05391	0.00343	1.06070	0.00351
	⑨	0.96584	0.00720	1.03994	0.00496	0.99104	0.00879
	⑩	1.18683	0.01322	1.00847	0.00600	1.14388	0.00489
	平均	0.92666	0.00967	1.01817	0.00419	0.99049	0.00557

抽出法のシミュレーション結果 (2)

補助変量 機械 10億 100億 1/2 抽出 $\bar{R}=0.83911$	Simulation NR	等確率		規模比例		回帰係数 (Y^R)	
		平均実現率	実現率の 二乗偏差	平均実現率	実現率の 二乗偏差	平均実現率	実現率の 二乗偏差
	①	0.99725	0.00230	0.94540	0.00744	0.98574	0.00971
	②	1.32889	0.01021	0.84734	0.01126	0.83193	0.01172
	③	0.84987	0.00194	0.92289	0.00161	0.97356	0.00518
	④	0.93880	0.00922	0.99399	0.00727	1.00418	0.00834
	⑤	0.99347	0.01251	0.98494	0.00479	0.92367	0.00848
	⑥	1.05502	0.01070	1.00973	0.00224	1.01957	0.00392
	⑦	0.77918	0.00654	0.95073	0.02109	1.03850	0.00170
	⑧	1.23427	0.01219	1.00675	0.00296	1.06025	0.00455
	⑨	0.98567	0.02366	0.85934	0.01185	0.88238	0.00688
	⑩	1.13293	0.00299	0.85379	0.00666	0.84691	0.02084
	平均	1.02954	0.00923	0.93749	0.00772	0.95667	0.00813

補助変量 電気 10億 100億 1/2 抽出 $\bar{R}=0.77533$	Simulation NR	等確率		規模比例		回帰係数 (Y^R)	
		平均実現率	実現率の 二乗偏差	平均実現率	実現率の 二乗偏差	平均実現率	実現率の 二乗偏差
	①	1.34996	0.01629	0.99157	0.00858	1.01200	0.01499
	②	0.82482	0.00258	1.02103	0.00472	1.01000	0.00718
	③	0.94718	0.00947	0.98016	0.00372	0.98405	0.00503
	④	1.00313	0.00618	0.94988	0.00815	0.93937	0.00235
	⑤	0.79130	0.00840	0.93653	0.01449	0.94958	0.00230
	⑥	1.00742	0.00272	1.02788	0.01265	0.97233	0.00730
	⑦	0.81502	0.00464	1.03546	0.00562	0.98361	0.00186
	⑧	0.93338	0.00496	0.86391	0.00662	0.91406	0.00255
	⑨	0.90189	0.00672	0.98250	0.00238	1.02756	0.00223
	⑩	0.83932	0.00684	0.95596	0.00604	0.98588	0.00841
	平均	0.94134	0.00688	0.97449	0.00730	0.97784	0.00542

抽出法 a シミュレーション結果 (3)

補助変量 - 資本金

機械 10 億 ~ 100 億 $\frac{1}{2}$ 抽出

項目	母平均	等確率		規模比例		回帰係数(平均)比例	
		$\hat{X}_{0,h}$	(実現率)	$\hat{X}_{1,h}$	(実現率)	$\hat{X}_{\bar{h},h}$	(実現率)
1. 資本金	2852	2887	(1.012)	2852	(1.000)	2919	(1.023)
2. 売掛金	9592	11411	(1.190)	9062	(0.945)	8819	(0.919)
3. 棚卸資産	6179	7596	(1.229)	6234	(1.009)	6560	(1.062)
4. 固定資産	8512	9135	(1.073)	7378	(0.867)	8204	(0.976)
5. 資産合計	31151	36509	(1.172)	29124	(0.935)	30374	(0.975)
6. 買掛金	7042	9191	(1.305)	6756	(0.959)	6618	(0.940)
7. 短期借入金	6126	7428	(1.213)	5226	(0.853)	5855	(0.956)
8. 長期借入金	3743	4379	(1.170)	3972	(1.061)	4134	(1.104)
9. 営業損益	574	671	(1.169)	578	(1.007)	711	(1.239)
10. 売上高	6628	8153	(1.230)	6100	(0.920)	6526	(0.985)
11. 従業員数	2574	2569	(0.998)	2359	(0.916)	2473	(0.961)
平均			(1.160)		(0.952)		(1.013)