

# 測定値誤差・測定値変動と相関分析

— 医学におけるデータ処理の一特性 —

林 知己夫 (統計数理研究所)  
福田 安平 (東鉄保健管理所)  
細谷 亮子 (統計数理研究所)  
林 文 (統計数理研究所)

(1967年10月受付)

## Response Errors and Correlation Analysis

— Some Problems on Data Analysis in Medical Science —

Chikio HAYASHI (Institute of Statistical Mathematics)  
Yasuhei FUKUDA (Tokyo Health Institute of Japan  
National Railway)  
Ryoko HOSOYA (Institute of Statistical Mathematics)  
Fumi HAYASHI (Institute of Statistical Mathematics)

The authors treat the problem of estimation in correlation analysis in the case where the relevant variables, for example  $X$ ,  $Y$ , are subject to errors or fluctuations in measurements. If errors or fluctuations are represented by random variables, the bias of estimation of  $Y$  by  $X$  can be calculated. The examples of distortion in data analysis in medical research are shown which originates in disregarding the response errors or fluctuations in spite of the variables  $X$ ,  $Y$  being subject to them. The authors present some statistical models in treating the errors or fluctuations and give the statistical method of obtaining valid conclusions through  $X$ ,  $Y$  which are subject to errors or fluctuations, in both cases where  $X$  and  $Y$  are quantitative and qualitative.

The Institute of Statistical Mathematics

測定値に測定誤差がある場合や測定値に変動がある場合、これを無視して、ある集団において統計的分析をほどこすとしたら、非常に誤った結末に到ることがある。以下このことについて述べてみようと思う。

従来の統計的分析では、どうしたことか、このことについて深刻な議論がなされていないように思われる。測定誤差に言及されてはあっても、形式的な取扱いに終わっている。さて、まず次のように考えよう。

$$x = x_0 + \epsilon:$$

$x$  は測定値,

$x_0$  は真の値,

$\epsilon$  は測定誤差をあらわす確率変数

$\epsilon$  は  $E(\epsilon) = 0, E(\epsilon^2) = \sigma_\epsilon^2$  とする.  $x_0$  はある集団における測定値をあらわす確率変数とする.  $x_0$  と  $\epsilon$  とは独立としよう.  $E(\epsilon) \neq 0$  の場合も当然あるが, ここでは簡単のため  $E(\epsilon) = 0$  としておく. このとき  $x$  の集団での分散  $\sigma_x^2$  は,

$$\sigma_x^2 = \sigma_0^2 + \sigma_\epsilon^2$$

:  $\sigma_0^2$  は  $x_0$  の集団での分散

となる.  $E(\epsilon) \neq 0$  とすれば  $x$  の平均自乗誤差  $\tau_x^2$  ( $x_0$  の平均のまわり) は

$$\tau_x^2 = \sigma_0^2 + \sigma_\epsilon^2 + E(\epsilon)^2$$

となる.

このようなことを基礎におく分析の場合には, さして致命的な誤りは生じないのであるが, 以下に述べるような分析をするときは「単なる測定誤差」が分析の「歪み」となってあらわれ結果に重大な影響を及ぼしてくることになる.

健康管理のとき得られるデータについての例をあげながら, このような問題についての分析法を示してみよう. 健康管理のときなどに測定し得られるデータには測定誤差が大きかったり, またそのときの測定値そのものは生体活動によって生ずる変動を含んでおり, しかも測定誤差や変動が相当大きいものがあるのである. まず測定値が量で与えられている場合, 次に質(分類, カテゴリー)で与えられている場合について述べる.

### 1. 測定値が数量である場合

健康管理において測定された最大血圧値に関する分析をみよう. ある集団のT年における最大血圧値を測定し, 同一人についてT+1年においても最大血圧を測定する. これをつきあわせたとき, どのような様相が出るか.

第1表 最大血圧の推移表 (40~50才)

T年 \ T+1年	0~79	80~89	90~99	100~109	110~119	120~129	130~139	140~149	150~159	160~169	170~179	180~189	190~199	200以上	計	%
	0~79	0		1	2	3	3	2								
80~89	1	8	33	41	22	5	3	1							114	0.13
90~99		38	357	567	385	142	38	7		2					1537	1.81
100~109	2	41	707	2565	2478	1118	363	86	18	15	5	2			7400	8.71
110~119		30	486	2950	6004	4196	1590	475	106	22	10	4	1		15884	18.70
120~129	1	5	128	1453	4722	6569	3820	1554	384	124	47	9	6	4	18826	22.16
130~139			31	397	1837	4183	4841	2819	951	359	128	27	11	6	15590	18.36
140~149		2	5	92	518	1499	2886	3215	1552	774	309	96	27	11	10986	12.94
150~159			2	8	104	370	940	1460	1404	871	402	146	67	34	5808	6.84
160~169				11	32	122	337	738	923	845	487	208	92	53	3848	4.53
170~179				6	13	33	127	277	456	486	459	243	110	83	2293	2.70
180~189				2	7	16	55	101	188	244	244	207	128	97	1289	1.52
190~199						7	17	37	61	109	132	120	103	119	705	0.83
200以上					4	2	18	23	53	78	90	96	118	279	761	0.85
計	4	124	1760	8094	16129	18265	15038	10795	6098	3929	2313	1158	663	686	85056	
%	0.01	0.14	2.07	9.52	18.96	21.47	17.68	12.69	7.17	4.62	2.72	1.36	0.78	0.81		100

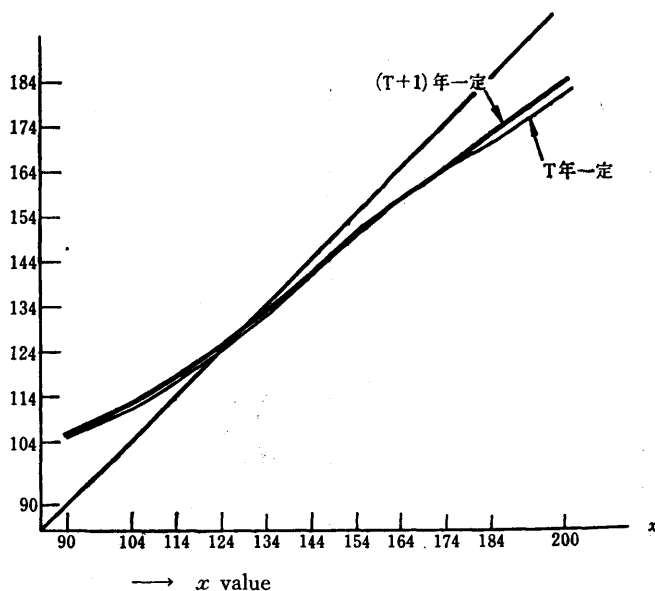
データとして次のものを示そう(第1表)。これは、T年、(T+1)年における同一対象の最大血圧測定値の関係を集計したものである。対象は、40才-50才の某企業体の職員についてのものである。血圧値は、定期健康診断時(4~6月)のもので、Tとしては昭和36年(したがって(T+1)年は昭和37年)のものである。なお、この対象は、昭和35年から血圧検診を始めている。また、同一時期に何回も測定したものもあったが、すべて第1回目の値を用いた。T年に検診を受けた対象の第2年目の受検率は98.9%であった。(T+1)年における未受検者は、受検者にくらべT年における血圧150 mmHg以上において高率となっていた(第2表)。

第2表 受検者・未受検者血圧分布(T年におけるもの)

	100 mm 以下	101~119	120~149	150~169	170~	計
受 検 者	1.96	27.37	53.38	11.35	5.94	100%
未受検者	1.66	23.42	49.74	14.82	10.36	100%

