創立第13週年記念講演会次第

昭和32年11月9日午後1時30分より統計数理研究所講堂において行われた.

1. 挨　拶　　所長事務取扱　松下嘉米男
2. マス・コミュニケイションの効果測定
——新聞記事内容の分析と一般の意見との関係——　　第2研究部長　林　知己夫
3. 大型自動計算機の現況とその発達　　第2研究部 第2研究室長　多賀保志

創立第14週年記念講演会次第

昭和33年6月7日午後1時30分より統計数理研究所講堂において行われた.

1. 挨　拶　　所　長　末綱恕一
2. 調査の信頼性　　第3研究部長　青山博次郎
3. 選挙予想　　第2研究部長　林　知己夫
4. インドの統計事情　　第1研究部長　松下嘉米男

講演後大型リレー計算機の公開を行った.

昭和32年度研究発表会アブストラクト

昭和33年3月26日，27日の両日統計数理研究所講堂において，32年度研究発表会を行った.

ある種のモーメント問題について

石井恵一

一次元 unimodal 分布に対するモーメント問題の criterion を導くのに分布の Stieltjes 変換を用いる方法は，研究所の Annals Vol. IX (1958) に述べたが，二次元の分布に対しても同様な方法が適用される．まず二次元の確率分布に対して Stieltjes 変換に相当するものを定義し，更に分布が unimodal である為の条件がこの変換の性質として表現されることを導く.

$d\phi(t_1, t_2)$ を $S=[a,\infty)\times[b,\infty)$ における確率要素とすれば（R^2 全体にわたる場合はその極限として考えればよい），複素数 z_1, z_2 に対し

$$* \qquad f(z_1, z_2)=\int_S \frac{d\phi(t_1,t_2)}{(z_1-t_1)(z_2-t_2)}$$

なる変換は明らかに次の定理の条件 ①—④ を満足するがその逆も成り立つ．すなわち

［定理 1］ z_1, z_2 の函数 $f(z_1, z_2)$ が * の形に表わされるための必要十分条件は次の ①—④ が満足されることである：

① $f(z_1,z_2)$ は K^2-S で z_1, z_2 に関して解析的である（K は複素数空間).

② $f(\overline{z_1}, \overline{z_2})=\overline{f(z_1, z_2)}$

③ $\Im z_1>0$, $\Im z_2>0$ なる領域では

$\Re[f(z_1,z_2)-f(\overline{z_1}, z_2)]\leqq 0$

④ $f(z_1,z_2)=\dfrac{1}{z_1z_2}+\dfrac{1}{z_1}W_2(z_2)+\dfrac{1}{z_2}W_1(z_1)+W(z_1,z_2)$.

ここに，任意の $\varepsilon>0$ に対し，扇形領域 $\varepsilon\leqq arg\, z_j\leqq \pi-\varepsilon$ で $z_j\to\infty$ のとき $W_j(z_j)=o\left(\dfrac{1}{z_j}\right)$，又同じ扇形領域で $z_1z_2\to\infty$ のとき $W(z_1,z_2)=o\left(\dfrac{1}{z_1z_1}\right)$.

次に二次元の unimodal 分布は次のように定義される．$E\subset R^2$ に対し，$pE=\{(pt_1, pt_2)\,;\,(t_1,t_2)\epsilon E\}$ と書けば，任意の E と任意の $p>1$ に対して

$$\phi(pE)\leqq p^2\phi(E)$$

であるとき，ϕ を，原点をモードとする unimodal 分布という．一般の点をモードとするものの定義は，これを平行移動させればよい．すると次の基本定理が成り立つ.

［定理 2］ ϕ が (α,β) をモードとする unimodal 分布であるための必要十分条件は，前記の如き $f(z_1, z_2)$ を作ると

$$g(z_1,z_2)=-\frac{1}{2}\left\{(z_1-\alpha)\frac{\partial f}{\partial z_1}+(z_2-\beta)\frac{\partial f}{\partial z_2}\right\}$$

が定理 1 に於ける ①—④ を満足することである.

この定理の意味は，(α,β) を mode とする unimodal 分布と，一般の分布との間に一対一の対応がつくことであり，これによって unimodal 分布に対するモーメント問題を一般のそれに帰着せしめることができる．実際，定理 2 の f, g に対応する分布を夫々 ψ, ϕ とし，それらのモーメントを夫々 $\{\mu_{j,k}\}$, $\{\nu_{j,k}\}$ とすれば，$\mu_{j,k}$ と $\nu_{j,k}$ の一方が存在すれば他方も存在し，且つ次の関係があることが，定理 2 に於ける式の両辺を展開することによって容易に証明される：

$$\nu_{j,k}=\frac{1}{2}\{(j+k+2)\mu_{j,k}-j\alpha\mu_{j-1,k}-k\beta\mu_{j,k-1}\}$$

一般のモーメント問題の条件にこの関係を代入することにより，$\{\mu_{j,k}\}$ に関する条件が得られるわけである.

カイ二乗分布の modal interval

渋 谷 政 昭

$f(x)$, $-\infty<x<\infty$ を連続，単峰の $p.d.f.$ とするとき

$$\int_L^U f(x)dx=1-\alpha,\quad f(U)-f(L)=0$$

という条件を満たす区間 $[L, U]$ を $f(x)$ にたいする $(1-\alpha)$-content modal interval とよぶ．ここではカイ二乗分布にたいする modal interval を求めた．これはガンマ型分布の scale parameter，正規分布の分散についての

i) 不偏最強力検定
ii) 相対精度限定点推定
iii) 平均最短区間推定

に用いられる．詳細は Ann. Inst. Stat. Math. Vol. 9, No. 3, に発表．

統計量の極限分布への拡散方程式の応用

本 尾 実

［定理］ $Y_1, Y_2, \cdots Y_n\cdots$ random variable

$$r_1\leq Y_i\leq r_2 \quad -\infty\leq r_1\leq r_2\leq\infty$$

(1) かつ $E(Y_{n+1}-Y_n|Y_n)=\varepsilon_n b(Y_n)+R(n)$

$$E((Y_{n+1}-Y_n)^2|Y_n)=\varepsilon 2a(Yn)+S(n)$$

$$E(|Y_{n+1}-Y_n|^3|Y_n)=T(n)$$

ただし $\varepsilon_n\downarrow 0$ $\sum\varepsilon_n=\infty$ $E(|R_n|)$, $E(|S_n|)$, $E(|T_n|)$ $=o(\varepsilon_n)$