しかし今回の分析においては、未受検者の問題は、関係がないことなので、分析の対象から除外し、T年・(T+1)年ともにデータのあるもののみを取扱うことにする。

第1表から直ちに気のつくことは、T年において、血圧が平均より低目のところでは、(T+1)年の血圧値はT年より高目に出るものが多い、一方、T年において、血圧の高かったものは(T+1)年において血圧がT年の測定値より低目に出るものが多い、ということが非常に顕著に見られることである。単に、これからみると、この1年間に、血圧の低かったものは高くなるように、血圧の高かったものは低くなるようになにかをした(あるいはそのようなことがあった)と結論したい気持ちになる。こうしたとき、いわゆる統計的検定を行えば、きわめて有意と出てくることは明らかである。ここになにか、実態的に原因を求めたくなるわけである。もう少し注意してみよう。(T+1)年の側からみるのである。このときも同様に、(T+1)年に血圧の低いものはT年ではそれより高い血圧を示したものが多く、(T+1)年で血圧の高いものは、T年では、それより血圧の低い値を示したものが多いということが読みとれる。このことは、T年をもとにしたときと全く同様のことで、裏返したようなことが成立している。

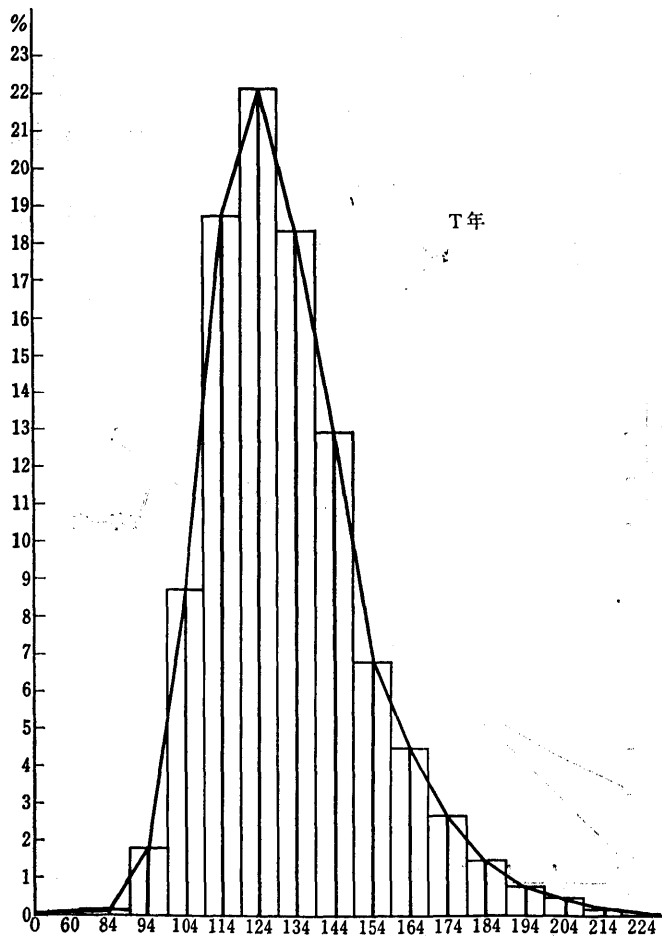


第1図

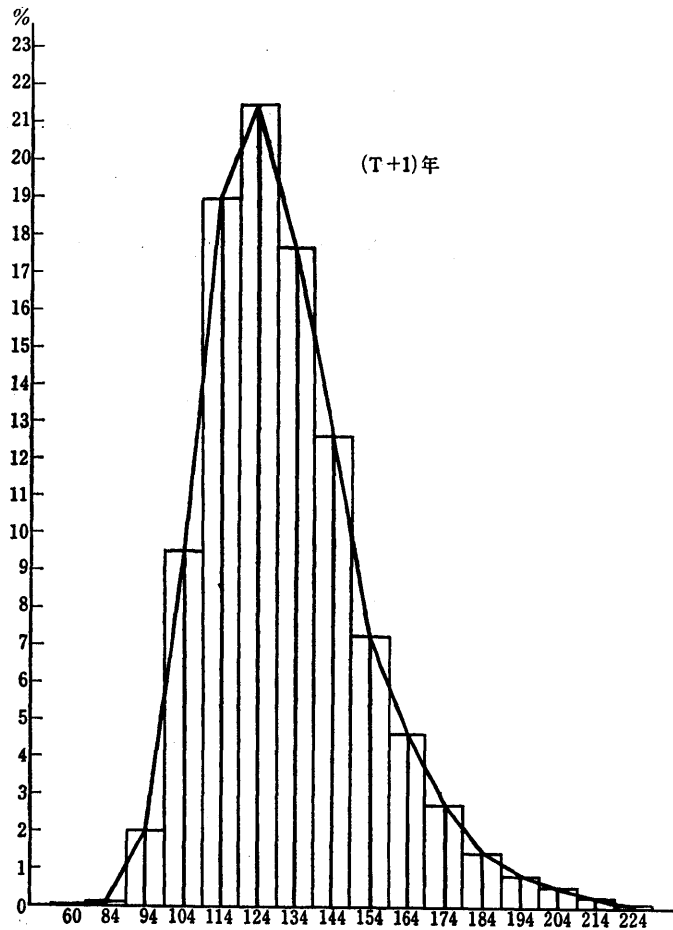
T年或は(T+1)年一定としたとき、それぞれ(T+1)年或はT年の血圧値の平均値を目もったものを第1図に示しておいたが明瞭な関係がみられよう。こうなると、T年をもとにしたときの現象に実態的に原因を求めるといふことに疑問を生じてくる。もし、T年をもとにしたときの現象に実態的な原因があるとすれば、逆に、(T+1)年の側からみたときにも、原因を求めなくてはならなくなる。しかも、これらが同一の原因であることはまずありえない。「時」は逆行しえないからである。ちなみに、T年、(T+1)年の集団平均値、分散、分布を比較してみると、まったく同一であった(第3表、第2図、第3図)。

第3表

	T 年	(T+1) 年
平均	131.21	131.63
分散	441.9442	440.9591



第2図



第3図

なお、分布をつくる時 200 mmHg 以上という分類しかデータとしてとられていないので、このままでは、分析が十分行えないので、曲線が滑らかになるようにかつ分布曲線の性質を失わぬように外挿して分布曲線を求めてある。

こうしたデータをみると、前述の事象のおこるのは、測定値が、誤差をもつ或は変動をもつためではないかとの予想が出てくる。

ちなみに測定値に変動あるいは誤差があるとき、相関係数にしても、大きく影響を受ける可能性がある。たとえば

$$x = x_0 + \varepsilon, \quad y = y_0 + \eta$$

:  $x_0, y_0$  は測定誤差のないときの値,  $\varepsilon, \eta$  はそれぞれ測定誤差で

$$E(\varepsilon) = 0, \quad E(\eta) = 0$$

$$E(\varepsilon^2) = \sigma_\varepsilon^2, \quad E(\eta^2) = \sigma_\eta^2$$

( $\varepsilon, \eta$  は平均0で、分散はそれぞれ  $\sigma_\varepsilon^2, \sigma_\eta^2$  としておく)、で、すべて相互に独立性を仮定しておく。

としたとき  $(x_0, y_0)$  間のある集団での相関係数を  $\rho_{x_0 y_0}$  としたとき、 $(x, y)$  のその集団での相関係数  $\rho_{xy}$  は、待望値の計算により

$$\rho_{xy} = \rho_{x_0y_0} \frac{1}{\sqrt{\left(1 + \frac{\sigma_\epsilon^2}{\sigma_{x_0}^2}\right) \left(1 + \frac{\sigma_\eta^2}{\sigma_{y_0}^2}\right)}}$$