(2) $a(x)$, $b(x)$ は (r_1, r_2) で連続かつ (r_1r_2) 内で $a(x)>0$ とする．今若し $(0\epsilon(r_1, r_2)$ として一般性を失わない)

$$S(x)=\int_0^x\exp\left(-\int_0^u\frac{b(z)}{a(z)}dz\right)du$$

$$m(x)=\int_0^x\frac{1}{a(z)}\exp\left(+\int_0^u\frac{b(z)}{a(z)}dz\right)du \quad \text{とおいたとき}$$

(3) $S(r_1)=-\infty$, $S(r_2)=\infty$; $m(r_1)>-\infty$ $m(r_2)<\infty$ であると．

$$\lim_{n\to\infty}P(Y_n\epsilon E)\to\frac{m(E)}{m(r_1r_2)} \quad \text{(法則収斂) が成立する．}$$

$$\text{ただし } m(E)=\int_E m(dx) \quad E\subset(r_1r_2)$$

証明は Khintchin [1] の方法で Y_n を $\frac{\partial}{\partial t}=a\frac{\partial^2}{\partial x^2}+b\frac{\partial}{\partial x}$ に対応する拡散過程を近似するのである．(1)の条件はもっと弱めることができる．

応用 1． 中心極限定理

$X_1, X_2, \cdots X_n\cdots$ 同一分布を持ち $E(|X_i|^3)<\infty$ な確率変数とすると

$$Y_n=\frac{1}{\sqrt{n}}(X_1+\cdots+X_n) \quad \text{とおけば}$$

$$b(x)=-x,\quad a(x)=\sigma^2 \quad \varepsilon_n=\frac{1}{\sqrt{n}} \quad \text{となり} \quad (E(X_i^2)=\sigma^2)$$

$$m(dx)=C\exp\left(-\frac{x^2}{2\sigma^2}\right)dx$$

応用 2． Pearson の χ^2 分布

状態の数を k ケとし

$$Y_n=n\sum\frac{\left(\frac{n_i}{n}-p_i\right)^2}{p_i}$$

$$\varepsilon_n=\frac{1}{n} \quad b(x)=-x-k-1 \quad a(x)=2x \quad (x\geq 0)$$

$$m(dx)=Cx^{\frac{k-3}{2}}\exp\left(-\frac{x}{2}\right)dx$$

応用 3． Stochastic Approximation

$Y(x)$ x の水準での実験結果を示す確率変数

$$M(x)=E(Y(x)) \quad M(0)=0$$

簡単のため $|Y(x)|<C$

$$M(x)=\alpha x+o(x) \quad (x\to 0),\ (\alpha>o),$$

$E(Y(x))=\sigma^2+o(x)$ とする

$X_1=x_1$

i) $X_{n+1}=X_n-\frac{C}{n}Y(X_n)$ とおくと

$Y_n=n^{\frac{1}{2}}X_n$ とおけば

$$\varepsilon_n=\frac{1}{n} \quad b(x)=\left(\frac{1}{2}-C\alpha\right)x$$

$$a(x)=\frac{C^2\sigma^2}{2}$$

$2C\alpha>1$ の時は

$$m(dx)=C'\exp\left(-\frac{(2C\alpha-1)}{2C^2\sigma^2}x^2\right)dx$$

ii) $X_{n+1}=X_n-\frac{1}{n^{\frac{\beta}{2}}}Y(X_n)$ とおけば $(0<\beta<1)$

$Y_n=n^{\frac{\beta}{2}}X_n$ とおき

$$\varepsilon_n=\frac{1}{n^{\frac{\beta}{2}}} \quad b(x)=-\alpha x,\quad a(x)=\sigma^2 \quad \text{となり}$$

$$m(dx)=C'\exp\left(-\frac{\alpha}{\sigma^2}x^2\right)dx \quad \text{が得られる．}$$

これは K.L. Chung [2] の結果である．

[1] A. Khintchine : Asymptotische Gesetze der Wahrscheinlichkeits=rechnung. (1933)

[2] K.L. Chung : On a stochastic approximation method, Annals of Mathematics Vol. 25 (1954)

ある two sample test について

藤　本　　煕

x 及 y をそれぞれ連続な分布函数 $F(x)$ 及 $G(y)$ をもつ独立な確率変数とするとき.

$$p=\mathrm{Prob}\{y<x\}=\int_{-\infty}^{\infty}G(t)dF(t)$$

において, 上式の右辺の $G(t)$ 及 $F(t)$ を empirical な $G_n(t)$, $F_m(t)$ で置換えた

$$\hat{p}=\int_{-\infty}^{\infty}G_n(t)dF_m(t)$$

ただし n 及 m は y 及 x の sample size を示す添字である――は, 丁度 Mann-Whitney の U-test の統計量と本質的には同じものであるが, この分散の推定値を求めるには $n=m$ の場合, Lehmann の two sample test で用いた統計量を用いるのが便利であること等, その他これに結付いた性質について述べた. また

$$p-\frac{1}{2}=\int_{-\infty}^{\infty}(G(t)-F(t)d\frac{G(t)+F(t)}{2}$$

と記せるが, これを empirical なものでおきかえて見ると

$$\int_{-\infty}^{\infty}(\hat{G}_n-\hat{F}_m)d\frac{\hat{G}_n(t)-\hat{F}_m(x)}{2}=\hat{p}-\frac{1}{2}+\frac{m-n}{4mn}$$

となって $n=m$ 以外の場合には $\frac{m-n}{4mn}$ だけ余分の項が現れるが, これを用いて, $G>F$ に対する $G\equiv F$ の判定の仕方を考察した.