となる。 $\sigma_{x_0}^2, \rho_{y_0}^2$  はその集団での誤差のないとしたときの測定値  $x_0, y_0$  の分散とする。

$\sigma_\epsilon^2/\sigma_{x_0}^2 \geq 0, \sigma_\eta^2/\sigma_{y_0}^2 \geq 0$  であるから、 $\rho_{xy} \leq \rho_{x_0y_0}$  となる。いま  $\sigma_\epsilon^2/\sigma_{x_0}^2 = 1/4, \sigma_\eta^2/\sigma_{y_0}^2 = 1/4$  つまり、誤差分散が、誤差のない測定値の集団での分散にくらべ、その  $1/4$  となっているとすれば、

$\rho_{xy} = \frac{4}{5} \rho_{x_0y_0} = 0.8 \rho_{x_0y_0}$  となり顕著に相関係数は落下する。誤差の程度が大きくなるとこの割合はますます激しくなる。見かけの相関係数（誤差を含む）は、真の相関係数（誤差なし）よりも小さくなるわけである。

さて、それはともあれ、われわれの血圧値は、どれほど変動するものであろうか。変動のあることは、従来言われていたが、健康管理のときの測定で——もちろんしかるべき一定条件（例えば；検定を経た血圧計で測定する；安静時間は5分以上をとって測定する；測定は2回目の値をとる；放尿後測定する；水銀のおろし方は1搏動 2~4 mmHg 程度とする；など）の下に測定する——どのくらい変動するかが、循環器管理研究協議会の血圧小委員会で研究され発表されている（文献 [1], [2] 参照）のでこの値を参考としてみよう。これによると一応最大血圧において 45 (mmHg)、最小血圧において 35 (mmHg) 程度の分散があることが示されている。これは非常に大きいものである。なお、そこでは、月別変動は別とすれば、週間変動、時刻変動（10時—16時の間）ではさして大きいものではないと示されている。上述の値は、これ以外の容易に捕捉しえない要因によって変動することを意味する。これは人間の生体変動、環境条件による変動を含むものと言えよう。なお上述の値は、前述の一定の条件の下での測定であるが、週間変動、時刻変動、その他の要因による変動を含んだものである。これらは、健康管理の測定の上ではさげられないものであろう。このように眺めてみると、血圧変動として大きい値を示すことがわかったので、この変動（誤差と言っておこう）のために、分析上の問題がおこるのではないかどうか、検討を加えてみよう。

なお、これまで、或はこれからも変動と言ったり誤差と言ったりしているが、分析の上では同じものと見做していただきたい。誤差とは、われわれの情報をおおいにかく雑音という意味であるから、生体変動による変動であっても、平均値としての情報をおおいにかく誤差とみなすわけである。いずれにしても真の値（あるいは平均値）のまわりの変動を誤差という表現で述べてみる。したがって、測定者の読みの誤差もあるし、測定条件や生体の変動によるものも含まれているわけである。同じ時期に2回以上測定して標識づけたとき同一の値が示されないときすなわちデータに変動あるときこれを「誤差があるため」と表現して取扱うのである。誤差をこのように値の変動を含めたひろい意味に用いることにする。

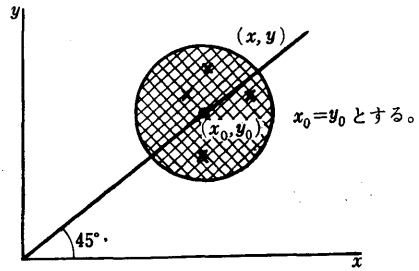
前述のように全体の集団の分散は 441 程度であるが、これには測定誤差が含まれている。測定誤差は変動の分散を上へのべた議論からよくコントロールされた条件下で一応 50 と見做しておき、かつ見透しをよくするため各人の変動の分散も一応同一と見做しておけば測定誤差のないときの分散  $\sigma_0^2$  は

$$\sigma_0^2 = 391$$

となる。前に述べた誤差分散と  $\sigma_0^2$  との比は、0.13 程度でかなり大きい。さて、こうした変動は一応ガウス分布と見做しておく。つまり平均 0、分散  $s^2$  のガウス分布とみなしておく。 $s^2 = 50$  としておくわけである。この確率的仮定も、循環器管理研究協議会の報告（文献 [1]）のデータからみて、妥当な仮定と考えられる。さて、こうなると2回測定を行えば第4図のように、測定誤差のないとき  $(x_0, y_0)$ 、ただし  $x_0 = y_0$ 、であるものが測定されたとき、それら

の値はある確率をもって円形の範囲にばらつくことが考えられる。

このような  $(x, y)$  が観測されるはずである。このような人々が多くあれば  $x_0 = y_0$  であっても  $45^\circ$  の上にのらずに散らばってくることになる。T年及び (T+1) 年において、測定誤差がなければ、等しい値を持つ、すなわち上の表現に従えば、各人について  $x_0 = y_0$  であるとしておこう。これは、前に示した T年、(T+1) 年の平均、分散、分布の形の恒常性から一応認められよう。



第4図

さて、測定誤差のないときの血圧値の分布を簡単のため、ガウス分布としておく、これは、さきのデータからみるとやや無理な仮定であるように見える——実際は分布が少し流れ、ペアソンⅢ型の恰好である——が、計算が容易で、見透しをよくするため一まずこうしておく。ペアソンⅢ型にしても以下の計算は可能であるが、見透しのよい形で書けないだけである。

これを

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-M)^2}{2\sigma^2}} dx = \Phi(x|M, \sigma^2) dx$$

としておく。M は平均値、 $\sigma^2$  は分散である。これまでの表現に従えば、本来  $x$  でなく  $x_0$  を用うべきであるが、煩雑をさけるため添字をおとして  $x$  としておく。T年の測定誤差のない血圧値の分布も、(T+1) 年のそれも同一であるとしておく。したがって、T年の M も  $\sigma^2$  も (T+1) 年の M も  $\sigma^2$  も同一であり、かつ T年である  $x$  を示すものは、(T+1) 年でもその  $x$  をもつとしておく。さて測定誤差  $\epsilon$  の分布は平均 0、分散  $s^2$  のガウス分布

$$\frac{1}{\sqrt{2\pi}s} e^{-\frac{\epsilon^2}{2s^2}} d\epsilon$$

であるとしておくので、測定誤差のないとき血圧値  $x$  を示すものの測定値  $z$  の分布は

$$\frac{1}{\sqrt{2\pi}s} e^{-\frac{(z-x)^2}{2s^2}} dz$$

となる。  $s^2 = g(x)$  と  $x$  の関数であることも予想されるが、一応  $s^2$  は常数としておく。  $s^2 = g(x)$  でも計算は、面倒になるが可能である。

さて、測定誤差のないとき  $x$  のものが測定で  $(u, u+du)$  の間に入る確率は

$$\varphi(u, x) du = \frac{1}{\sqrt{2\pi}s} e^{-\frac{(u-x)^2}{2s^2}} du$$

となるから、本来  $x$  のものが T年で  $(u, u+du)$  の値を示し、(T+1) 年で  $(u+d)$  以上の値を示す確率は、測定が独立であると考えてよいから、

$$E(u, d) = \int_{-\infty}^{\infty} \varphi(u, x) \Phi(x) dx \int_{u+d}^{\infty} \frac{1}{\sqrt{2\pi}s} e^{-\frac{(z-x)^2}{2s^2}} dz / C$$

$$C = \int_{-\infty}^{\infty} \varphi(u, x) \Phi(x) dx$$

であらわされる。  $\Phi(x)$  は  $\Phi(x|M, \sigma^2)$  を略記したものである。

これはガウス分布の仮定のもとに計算を実行すれば、

$$E(u, d) = \frac{1}{\sqrt{2\pi}} \int_{A+Bd}^{\infty} e^{-\frac{t^2}{2}} dt$$

となることが示される。

ただし、

$$A = \frac{(u-M)s}{\sqrt{(s^2 + \sigma^2)(s^2 + 2\sigma^2)}}$$

$$B = \frac{(1 + \sigma^2/s^2)s}{\sqrt{(s^2 + \sigma^2)(s^2 + 2\sigma^2)}}$$

T年のときの測定値で  $u$  を示し、(T+1)年 で  $u$  以上の測定値を示す確率は

$$E(u, 0) = \frac{1}{\sqrt{2\pi}} \int_A^\infty e^{-\frac{t^2}{2}} dt$$

となる。

$u-M > 0$  のとき

$A > 0$  となるから、 $E(u, 0) < \frac{1}{2}$  となる。これは T年に  $u$  を示したものは (T+1)年 で  $u$  以上の測定値を示す確率が  $1/2$  より小となる、すなわち、T年の測定で集団平均以上を示したものは (T+1)年の測定ではそれ以上の値を示すものが少なくなる、つまり  $u$  以下の値を示すものが大きくなることを示している。

また  $u-M < 0$  のとき  $A < 0$  となるから  $E(u, 0) > \frac{1}{2}$  となる。これは、T年において  $u$  を示したものは、(T+1)年 で  $u$  以上の測定値を示す確率が  $1/2$  より大となる、すなわち T年の測定値が集団平均以下であるものは、(T+1)年の測定値が  $u$  より大となる確率は  $1/2$  より大となる。

また  $u-M = 0$  のとき  $A = 0$  となり  $E(u, 0) = \frac{1}{2}$  となる。すなわち T年の測定値が集団平均に等しかったものは (T+1)年の測定値は  $u$  以上を示すものと  $u$  以下を示すものとは等しいということになる。

以上示した計算結果は、前に示した最大血圧の測定値間の関係を如実に示しているものと思えるのである。測定値に誤差があるとき、その悪戯として第1表のような傾向があらわれてくることになる——測定値に誤差がないとき、T年、(T+1)年 でまったく等しい値をもつものであっても。

それでは定量的にどうなるかを計算してみよう。

このため、T年の  $u$  一定のもとでの平均値を計算してみればみやすい。

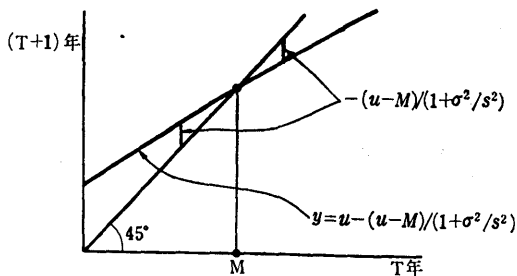
$d$  についての密度関数は

$$B \frac{1}{\sqrt{2\pi}} e^{-\frac{(A+Bd)^2}{2}}$$

となるから  $d$  を乗じて平均を求めれば、 $-(A/B)$  となる。すなわち  $-(u-M)/(1+\sigma^2/s^2)$

となる。これは  $u$  に対するズレの平均である。これが、 $u$  に関して直線となるのは注目される。さて、もちろんこれは、 $u-M > 0$  のとき負であり  $u-M < 0$  のとき正であり  $u-M = 0$  のとき 0 となる。概念図を描けば第5図のようになる。

なお、この分散は  $1/B^2$  によって与えられる。



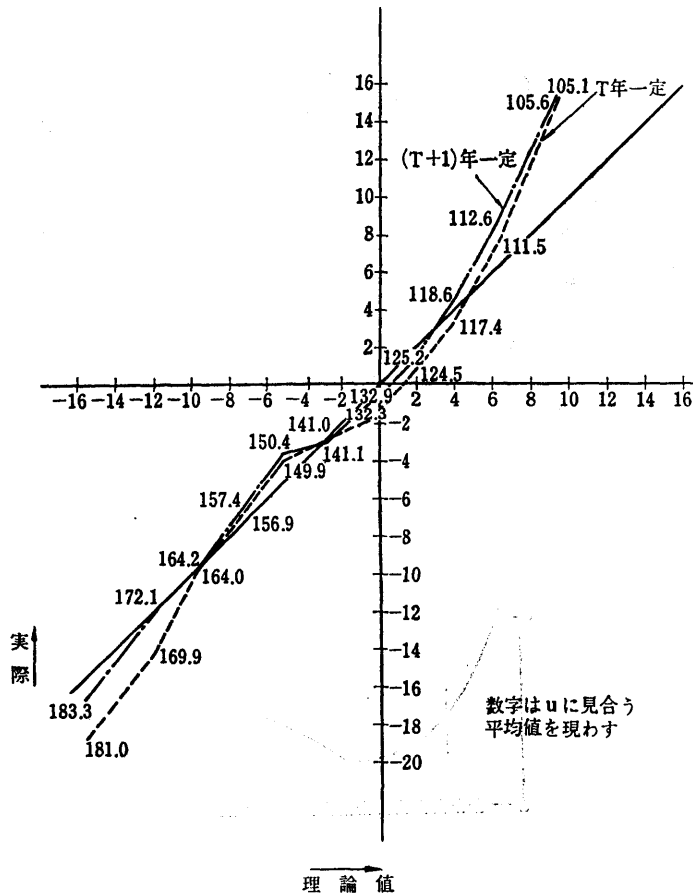
第5図

データと合わせてみることにしよう。前に述べたように  $s^2 = 50$ ,  $\sigma_0^2 = 391$  として計算して



みると、理論値より実測値の方がはるかに大きく、約2倍程度出てくるのである。また、われわれのデータのちらばりをながめてみると、 $s^2=50$ では考えられぬくらいちらばっているのである。 $s^2=50$ が過少と思われるのである。 $s^2=50$ は、前掲の報告のように相当コントロールされた条件下での値であり、一般の健康診断ではここまでコントロールされない（例えば安静時間はまちまち、測定値は2回目の値とは限らぬ、測定者は多数、時期は短期間ではないなど）ので、これらの分散を見積れば一応  $s^2=100$  と考えることも無理ではない（前掲循環協の報告書 [1] 参照）。こうすると  $\sigma_0^2=341$  となる。これを用いて計算しデータとあわせてみると第4図のようになった。横軸に計算値、縦軸に実際のデータをとったので両者が合致していれば  $45^\circ$  の線上にのることになる。

第4図で両端において  $45^\circ$  の線から外れるのは、真の血圧値の分布の型がガウス分布と異なるためである。測定値の分布の曲線は  $x$  の小さいところで急になっている。このことはデータでは  $(u-M)$  の小さいところ、すなわち  $u$  の小さいところで  $d$  の平均は理論値より大きくでるはずである。このことは、第6図ではっきりあらわれてきている。



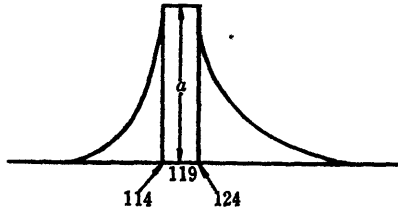
第6図 d に関する理論値と実際のデータ

そこでさらに考えを進め、これまでは個人の血圧（平均値）の集団分布がガウス分布に従うと考えてきたが、これを「ガウス分布」とせず一般の分布とし、これをデータから求め、前に述べた基本式を用い  $s^2=g(x)$  として数値計算によって出してみた。これは、電子計算機を用いれば容易なものである。

さて、データから測定誤差のないときの密度函数  $\Phi(x|M, \sigma^2)$  を求めるには、

$$\int_{-\infty}^{\infty} \Phi(x|M, \sigma^2) \cdot \frac{1}{\sqrt{2\pi s}} e^{-\frac{(x-x)^2}{2s}} dx = F(x),$$

$F(x)$  はデータから求めた分布  $G(x)$  と対応するものであるから、 $\Phi(x|M, \sigma^2)$  を求めればよいので、これは  $\Phi(x|M, \sigma^2)$  を相対度数分布とみなし  $F(x)$  もこれに見合って相対度数分布とすれば一次方程式を解いて  $\Phi(x|M, \sigma^2)$  を求めることができる。しかし、われわれの場合は  $\Phi(x|M, \sigma^2)$  を簡単化し下の形と予想する (第7図)。