$G(y)=N(y\,;m_1, \sigma_1^2)$, $F(x)=N(x\,;m_2, \sigma_2^2)$ の Gauss 分布については,

$$p=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{m_2-m_1}{\sqrt{\sigma_1^2+\sigma_2^2}}}e^{-\frac{t^2}{2}}dt$$

で, これは $G(x_0)+F(x_0)=1$ なる x_0 を以て discriminant point とする分類の問題で, その判別の正しい確率

$$C=\frac{1}{\sqrt{2\pi}}\int^{\frac{m_2-m_1}{\sigma_1+\sigma_2}}e^{-\frac{t^2}{2}}dt$$

と類似性をもつのは面白い. この場合の p の分散も容易に求まる.

再生的な過程と tandem type の queue について

赤　池　弘　次

自動車の流れの観測結果をもとにし, その統計的構造の表現に関係する再生的な過程を考察した.

1. 間隔過程. 前年度の研究発表会においてすでに発表した模型であるが, 農林省蚕糸試験場の嶋崎技官等とその定粒繰糸工程への適用を発展させた (彙報5巻2号参照).
2. tandem type の queue. 自動車の流れが単純なポアッソン過程からはなれるのを, 車間距離の制限によるものとしてひとつの模型を考えた. これは tandem type の queue の変型と見做される. これについてその安定性の条件を論じた. (Annals of the Institute of Statistical Mathematics Vol. IX 参照) なおこれに関連して通常の tandem type の queue における窓口の配列の順番がどのような影響をもつかを FACOM-128 によりモンテカルロ法によって検討した. この結果によれば input の間隔が平均 $E(t)$ 分散 $D^2(t)$ なる負の指数型分布に従う場合の output の間隔の時間平均の意味での分散は窓口のサービスが平均 $E(s)$ 分散 $D^2(s)$ なる場合 $E(s)/E(t)=\rho$ とすると ($\rho<1$ とする)

$$D^2(s)+(1-\rho)D^2(t)+\rho(1-\rho)E^2(t)$$

となるが, これが更に次々の窓口を通って行く場合にも十分良い近似を与えることが認められた.

通話路の伝送容量と伝送速度について

高　野　金　作

$[A_0, \mu]$ をエルゴード情報源とし, その1文字当りのエントロピーを H とする. $[A, \nu_x, B]$ を有限記憶 m をもち m_1 歩従属な定常通話路とし, そのエルゴード伝送容量を C とする. r, s を自然数とし $\varphi_{r,s}$ を $A_0{}^r$ から A^s への一意写像とし, $A_0{}^I$ から A^I への符号化はこの $\varphi_{r,s}$ を用いて r 項毎に行うものとする. ここに $I=\{1, 2, 3, \cdots\}$. 情報源から出る長さ r の通信文 $\alpha=(\theta_1, \theta_2, \cdots, \theta_r)\epsilon A_0{}^r$ を $\varphi_{r,s}(\alpha)=(x_1, x_2, \cdots, x_s)\epsilon A^s$ に符号化し, それを通話路の送信側から送信したとき受信側で受信されるものを $y_1, y_2, \cdots, y_s$, その最後の $s-m$ 項から成る系列を $\beta=(y_{m+1}, y_{m+2}, \cdots, y_s)$ とおく. α, β の同次分布を ω とおき, α, β のとりうる値を $\alpha_1, \alpha_2, \cdots, \beta_1, \beta_2, \cdots$ とし, 各 β_k に対し α_{ik} を $\omega(\alpha_{i_k}, \beta_k)=\max\omega(\alpha_i, \beta_k)$ によって定義し, $M=U(\alpha_{i_k}, \beta_k)$ とおく. 以上の準備の下に Shannon の基本定理は次の形に定式化することができる.

定理 1. $H<C$ ならば各自然数 n に対し符号化 $\varphi_{n,n}$ を適当にとって

$$\lim_{n\to\infty}\omega(M)=1,\quad \lim_{n\to\infty}\frac{H(\alpha)-H_\beta(\alpha)}{n}=H$$

を成り立たせることができる.

多くの著者達は一定の符号化 $\varphi_{n,n}$ (n 固定) の下で通信文の長さを限りなく大きくするときの平均伝送速度の漸近的性質を論じているが，これは $\omega(M)$ が1に近いことを利用して n 項毎に復号する限り無意味である.

定理1の最後の関係式は雑音があっても平均伝送速度を情報源のエントロピーにいくらでも近づけ得ることを示している．しかし，これは通話路をフルに利用したことにはならない．平均伝送速度が通話路の伝送容量にいくらでも近づき得るような符号化が望ましい．これに関し次の定理が成り立つ.

定理 2．$H>0$, $C>0$ ならば，各自然数 r に対して自然数 s と符号化 $\varphi_{r,s}$ を適当にとって

$$\lim_{r\to\infty}\omega(M)=1,\quad \lim_{r\to\infty}\frac{H(\alpha)-H_\beta(\alpha)}{s}=C$$

を成り立たせることができる.

H と C との間の大小関係は問わないことに注意せよ．詳細は Ann. Inst. Stat. Math. 9巻2号参照.

昭和32年度第一研究部の研究概要

代理　林　知己夫

第一部長松下嘉米男はインド・ニューデリーの統計研究所 (I. C. A. R. の Statistical Wing) へ出張し，講議・研究指導を行っている．本年度に行われた研究室の研究状況は次の通りである.

松下；最小距離法による推論方式，決定函数理論の基礎仮定，距離概念に基く決定方式：藤本；分類の問題，製造工程を異にする鉄筋の強度に及ぼす影響の統計的判定の問題，ノン・パラメトリック検定の検定力の比較：本尾；統計量の極限分布 (現在，特に確率近似にあらわれる統計量及び順位統計量に対して)，連続マルコフ過程について (現在特に一次元連続マルコフ過程の極限的性質及び多次元連続マルコフ連続と調和函数との関係について)：石井；ノン・パラメトリック推論に現われるモーメント問題，特殊な分布の解析的特徴づけについて，マルコフ過程特に拡散方程式について：赤池；再生的な過程 (特に繰糸工程及び待合せ行列 (縦型タンデム) に関連して)，FACOM-128 によるモンテカルロ法の実用化，$Hx=\lambda Fx$ に関する疑似乱性による計算法：渋谷；FACOM-128 による実験データの解析，機械計算のための実用解析 (i 連分数による函数近似　ii 級数の収束を早める方法)，順序統計量の標本分布論；高野：情報理論の数学的研究，確率論における極限定理.

Non-sampling error (サンプリング台帳の精度)

大　石　　潔

一般にあるリストからサンプルを抽出して，観察または調査し集計分析する場合，サンプリング以後の操作を統計理論に従って如何に厳密に行ったとしても，サンプルを抽出するリスト (サンプリング台帳) が精確性を欠けば，得られる結論も信頼性が薄くなるであろう.

われわれは東京都23区民を対象として，屡々面接調査を行っているが，多くの場合，選挙人名簿または住民票を基にして，サンプリングを行っている．従ってこれらの台帳がどの位精確であるかを知る必要がある.

このため，都23区を地域，産業構成により層別し，第一段階として両台帳から独立にサンプルを抽出し，それらが他の台帳に登録されているかどうかをみる，(昭和32年8月実施)．第2，3段階として郵便調査，訪問調査を計画している.

第一段階の (単純な) 結果についてみると，選挙人名簿から抽出したサンプルに関しては，住民票に登録されているものが86.1%，選挙人名簿作成以後略1年間の移動 (転出) が7.5%，該当者なし2.7%，その他である．住民票から抽出したサンプルに関しては，選挙人名簿に記載されているものが84.1%，選挙人名簿作成以後略1年間に移動 (転入居) したものが，12.2%，該当者なし3.3%，その他である.

なお，1年間の移動の状態は，10月から翌年2月にかけて少なく，3月から8月にかけて多い.

(詳細な分析は今後に俟つとして) 以上の叙述から9月から翌年2月頃までは選挙人名簿・住民票のいずれからサンプルを抽出しても大差はないが，それ以後は住民票からサンプルを抽出することが望ましい.

The extreme deviate from the sample mean in the multivariate case

塩　谷　　実

$x_\alpha=(x_{1\alpha}, x_{2\alpha}, \cdots, x_{p\alpha})$, $\alpha=1.2, \cdots, n$ を平均 m, 共分散行列 Λ をもつ p 変数の正規分布からの大きさ n のランダム・サンプルとする．此の時標本平均 $\bar{x}$ からの偏差を

$$\chi_\alpha^2=(x_\alpha-\bar{x})\Lambda^{-1}(x_\alpha-\bar{x})' \quad (\Lambda \text{ が既知の時})$$

或いは

$$T_a{}^2=(\boldsymbol{x}_a-\overline{\boldsymbol{x}})\boldsymbol{L}^{-1}(\boldsymbol{x}_a-\overline{\boldsymbol{x}})' \quad (\Lambda \text{が未知の時})$$

で定義する．但し $\boldsymbol{L}$ は或る自由度を持つ Λ の不偏推定量で $(\boldsymbol{x}_a-\overline{\boldsymbol{x}})$ とは独立であるものとする．

茲では $\hat{\chi}^2{}_{MAX}=\max_a\{\hat{\chi}_a{}^2\}$ 及び $T^2{}_{MAX}=\max_a\{T_a{}^2\}$ について，$P_r\{\hat{\chi}^2{}_{MAX}>A\}$.，$P_r\{T^2{}_{MAX}>B\}$ を，A，B が大きい所で近似的に評価することを考えた．此れは普通の有意水準に対して臨界点を与えることになる．此等 $\hat{\chi}^2{}_{MAX}$，$T^2{}_{MAX}$ の応用としては処理が多次元的に測られる場合の analysis of dispersien に於いて，処理変動が有意に大きいと判断された場合，一番大きな平均を持つ処理が他のものに比して有意に大きいかどうかを検定するのに役立つものである．

偏微分方程式と確率過程との関係（特に多重散乱について）

横　田　紀　男

昭和32年度の研究方針は，昭和30年度研究した数学を，統計力学に応用することを目的として，昨年度に引続き研究して来た．その成果は，日本物理学会統計力学分科会で発表し，(イ) 32 年 5 月「圧力幅の密度依存について」 (ロ) 10 月「摂動計算によらず response function を求める試み」 (ハ) 33 年 3 月「密度行列の展開定理，virial 展開及び多重散乱の公式について」等である．

従来 path integral 法によって scattering matrix を求める試みは，多くの人の夢であったが，之を遂行するには多くの困難な点があった．特に energy 保存則等が満されてなく，之による計算の結果は信用し得なかった．筆者はその原因を究明し，実際への応用例でその夢を実現することに成功した．之は (ロ) である．

量子統計力学で実際に応用するにあたっての根本的な問題は，集団が同一要素から成っているとき，その要素間に correlation がなければ，各要素が独立に確率法則によって行動し，何等問題がないが，実際には各要素間に相互作用があって，独立な確率法則からのずれが生じる．そのずれを正確に取出すことであって，昨年度取上げた theme (イ) もそれが目的であった，然し応用例題 (イ) で実際に計算するに当って，之等を取扱うに重要な基本公式が出来てなく，満足に結果を得ることが出来なかった．筆者はそれに必要な基本公式を導出し． (ハ) で発表したのがそれである． (ハ) 述べたことは， (イ) の外に多く応用があって，量子で統計力学の今後の大きな発展が予想される．