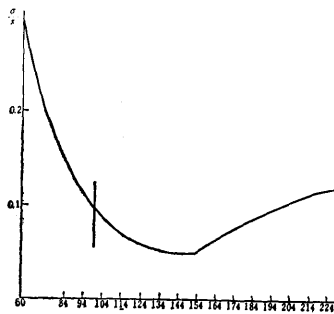


第7図

$$\begin{cases} ae^{b(x-114)} & x \leq 119 \\ a & 119 < x \leq 129 \\ ae^{c(x-124)} & 129 < x \end{cases}$$

ここで当然のことながら面積が1となるようにし、 $F(x)$  の平均が  $G(x)$  のそれと一致するものとし (これにより  $b, c$  を  $a$  の函数として求める) 次に各16の区間——データを16の区間にわける——において  $F(x)$  と  $G(x)$  とがなるべく一致するように (区間にわたった上で最小二乗法によって)  $a$  を求めるのである。

なお、このとき  $s^2=g(x)$  に関してであるが、データとにらみあわせ適切と思われるものを想定した。実測のデータからは、一般に集団平均に近いものの  $s^2$  は小さい (前掲のごとく50とする)、大きな血圧値を示すものの  $s^2$  は大きい、ことがわかっているが、血圧値の非常に小さなものに対しては、それを知るデータがなかった。しかし、われわれのいま取り扱っている



第8図

データからは、 $s^2$  が50よりもずっと大きくなることが予測された。これらを考えあわせて  $s^2$  の形を放物線と直線との混合と考え、血圧値が154以下は放物線——血圧値134で  $s^2=50$  とする——であるとし、それ以上では直線的に増加し174で  $s=0.09 \times 174$  となるものとして  $s^2=g(x)$  を特性化してみた——これは試行錯誤の結果こうするとデータがわれわれの誤差モデルによくあうことが了解された。こうみると、 $s/x$  の形は第8図に示すようになる。通常あら

われる  $x=104$  以上でみるとデータから首肯できる形をしている。詳細には、誤差測定 of データを数多くとり検討しなければならないが、これは今後の問題である。

こうすると  $s^2$  の平均は 73.9572 となり、 $\sigma_0^2=367.9847$  となる。

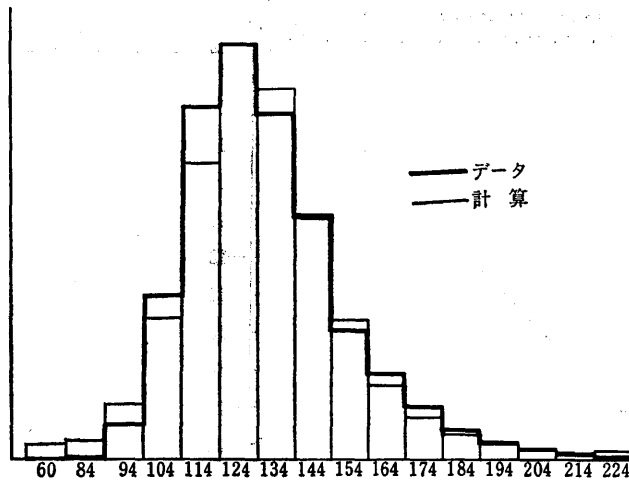
これを利用して  $\Phi(x|M, \sigma^2)$  を計算してみると

$$a=0.0262319$$

$$b=0.1008924$$

$$c=-0.0549150$$

となった。これを用い、データの分布とあわせみると第 9 図のようになり、かなりよい一致である。参考のため理論にもとづくものの標準偏差を出してみると 22.9 となり、データの 21.0 とくらべ少し増大した。



第 9 図

この  $\Phi(x|M, \sigma^2)$  および  $s^2=g(x)$  を用い、

$$E(u, d) = \int_{-\infty}^{\infty} \varphi(u, x) \Phi(x) dx \int_{u+d}^{\infty} \frac{1}{\sqrt{2\pi} S(x)} e^{-\frac{(x-x)^2}{2S(x)^2}} dx / C$$

$$C = \int_{-\infty}^{\infty} \varphi(u, x) \Phi(x) dx, \quad \varphi(u, x) = \frac{1}{\sqrt{2\pi} S(x)} e^{-\frac{(u-x)^2}{2S(x)^2}}$$

の数値計算を実行する。これから  $d$  の平均も計算できる。

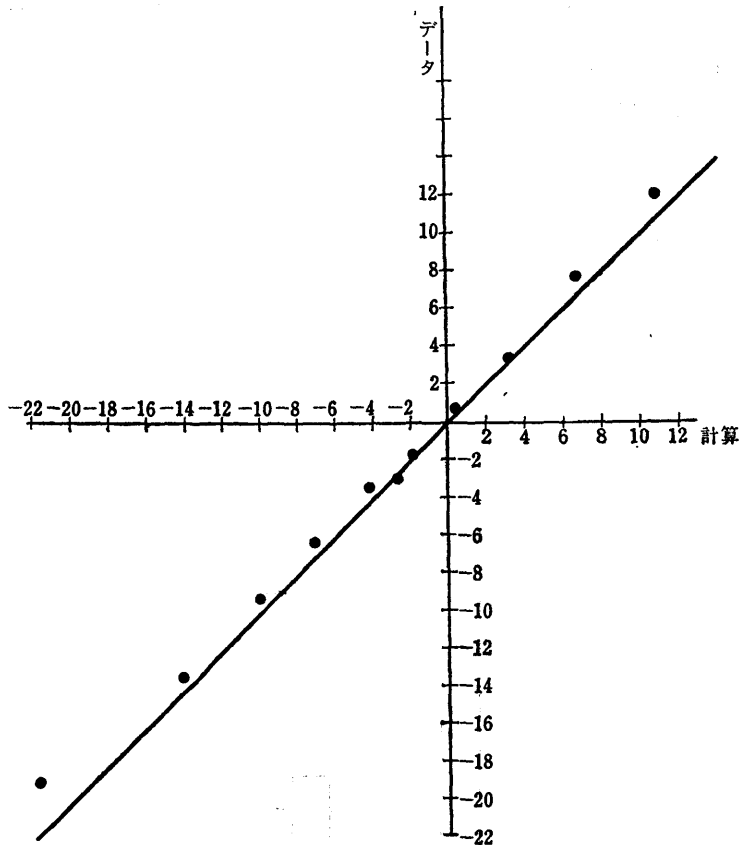
これを実際に行なってみた結果、 $d$  の平均として次の第 4 表を得た。

これをグラフに目もとると第 10 図のようになりきわめてよい一致が示されたことがわかる。

以上の議論からみて測定誤差の悪戯によって思わぬゆがんだ様相が現われてきたことが示された。つまり、測定誤差によってわれわれの得た「データの相関関係」の様相がよく説明されたことがわかった。T 年も (T+1) 年の測定も、もし測定誤差がなければまったく相等しい値を示すものとしても、T 年からみた (T+1) 年の平均

第 4 表

$x$	$d$ の平均	データ
93	10.6	12.1
104	6.6	7.5
114	3.2	3.4
124	0.2	0.4
134	- 2.0	- 1.9
144	- 3.1	- 2.9
154	- 3.7	- 4.1
164	- 6.4	- 7.0
174	- 9.5	-10.0
184	-13.6	-14.1
201	-22.5	-19.0



第 10 図

値は 45° の線に乗らないで、かたよった直線上に乗ることになるわけである。つまり回帰直線は 45° とならないわけである。これが 45° に乗ったとしたら、むしろ T 年と (T+1) 年とはまったく相等しい——測定誤差がないとき——値を示すものではない、と言えることになる。

ここでまた見透しをよくするために  $\Phi(x|M, \sigma^2)$  がガウス分布、 $s^2$  は一定であるとして議論を進めよう。前述のように T 年と (T+1) 年との経過を論ずるにあたっては、45° の直線を中心にするのではなく、偏った回帰直線  $y = u - (u - M) / (1 + \sigma^2 / s^2)$  ( $u$  は T 年における値)、を基にして考えなくてはならないのである。この回帰直線よりさらに上まわって、回帰線が引かれるならば、そこに何か経過変動を見るべきである。また、経過変動をみるとき、如上の回帰直線のまわりの分布

$$\frac{1}{\sqrt{2\pi}(1/B)} e^{-\frac{(d+A/B)^2}{2(1/B)^2}}$$

を用いて、検定を行う必要がある。これらの知識を用い、検定を行い、一定の範囲を超えるものを有意としてその要因をさぐることは有効なことであろうと思われる。

これまで T と (T+1) 年と同一の測定値をもつとしたが、測定誤差がないとき T 年で  $x$  のものは (T+1) 年で  $kx$  ( $k > 0$ ) をもち、また誤差分散が T 年で  $s_T^2$ 、(T+1) 年で  $s_{T+1}^2$  である——ガウス分布とする——としてもまったく同様に計算を行い、回帰直線を出すことができる ( $s^2 = g(x)$  のとき回帰線は一般に直線とはならない)。

これまでの理論計算では  $x$  の分布がガウス分布  $\Phi(x|M, \sigma^2)$  としたが、 $x$  の分布に、山が

ない一様分布であったとしたならば、上述の回帰直線は常に  $45^\circ$  の直線になることは注意しておく必要がある。

さて、以上のところでは、1回目と2回目との値が等しいとしたがこんどは、1回目に  $x$  であったものは2回目には  $kx$  ( $k > 0$ ) の値を示すものとし、誤差分布が1回目のとき、

$$\frac{1}{\sqrt{2\pi} s_1} e^{-\frac{(v-x)^2}{2s_1^2}} dv$$

とし2回目のとき

$$\frac{1}{\sqrt{2\pi} s_2} e^{-\frac{(w-kx)^2}{2s_2^2}} dw$$

であるとしてみる。  $x$  の分布は前と同様  $\Phi(x|M, \sigma^2)$  とすると  $kx$  の分布は  $\Phi(kx|kM, k^2\sigma^2)$  となる。

このとき1回目で  $u$  を示したものが2回目で  $ku+d$  以上を示す確率は、

$$\frac{1}{\sqrt{2\pi}} \int_{A'+B'd}^{\infty} e^{-\frac{t^2}{2}} dt$$

となる。

ここに

$$A' = (u-M)k \left( \frac{s_1}{s_2} \right) s_1 \frac{1}{\sqrt{(s_1^2 + \sigma^2) \left\{ s_1^2 + \sigma^2 \left( 1 + \frac{s_1^2}{s_2^2} k^2 \right) \right\}}}$$

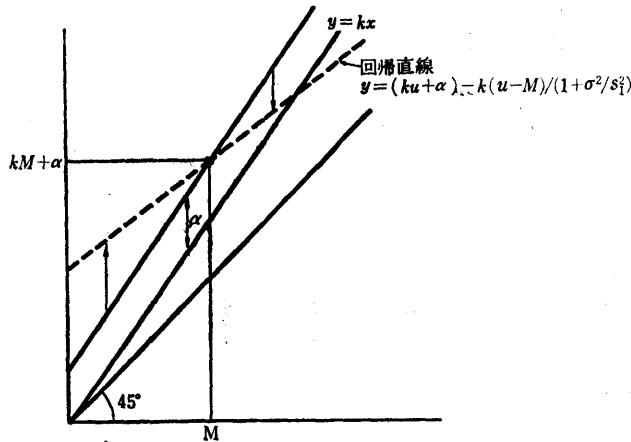
$$B' = \left( 1 + \frac{\sigma^2}{s_1^2} \right) \left( \frac{s_1}{s_2} \right) s_1 \frac{1}{\sqrt{(s_1^2 + \sigma^2) \left\{ s_1^2 + \sigma^2 \left( 1 + \frac{s_1^2}{s_2^2} k^2 \right) \right\}}}$$

したがって  $ku$  に対するズレ  $d$  の平均は

$$-A'/B' = -\frac{k(u-M)}{1 + \sigma^2/s_1^2}$$

で、やはり、 $u$  に関して直線となる。  $s_1^2$  にのみより  $s_2^2$  に依存しないことは注目してよい。また分散は  $1/B'^2$  となる。

また  $x$  が次のとき  $kx$  となるのではなく  $kx+\alpha$  となるときは、ただ  $45^\circ$  の直線に  $\alpha$  だけのゲタをはかせたうえで、前述の議論をそのまま利用すればよいことが容易に示すことができる。したがって、 $x$  であったものが次のとき  $kx+\alpha$  となった場合には、両者の議論をあわせ傾きをかえ、 $\alpha$  のゲタをはかせて ' $kx$ ' のときの議論を用いればよい。



第 11 図  
 $k > 1, \alpha > 0$  のときの図

2. 測定したものが分類で与えられている場合

こんどは、測定したものが、分類の形で与えられているときを考える。例えば、尿検査などで測定結果が、+、±、-等で与えられている場合である。このときの測定誤差は第5表のよう

第5表

見かけ (測定) 本当	+	±	-	計
	$P_{++}$	$P_{+±}$	$P_{+-}$	1
+	$P_{++}$	$P_{+±}$	$P_{+-}$	1
±	$P_{±+}$	$P_{±±}$	$P_{±-}$	1
-	$P_{-+}$	$P_{-±}$	$P_{--}$	1

確率の表で与えられるものとしよう。これは、本来(本当)のところ+のものが、測定で正しい値+を示す確率が $P_{++}$ 、あやまった値±となってあらわれる確率を $P_{+±}$ 、-となつて大いにあやまったものとして測定される確率を $P_{+-}$ とする。これが測定誤差の表現である。 $P_{++}=1, P_{+±}=0, P_{+-}=0, P_{±+}=0,$

$P_{±±}=1, P_{±-}=0, P_{-+}=0, P_{-±}=0, P_{--}=1,$ であれば、つまり第6表のとおりであれば、測定誤差がないことになる。

このように測定誤差があるとき、結論で誤りを犯すことになる。これからは誤差が確率的に独立におこるものとして話を進めてみよう。

第6表

見かけ (測定) 本当	+	±	-
	1	0	0
+	1	0	0
±	0	1	0
-	0	0	1

第7表

	見かけ (測定) 本当	+	-
		0.7	0.3
200	+	0.7	0.3
800	-	0.2	0.8
		300	700

簡単のため、+、-の2つとして数値例で説明しよう。本当は+のものは、測定誤差があるため本当の値+を示す確率は0.7、-とあやまった値を示す確率が0.3としておく(第7表)。本当は-であるものは測定誤差により-と正しい値を示す確率は0.8、+とあやまった測定を出す確率を0.2としておこう。いまかりに、本当は+である人数は200人、本当は-である人数は800人(それぞれ20%、80%)であるとするが、これは未知なのである。測定結果をまとめてみると平均的な意味で+とあらわれる人数は

$$200 \times 0.7 + 800 \times 0.2 = 140 + 160 = 300 \text{ 人}$$

$$200 \times 0.3 + 800 \times 0.8 = 60 + 640 = 700 \text{ 人}$$

となる。この300人、700人が観測された値となる。観測された値が300人、700人であるから、+が30%、-が70%と結論しがちであるがこれはあやまりである。本当は+が20%-が80%なのである。これはいわゆるサンプリングによる誤差ではない。測定が確率的におこるための誤差なのである。このようなとき、測定誤差の確率がわかっていたら観測から正しい値を求めることができる。本当の+の人数を $n_+$ 、-の人数を $n_-$ とすれば、みかけの+、-の人数、すなわち観測によって得られる人数は平均的に、

$$\left. \begin{aligned} 0.7n_+ + 0.2n_- &= 300 \\ 0.3n_+ + 0.8n_- &= 700 \end{aligned} \right\}$$