(ハ) の公式を用いて実際に応用した未発表の成果を述べよう．一価金属の抵抗は不純物が入っていると極低温附近で，抵抗 minimum を示すことが分っている．その原因は未だ判ってなかった．之は各不純物が電流を独立に散乱する影響 (Markoff process) からのずれ，即ち各不純物が独立に散乱しない影響であって，その補正項を取り入れれば説明される．

之等については引き続き 33 年度も行う予定であり，(ハ) の応用による 33 年度の量子統計力学の発展は大いに期待されるものがある．

実験の計画について

樋　口　伊　佐　夫

いわゆる実験計画法と呼ばれるものは，狭い一定の考え方に依っているので，これを農学に於ける圃場試験以外の分野に適用する場合，根本的に検討すべきことが多い．しかるにこれのもつ形式的斉合性により，恰も一般性をもつかの如く思われる所に問題がある．そこで広い立場に立って問題点をあげ，経験の蓄積のない場に適用し得ると考えることが幻覚である可能性が多いことを示した．また例えば直交化の原理にしても，randomization の原理にしても，実験操作迄考慮すれば，一般原理といえる程のものではないことが，極くありふれた簡単な例によってもわかる．現存の計画法体系とは別の観点から研究すべきことが多い．

Discrete Decision Problems

鈴　木　雪　夫

問題とする分布が有限箇の点にのみ正の確率を与える離散的分布である場合の decision problem を取扱った．詳細は Ann. Inst. Stat. Math. Vol. 9, No. 3 に掲載してあるので，此処では，結果を要約しておく．

定理 1. Optimal decision function と同等な non-randomized decision function が常に存在する．

問題とする分布が atomless という条件の下で A. Dvoretky, A. Wald, J. Wolfowitz は optimal decision function と同等な non-randomized decision function の存在を証明している． (elimination of randomation).

吾々の結果は，atomless という条件が全く成立しない discrete の場合にも，このことが成立つことを示している．

次に，仮設検定の問題を吾々の手法で取扱った場合と，第一種の error を固定して，第二種の error を最小（或は minmax）にする場合とを取扱った．後者の場合，一般には randomized decision function を必要とするのであるが，特に randomized decision を必要とする sample point の数について次の定理が成立つ．

定理2．仮設に含まれる分布の数を m, 対立仮設に含まれる分布の数を m' とし，これらの分布の少くとも一つによって正の確率を与えられる sample point の数を J とすると，optimal decision function を用いたとき，randomized decision を必要とする sample point の数は高々 $\min(J, m+m'-1)$ である．

これらの結果を導くのには，linear programming の概念が有効に用いられた．特にすべての問題が linear programming の型に完成化されたことは興味深いことである．

32年度は上記の研究の他に，日本経済分析研究会及び多賀研究室との協同研究として経済計画の linear programming model について，理論と Facom 128 による計算法の研究を行っている．特に，この種の計算に大型自動機が必要欠くべからざるものであることを痛感する次第である．この研究の結果は，統計数理研究所彙報に掲載する予定である．

流量を予知する機構について

菅 原 正 己

流量を予知する方法については，数年来，種々の方式を試みて来たが，仮りに貯溜型の直列形式と名づけている方式が，最も有望であるように思われる．

貯溜型は，模型的に図-1 によって示される．2種類の out-put, 流出高 $y(t)$ および滲透高 $\eta(t)$ は，貯溜高 $X(t)$ の函数として定まる．

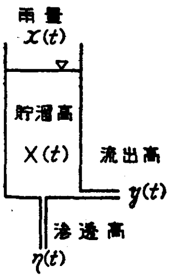

図-1

従って in-put と out-put との関係は，次の式で示される．

$$X(t)=\int_{-\infty}^{t}(x(\tau)-y(\tau)-\eta(\tau))d\tau$$

$$y(t)=f(X(t)),\quad \eta(t)=g(X(t))$$

以上の関係は，任意の函数 $f(X)$, $g(x)$ を含み，甚だ一般的であるが，目下の所，$f(x)$ としては折れ線で示される下に凸な増加函数を，$g(x)$ は X に比例するものを用いている．（淀川の日流量推定の時は，$g(x)$ として，折れ線で示される上に凸な増加函数を用いた）．これは，図-2 に示される機構で実現することができる．

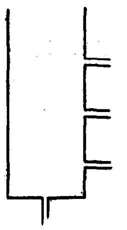
図-2

貯溜型の直列形式と云うのは，図-3 に模型的に示されるもので，貯溜型を縦に何段か組み合わせるものである．これは，地下の帯水層の構造とも対応し，また，特殊な場合として，実質的に従来の単位流量図法を含んでいる．この方式は，初期欠損の履歴を表わす点，流出高の大きさが雨量強度によって変化する点，中間流出の時間的遅れを自働的に与える点等の利点を持っている．

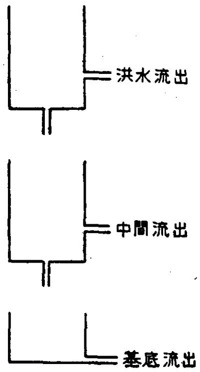

図-3

今年度は，この貯溜型の直列方式によって，宝川試験地（19 km^2），石狩川伊納（3140 km^2）の洪水流量の解析を行い，満足すべき結果が得られた．

交通現象の確率過程的分析，特に信号燈による交通整理について

植 松 俊 夫

交通に関しては，或種の確率過程のモデルにより数学的に処理出来る色々の問題があるが，ここではその一つとして，交叉点での交通整理の問題を取上げた．

十字路の交叉点に於て，2方向の自動車の流れに対し，普通の交互に青信号と赤信号とが夫々一定時間毎に繰返される如き信号燈により交通整理が行われる場合を考え，この場合青信号の長さと，赤信号の長さをどの様に割りふれば一番効率がよいかという事を問題とした．一番効率がよい場合という意味を，混雑がひどくなる迄の時間が二つの流れに対して釣合う如き場合なりと解する事とした．今各々の流れの交叉点での待行列の長さが，道路の容量によってきまる或る一定の限界を越える迄の時間を考え，それ等をあらわす確率変数を m, n とした時，$E(m)=E(n)$ となる様に青及び赤信号に与うべき時間を決める問題であると formulate した．こうすれば，自動車の到着に関し適当なモデルを仮定する事により，問題を random walk の問題として取扱う事が出来る．ここでは上の様な formulation により交通整理の問題を解く事が出来る為の条件を調べ，又それ等の条件の下に，実際