となるから、この連立方程式をとけば、

$$n_+ = 200$$

$$n_- = 800$$

と正しい値を得ることができる。これは一般化して表現し

$$\begin{pmatrix} 0.7 & 0.3 \\ 0.2 & 0.8 \end{pmatrix}' \begin{pmatrix} n_+ \\ n_- \end{pmatrix} = \begin{pmatrix} 300 \\ 700 \end{pmatrix}$$

但し  $\begin{pmatrix} \quad \\ \quad \end{pmatrix}'$  は転置行列  $\begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix}$  をあらわす、

となることから

$$\begin{pmatrix} n_+ \\ n_- \end{pmatrix} = \begin{pmatrix} 0.7 & 0.3 \\ 0.2 & 0.8 \end{pmatrix}'^{-1} \begin{pmatrix} 300 \\ 700 \end{pmatrix}$$

として求められる。

以上は、平均的な関係式をつかって求めたのである。しかし必ずしも平均値が出てくるわけではないので、そのときはどうするかとなる。このときは上の関係式を形式的につかって求めたものが、 $n_+$ 、 $n_-$  の偏りのない推定値となることが証明でき、その分散も容易に計算できる。煩にわたるので、一応これを割愛し、実質的な議論を進めよう（詳細は [3] 参照）。

上述のことは、分類がいくつあっても同じであり、尿検査の3つの場合は、本当の値  $n_+$ 、 $n_{\pm}$ 、 $n_-$  の推定は、測定で得られた値  $m_+$ 、 $m_{\pm}$ 、 $m_-$  から

$$\begin{pmatrix} \hat{n}_+ \\ \hat{n}_{\pm} \\ \hat{n}_- \end{pmatrix} = \begin{pmatrix} P_{++} & P_{+\pm} & P_{+-} \\ P_{\pm+} & P_{\pm\pm} & P_{\pm-} \\ P_{-+} & P_{-\pm} & P_{--} \end{pmatrix}'^{-1} \begin{pmatrix} m_+ \\ m_{\pm} \\ m_- \end{pmatrix}$$

によって求められる。 $\hat{n}_+$ 、 $\hat{n}_{\pm}$ 、 $\hat{n}_-$  はそれぞれ  $n_+$ 、 $n_{\pm}$ 、 $n_-$  の偏りのない推定値である。この推定値の分散も容易に計算できる（文献 [3] 参照）。但し逆行列は存在するものとする。

ここまでのところは、ちょっと注意すれば解るのであるが、こうした誤差がある場合、1. でのべたように、2回の測定を組み合わせたとき、非常に誤った結論にみちびかれる可能性がある。数値例で示そう。

$P_{++}$ 、 $\dots$ 、 $P_{--}$  の表として次のようなものを考えてみよう（第8表）。

本当の +、 $\pm$ 、- の数は 100、200、700 としておこう。こうした測定誤差があるとき、2回の測定結果のクロス表は平均的に第9表のようになる。こうなる理由を少し説明してみよう。

1回目の測定で + を示すものは

$$100 \times 0.7 + 200 \times 0.2 + 700 \times 0.1$$

$$= 70 + 40 + 70 = 180 \text{ 人}$$

である。100人が180人と出るところからしてあやまっている。このうち2回目の測定で + を示すものは平均的に

$$70 \times 0.7 + 40 \times 0.2 + 70 \times 0.1 = 64 \text{ 人}$$

である。1回目 + で2回目  $\pm$  を示すものは、

$$70 \times 0.2 + 40 \times 0.6 + 70 \times 0.1 = 45 \text{ 人}$$

となる。- を示すものは、

$$70 \times 0.1 + 40 \times 0.2 + 70 \times 0.8 = 71 \text{ 人}$$

となる。こうして第9表が出来上がる。

測定誤差が1回目、2回目とも同様であるならば周辺分布は同一となっている。ここに注意すべきは、周辺分布が真の分布 0.10、0.20、0.70 と大きく異なっていることである。もし、測定誤差がなければ、2回測定をやったとき第10表のようになるべきである。それが測定誤差のため第9表のようにばらついた表が得られる。ばらついているだけならば、それまでであるが第9表をもとにし、測定誤差がないものとして、実体的な結論を引き出そうとすれば大き

第8表

		見かけ		
		+	$\pm$	-
本当	+	0.7	0.2	0.1
	$\pm$	0.2	0.6	0.2
	-	0.1	0.1	0.8

第9表

	2回目				
1回目		+	±	-	計
+		64	45	71	180
±		45	83	82	210
-		71	82	457	610
計		180	210	610	1000

第10表

	2回目			
1回目		+	±	-
+		100	0	0
±		0	200	0
-		0	0	700

な誤りをおかす。1回目に+であったものが180人である。2回目の測定で+にとどまるものは64人で約1/3にしかすぎない。そこで+であったものは±或は-へ、とくに-へ変化したものが多くと結論できる。ここに、測定誤差を無視して、いわゆる検定論を用いたところで結論は同様である。ところがこれは、測定誤差の悪戯にすぎないのである。1回目±のところも2回目の測定で±にとどまるものは(+, -)になるものにくらべ少なくなっているのである。

こうしたことも、測定誤差のことを心得ておけば、妥当性ある結論をみちびくことができるが、これを無視して分析していたとすれば、非常にあやまりを犯すことになる。

妥当性を得る分析ができるためには誤差を示す確率の表  $(P_{ij})$ ,  $i, j = +, \pm, -$  を知っておかなくてはならない。知るためには、同一人の2回以上の測定値をもとにして求めることを考えなくてはならない。このためには、基礎研究を行なって、その結果から計算して出してこなくてはならない。基礎研究を通して、妥当性ある結論を導くことができるのである。この具体的方法については、文献 [3] を参照されたい。

平均的にみれば、1回目、2回目のクロス表の結果を  $m_{++}, m_{+\pm}, m_{+-}, m_{\pm+}, m_{\pm\pm}, m_{\pm-}, m_{-+}, m_{-\pm}, m_{--}$  とし、本来+のものを  $n_+$ , ±のものを  $n_{\pm}$ , -のものを  $n_-$  としておこう。

もちろんこのモデルに従えば、 $m_{+\pm} = m_{\pm+}, m_{+-} = m_{-+}, m_{-\pm} = m_{\pm-}$  となるべき筈のものである。

そうすると計算によって平均的關係として

$$R \begin{pmatrix} n_+ \\ 0 \\ 0 \\ 0 \\ n_{\pm} \\ 0 \\ 0 \\ 0 \\ n_- \end{pmatrix} = \begin{pmatrix} m_{++} \\ m_{+\pm} \\ m_{+-} \\ m_{\pm+} \\ m_{\pm\pm} \\ m_{\pm-} \\ m_{-+} \\ m_{-\pm} \\ m_{--} \end{pmatrix}$$

が得られる。ここに  $R$  は  $9 \times 9$  マトリックスで

$$R = \left\{ \begin{pmatrix} P & 0 & 0 \\ 0 & P & 0 \\ 0 & 0 & P \end{pmatrix} \begin{pmatrix} P_{++}I & P_{+\pm}I & P_{+-}I \\ P_{\pm+}I & P_{\pm\pm}I & P_{\pm-}I \\ P_{-+}I & P_{-\pm}I & P_{--}I \end{pmatrix} \right\}$$

ただし

$$P = \begin{pmatrix} P_{++} & P_{+\pm} & P_{+-} \\ P_{\pm+} & P_{\pm\pm} & P_{\pm-} \\ P_{-+} & P_{-\pm} & P_{--} \end{pmatrix}, \quad 0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$



をあらわし  $\{ \}$ ' は転置行列である。

こうした関係式を把んでおくと、 $(P)$  及び  $n_i$  ( $i=+, \pm, -$ ) を知ることによって、2回の測定結果のクロス表が測定誤差によってどう現れるかを上述の平均値関係モデルに従って求めることができる。これと実際のデータとつきあわせ、 $m_{ij}$  ( $i, j=+, \pm, -$ ) の確率的な分布——これを知るとは全く容易である——と比較することによって、データの  $m_{ij}$  が確率的変動内にあるか否かをみることができる。これを外れていたとすれば、測定誤差以外に何か実質的な要因がはたらいたとみなせばよいことになる。このときも  $(P)$  をあらかじめ知っておくこと、 $n_i$  を求めておくことは欠くべからざることである。 $n_i$  の推定は、第1回、第2回の調査結果の周辺分布から直ちに推定することができる。

以上のことは別にして、 $n_+$ 、 $n_{\pm}$ 、 $n_-$  の推定は、上述のモデルの妥当性が確かめられていれば、上の関係を利用しても求められる(文献[3]参照)。

以上は考え方の一つであるが、データを通して  $P$  はどのくらいのものかみてみよう。

肺結核をみるためのレントゲン写真の判定についての例を示して論じてみよう。

まず空洞の有無 (*WHO* の分類をまとめたもの) の判定データをみよう。脚注論文の第10表である。協議決定解とあるのを一応正しい値とみなしておこう。そうすると上記第10表は不明を除くと第11表のようになる。これは多くの人の判定の合計であるが一応、平均的な意味で  $(P_{ij})$  をあらわしているとみなしてみることにしよう。こうした場合、前の理屈にしたがえば、何回かの測定値をつき合わせ、クロス集計した場合、そのバラツキを計算することができる。480人のケースについて、空洞有無について、1回、2回の判定をしたものがあげられている(前掲論文)。

第11表

見かけ		正		
		有	疑	無
正	空洞有	0.84	0.12	0.04
	空洞疑	0.21	0.55	0.24
	空洞なし	0.04	0.09	0.87

これは第13表のようになる。周辺分布は第12表のようになり、さして差はみとめられない。

第12表

	有	疑	無	計
第1回	17.1	27.1	55.8	100
第2回	18.8	30.2	51.0	100

第13表

2回		1回			計	%
		有	疑	無		
有	有	63	12	7	82	17.1
	疑	6	120	4	130	27.1
	無	21	13	234	268	55.8
計		90	145	245	480	
%		18.8	30.2	51.0		

そこで、この表から  $n_+$ 、 $n_{\pm}$ 、 $n_-$  を推定する。そのため周辺分布の平均をとれば、

\* 労働結核研究協議会病型研究班による肺結核の病型分類について。

$$m_+ = (90 + 82)/2 \doteq 86$$

$$m_{\pm} = (130 + 145)/2 \doteq 137$$

$$m_- = (268 + 245)/2 \doteq 257$$

として,  $n_+$ ,  $n_{\pm}$ ,  $n_-$  を推定する.

$$\begin{pmatrix} \hat{n}_+ \\ \hat{n}_{\pm} \\ \hat{n}_- \end{pmatrix} = (P)^{-1} \begin{pmatrix} 86 \\ 137 \\ 257 \end{pmatrix}$$

その結果  $\hat{n}_+ = 41$ ,  $\hat{n}_{\pm} = 201$ ,  $\hat{n}_- = 238$  となる.

これらを用い, 前の理屈によって, 平均的クロス表を計算してみよう. そうすると, 2回の測定値をつき合わせてクロス集計した場合, そのクロス表の数字のちらばり度合をみるのが

第 14 表

1回 \ 2回	有	疑	無	計
有	38	28	20	86
疑	28	64	45	137
無	20	45	192	257
計	86	137	257	480

できる. これによると第 14 表が得られることになる.

これを観察すると, 第 11 表ぐらいの測定誤差——かなり小であると思われる——があったとき, 相当ばらついた表が得られ, 誤差の影響を無視することはいささか危険なことがわかる. このデータは非常に注意深く行われたものであるが, 一般の健康管理のときはさらに大きな測定誤差が予想されるので, 時を変えた 2 回の測定をもとにして, 軽快した, 重くなった, というときにも如上の配慮を怠ってはならない.

さて, 実際に, 得られている 2 回測定のデータは第 13 表のとおりで, 計算から得られた表は第 14 表のとおりで結果は甚しく異なったものになっている. 対角線上にあるもの, つまり 2 回測定の一一致率は, 計算の方で 61%, 実際の方は 87% と実際の方がはるかに高い. 確率計算より現実の方の 2 回判定の一一致率が高いことになる. このことは, 個人内では 2 回の判定にそれほどくい違いはないが, 個人間のくい違いが多いという事実を裏書きしていることになる. こうした問題では前の協議解と各人の判定の不一致から計算した確率がすべて各人に同一のものとして, 確率的に判定されるものでないことを示していることになる. つまり, 個人内の変動はそれほど大ではないが個人間の判定のくい違いはそれに比して大であることが知られる. こうした問題では, 各個人の確率判定のほか, 判定の偏り (判定確率マトリックスが人によって異なる) を考慮に入れて分析しなければならないことになる. 次の例に進もう.

ツベルクリン反応の変動——測定のしかた, 個人変動, 測定者の判定誤差——というものを考えるとき, しかも, その変動の大きいことが予想されるので 2 回の反応をつき合わせて, 陽転, 陰転に関して論じようとするとき, ここで述べたような推論の方法を無視しては, 大きな誤りを犯すことになる.

ちなみにツベルクリン反応におけるデータ\*のうち 50~54 歳のものをみよう (第 15 表).

33 年と 34 年とでは, さすがに 50~54 才であるため周辺分布はそう異なっていない. つまり, この間にツベルクリンの反応分布が全体としてそう著しく異なっていないとみられるのである. 対角線に関してその上下方すなわち  $m_{ij}$  と  $m_{ji}$  の数字を比較するとかなり似通ってい

\* 厚生省結核実態調査より. 男女計 BCG 接種なし,

第 15 表

33年	34年					計	%
	—	±	+	++	+++		
—	32	24	21	4	0	81	10.4
±	21	32	24	12	1	90	11.6
+	15	20	97	57	12	201	25.8
++	2	17	70	126	53	268	34.5
+++	1	1	9	45	82	138	17.7
計	71	94	221	244	148	778	100
%	9.1	12.1	28.4	31.4	19.0	100	

る。しかも、33年—のものの陽転率は30%と高く、±を含めると—に止まるより多く60%程度になっている。34年からみても同様に34年に—であったものは33年では(±、+以上)のものが多かったということになる。以上のような傾向からみて、これまで述べてきたような測定誤差(変動)による悪戯が主要な要因となって陽転などに大きな値が出てきているのではないかと考えられるのである。この陽転陰転(33年からみる、或は34年からみる)の傾向に実質的な要因を見出すより前に、変動にもとづく要因からの分析をほどこして見る必要がある。

われわれはツベルクリンの測定変動(誤差)のデータがないため、如上の傾向をデータとして示すことはできないが、これまで述べてきたわれわれの考え方によって、相当のところまで裏づけられるように思われるのである。

本文の数値計算に関しては日本ビジネスコンサルタント(NBC)科学計算課の本間氏、宮下氏にお世話になった。厚く感謝の意を表するものである。

#### 参 考 文 献

- [1] 循環器管理研究協議会：報告第1集，循環器管理研究協議会，1966.
- [2] 田中久恵，山口明子，秋山房雄，林 知己夫：血圧値・心電図・尿所見の変動について——健康管理の場における——公衆衛生学雑誌，第14巻，第2号，1967.
- [3] 林 知己夫：社会調査における回答誤差——その歪みをどう補正するか——20周年記念論文集，NHK放送文化研究所・1967.