には一次方程式を解く事により，上の様な意味の一番効率のよい交通整理方法を決定出来る事を示した．詳細は当研究所の Annals (Vol. IX, No. 2) に述べられている．

経済活動指数数量化の試み

橋爪浅治

経済活動の活溌性を表すために，従来，経験的な判断から好況，不況という呼名で云われている．これは主観的でそのものづばりのこともあるが，また時には混乱され不明確となることもありうる．

われわれは経済活動に寄与すると思われる8箇の経済指数を1951年より1956年迄，月別にとりあげ，これに因子分析法を施こし，これの第1因子推定値をもって，経済活動指数と定義ずけることを試みた．以上のことを日本，米国，英国，西独の4ケ国について分析を行った．

Epidemic Model について

崎野滋樹

窮極目的は地域社会に於ける伝染病の伝播現象を統計的に推論することであって，その第一歩として先づ流行性感冒についての伝播現象を調べるために1957年10月20日～11月30日まで群馬県伊勢崎市で流感調査を行った．我々は伊勢崎市の2学区（殖蓮学区，三郵学区）に於ける学童を持つ全家族を対象として全数調査を行った．先づ集団に於ける伝染病一般の伝播模型を考えて見よう．時刻 $t=0$ で病人 a 人，又時刻 $t=t$ では病人 s 人，removal（免疫者，死亡者，隔離患者）q 人あるとする．そして時刻 t で病人が s 人あるとき Δt 時間に1人病人が増える確率を $\beta_1 s\Delta t$，次に removal に1人入る確率を $\beta_2 s^2\Delta t$ とすると次の階差方程式を導くことが出来る．今時刻 t で病人が s 人いる確率を $P_s(t)$ で表わすとき

$$\left.\begin{aligned}\frac{dP_s(t)}{dt}&=-(\beta_2 s+\beta_2 s^2)P_s(t)\\&\quad+\beta_1(s-1)P_{s-1}(t)+\mu(s+1)^2P_{s+1}(t)\\\frac{dP_0(t)}{dt}&=\mu P_1(t)\end{aligned}\right\}\quad(1)$$

probability generating function を $\varphi(v.t)=\sum_s v^s P_s(t)$ とおくと，

$$\frac{\partial\varphi}{\partial t}=\beta_2 v(1-2)\frac{\partial^2\varphi}{\partial v^2}+(\beta_1 v-\beta_2)(v-1)\frac{\partial\varphi}{\partial v}\quad(2)$$

但し境界条件は $\varphi(v.0)=v^a$

(2) の近似解として（$a=1$ として）

$$\varphi=\frac{1}{\sqrt{\pi t}}e^{\frac{1}{2}\int^V B_1(\xi)d\xi}\int_0^{\frac{\pi}{\sqrt{\mu}}}e^{-\frac{(V-\eta)^2}{4t}}\left(\sin^2\frac{\sqrt{B}}{2}\eta\right)\times e^{-\frac{1}{2}\int^\eta B_1(\xi)d\xi}d\eta\quad(3)$$

但し

$$\frac{1}{2}\int^\eta B_1(\xi)d\xi=\frac{1}{16}\left(-\frac{1}{2}\sin 2\sqrt{\beta_2}\eta+\sqrt{\beta_2}\eta\right)\frac{\beta_1}{\sqrt{\beta_2}}$$

これから時 t に於ける平均分散は容易に計算される．吾々は第一近似として上の模型に従って流感を分析し，更に一般的な伝染病の模型を推論したい．

物価指数について

石田正次

今までの物価指数についての一般的な考え方は，相対的な代表価格とでも云うべきものであって，物価の動きを表現するための一意的に定まる，ある種の指標であると云うことができよう．この代表という意味はモード的な意味での一般性と重要性の高い品目と標準的な取引（それは多くの場合現金取引と考えられている）に於けるそれらの価格を指すものであって，指数はこれ等の綜分によって算出されている．そして品目の選定を又多ゝの場合取引額の高い品目の中から主観によって行われている．そのような物価指数が実際の経済現象について科学的な数量として何を意味するかは明確でなく，これを無反省に経済の数量的な分析に取り入れたとするといくつかの不合理な点を生ずる．又価格が商社，取引量，取引方法などの条件によって大きく変るので一律に標準価格で通すことは困難であり，又利用度の代表として選定された品目は必ずしも価格変動の代表ではないということも問題である．

そこで物価指数の算出を次のようにして行う方がよいと考える．

1. 物価指数の利用目的からの定義づけ，利用目的に最もよく適合した場に於て，取引量，取引法，取引の相手，などを含めて物価指数を定義する．
2. 調査範囲の決定
   物価指数の定義に従って，どのような種類の取引を調査の対象にするかを決め，母集団の構成をする．この場合取引きを調査の単位とすることが多くの場合都合がよい．一つの取引には，品目，数量，価格，取引法，相手，期日が標識として対応する．
3. 層別
   取引はこれを直接調査することが困難であるから，商社を介して抽出する．そこで指数の目的によっ

て，商社の層別を行いその抽出を行う．

4. 物価調査

抽出された各商社に含まれている調査対象となる取引について，前記の標識を調査する．この場合，もし必要があれば取引の層別抽出を行う．

5. 集計計算

上のようにして調査された全標識を綜合して指数を算出する．

ここで品目の選定は相関分析，数量化によって定められ，指数の算定に於ても数量化的な手法が有効であると思われる．

これらの点について我々は二三の予備調査を行いつつある．

以上のような観点から物価指数を計算するには高度の調査技術と能率的な集計計算機構を必要とするので早急に実現することは不可能であるが我々はこれを一つの理想案として研究し，それを逐次的に実際の指数に接近させるよう努力したい．

昭和32年度第2研究部の研究概要

林 知己夫

第2・第3部共同の研究としては，前年度から引きつづきマスコミュニケイションの効果に関する問題がとりあげられ——我々のマスコミュニケイションの研究一般については彙報第5巻，第1号，1957, 80 p. を参照せられたい——EF VIII, EF IX の調査を実施した．この結果の概要は数研研究レポート，No. 3「マスコンの効果」にこれまでの結果と共に発表されている．また新聞記事の内容分析されたものと一般の人々の意見の流れとの関係（特に日ソ交渉に関する問題，原爆実験に関する問題）については創立第13週年記念講演会において，またその概略を「マスコミの影響」(1957年11月15, 16日，朝日新聞学芸欄）に発表した．なお内容分析に関しては，東大・新聞研究所の池内一氏に，日ソ交渉問題についての新聞記事の氏の方法による分析を依頼した．この結果と意見との対応関係を統計的に分析することを考えている．各研究者個人の発表はこれまでにある通りであるが，全般を通して整理してみると

（1） 基礎理論的なもの がまずあげられるが，これには，分類と言うことの統計数理的考察，予測と言うことの統計数理的考察，行為決定についての考察，実験の計画法についての統計数理的考察に関する研究があげられる．

（2） 社会調査法・標本調査法に関する研究

ソシオメトリック・パタンにおける凝集力指標推定（林）及びコムミュニケイションパタンのある特性（知識の伝達される鎖の長さ）の推定（多賀・鈴木達三）に関する標本調査理論の研究，非標本抽出誤差評価に関する研究（大石），面接調査法に関する誤差評価の研究（西平），社会調査における D.K. 群の特性に関する研究（西平），その他世論調査，市場調査，マーケッテングにおける統計的方法応用に関する研究（林・西平），労働組合の調査（多賀・鈴木雪夫・大石），工場立地条件の調査（多賀），村の類型決定に関する研究（多賀・大石），医師の態度調査（鈴木達三），厚生問題に関する調査（西平），年金問題に関する統計的研究（内田）がある．

（3） 実験の計画法に関する研究

実験の計画法を広く考え，最適な計画法はいかなるものであるかを，粒子統計，特にパッキングの問題・量子統計に関して考察した．具体的には膠質黒鉛の材料問題，洗炭等の粒子選別問題，物性の問題がとりあげられた．

（4） 分析法に関する研究

これについては第三研究部青山博次郎の項を参照せられたい．O.R. に関する諸問題もここでとりあげられている．この他経済における計画・物価の問題等も取扱われた．

（5） 数量化・予測に関する研究

多次元解析に関する理論的研究（林，塩谷，植松），類型決定の問題（林，石田，崎野），拡散に関する統計的研究（石田），経済量の取扱いに関する統計的研究（石田，橋爪），森林調査における幹材の特性と歩止りに関する研究及び航空写真利用に関する統計的研究（石田），多次元解析と数量化に関する研究（林），数量化に関する理論的研究及び実証的研究——政治的態度，デザイン，広告効果に関するもの——（林），感覚の量化に関する研究，特に味の問題について（塩谷）の研究・音響についての研究（林），態度測定法に関する統計数理的研究（林），因子分析法に関する統計的研究（橋爪），観測誤差の数量化的取扱いに関する研究（林，高倉），色空間の統計的構成法に関する研究（林），色彩統計に関する諸問題（林，石田，高倉，越谷，佐藤），危機場面における行動の統計的分析（石田）がある．

（6） 現象の経過過程模型に関する研究

今年度は交通現象に関する統計数理的分析，特に信号燈の赤青変化時間間隔の与える影響についての確率過程的分析法の研究（植松），伝染病伝播の統計的模型に関する研究（崎野），インフルエンザの実態調査結果の分析（崎野），心弾図の統計的分析法の研究（崎野）

があげられる．

（7） 計算法に関する研究

乱数発生機によるモンテカルロ法の研究（石田），$HX=\lambda^2FX$ の解法に関する研究（林，赤池，植松），熱気流とトルクに関する計算法の研究（石田）がある．

確率要素について

樋口順四郎

ここで確率要素とよぶのは，Fréchet にしたがって，抽象空間の値をとる確率変数のことである．すでに何人かの人がこのような抽象的な確率変数の議論をしている（S. Doss, R. Fortet, E. Mourier, S.K. Nasr など）．Banach 空間の値をとる場合は，Mourier の thèse で論じられているが，そこでは dual space の議論が対称的にいかない．dual space をもっと対称的に議論するには，Banach space でなくもっと一般な linear convex topological space で話をすすめた方が都合がよいかと考えられる．この空間はすでに前からとり上げられている（たとえば J. von. Neumann）．この空間の complification については V. Pták の議論を参照した．詳細は別に発表する．

逆行列の精度について

多賀保志

リレー計算機を用いて，60次および32次の Leontief 行列の逆転を行ったが，データ誤差の影響はともかくとして，つまり与へられた投入産出表の数量は全く正しいとしても，計算誤差（rounding-off error）を精密にチエックしておく必要がある．A^{-1} を与えられた行列 A の真の逆行列，$\widetilde{A}^{-1}$ を計算結果としてえられる A^{-1} の近似逆行列とし，誤差行列 $E=\widetilde{A}^{-1}-A^{-1}$ の各要素の大きさを知ることが大切である（普通は残差行列 $R=A\widetilde{A}^{-1}-I$ を作ってチエックする）．簡単な計算結果より，

$$E=\widetilde{A}^{-1}R(I+R)^{-1}=\widetilde{A}^{-1}\cdot\sum_{k=1}^{\infty}(-1)^{k-1}R^k$$

となることは容易にわかるが，実際に $R=A\widetilde{A}^{-1}-I$ を求めてみると，R の要素は高々 10^{-7} の order であるから，$E\fallingdotseq\widetilde{A}^{-1}\cdot R$ なる第1近似で十分 E の各要素の大きさがわかる．ただし，これは Leontief 行列の特殊な性質によるものと考えられるから，一般の場合には第2近似あるいはそれ以上の近似が必要となることはありうる．

面接調査の諸問題　その4

西平重喜

1957 年度には第2研究部第1研究室関係としては，表記の研究を続行した．例によって他の目的の調査に便乗して各種のデータを得た．今年度は主としてパネル調査の問題を扱った．そのくわしいことは，「彙報」5巻2号の同じ題目（ただしその3となっている）で発表した．

マス・コンの効果研究のグループの幹事役をつとめているが，今年も春，秋に調査をおこない，いままでに 13 回の調査をおこなったことになる．（ただし第 VI 次調査は4回に分けてあるので，第 IX 次調査まで）．今年は新聞記事の内容の分析にも着手した．その結果は「数研研究リポート」No. 3 として 1958 年4月始めに発表する．

また，層別の効果をみるために，日ソ関係の世論を投票区別のデータについて扱ってみた．この結果では層別の効果は小さい．（青写真資料あり）

さらに，上記の効果の調査のデータを通してみると，性，年齢，学歴，職業，支持政党別にみたときの各カテゴリ間の差は，主として女，老年，低学歴，職人・工員，保守派に無答が多いことによって起っていることが分かった　また，調査の回次によって，カテゴリ間の差が逆向きになることは，ほとんどない．なお，無答グループが行動をしいられたときどうするか，無答をのぞいた回答分布について，研究を続行中である．

研究指導普及室としては，「統計通信」第1号をつくり（3月 31 日出来），統計研究者の名簿ができた．

研究所外への援助としては，労働者婦人少年局の「主婦の生活と意見」（婦人関係資料シリーズ No. 23），最高裁判所調査官研修所の「非行少年の職業に対する態度」，「不和のある主婦の生活と意見」（同研修所のプリント）および，四国電力の従業員の世論（厚生，教育・訓練，モラール，労使への帰属性）調査をおこなった．

四国の調査は 603 事業所，6980 人の母集団から 1681 のサンプルをとったが，10 調査班，40 調査場，延 61 回の集合調査で，出席率 97% を得た．このような広くちらばった事業所の調査でも，会社と労働組合に十分話し合いができ，協力をあおげば，ほぼ完全なサンプルをつかむことができるわけである．なお，調査に対する意見では，有益，およびいくらか役に立つだろうか，40% ずつをこえており，否定的に考える

ものは数 % にすぎない.

そうして，賃金，年齢，入社時，学歴，住宅状況などについて，会社の記録とサンプルの答の分布はだいたい一致している. ただし賃金は会社記録平均22,200円であるが. 調査では19,800円となって，多少低い. 以上のことは日本鋼管の調査の場合にもいえることで——ただし，学歴の小学校卒業と中学（旧制高小をふくむ）の間にくるいがある——この程調査の，validity をあらわすものと考えられる.

昭和32年度第三研究部の研究概要

青山博次郎

第三研究部は計算法の研究，計算機械化の研究を行い，併せて研究指導普及を行っている. 本年度に於ては特有方程式の数値解法に関する研究，乱数作成機の利用によるモンテカルロ法の研究を行い，また具体的な問題としては Leontief 行列の逆行列（通産省60元，経済企画庁 32 元）を行い 30 時間余で 60 元の逆行列を計算することを得た. 外部機関との協同研究として日本経済分析委員会による Linear programming の計算を行い，また気象研究所と協同して数値予報を，東大鉱物研究室及び物性研究室と協同して結晶構造解析を行った. これらは何れも当研究所の継電器式万能自動計算機によった.

研究指導普及室では四国電力株式会社の従業員モラール調査，最高裁判所・家庭裁判所・調査官研修所の夫婦間に不和のある妻の調査，非行少年の職業に対する態度調査，労働省婦人少年局の主婦の社会的態度調査などの援助指導を行い，また我国における統計研究者の研究連絡をはかるため統計研究通信（研究テーマ，実績，雑誌論文目録など）を作成した.

第二研究部と共同してマス・コミュニケイションの効果の測定，内容分析を継続して行った.

私自身の研究としては前年に引つづき氷冷蔵庫の保有率，使用状況の分析による経済予測の問題，テストの信頼度の標本変動，ソシオメトリーにおける統計量の研究（Annals に報告）を行ったほか，日本鉄道技術協会の「踏切事故発生率算定方式研究」に協力し，踏切整備基準の統計数理的研究を行った（昭和 33 年 3 月 JREA 報告書及び講究会で発表). また総合研究「社会調査における bias の統計的研究」の代表者として，第二，第三研究部の援助を得て面接調査員による調査結果の偏りを研究した. この結果のうち調査員のゴマカシ（サンプル本人に会わないで作り上げたもの，サンプルを調べずに代人を調査したもの）がどのような要因によって生じたかを分析した結果だけについて記しておこう.

東京都から住宅的，商業的，工業的，混合的な地域を各1つづつえらび，48 人の調査員を用いて世論調査を行った. 各調査員は 12 人のランダム・サンプル，2人の植付けサンプル（さくら)，1人の移転サンプル計 15 人を調査し，その結果を，郵便調査，植付けサンプルからの報告，パネル調査などから検討し，ゴマカシの有無をしらべた. ゴマカシのあった調査員は 22 人，サンプル数にして 5.2% であった. このゴマカシを行った調査員を調査における involvement の程度，集合状況，Sheldon の形態学的類型（外胚葉型，中胚葉型，内胚葉型）を用いて簡単な数量化を行うだけで 81.8% の判断成功率を以て弁別することができた. なお事前に判明する要因だけを用いると 73.3% の判断成功率である. この外に趣味・娯楽のスコアだけを用いてゴマカシの調査員を弁別するときは判断成功率が 70.3% となった. 更に有効な数量化によって判断成功率は上昇するものと考えられる.