チェビシェフ型不等式について

石　井　恵　一

(1959 年 2 月 受 付)

On Tchebycheff Type Inequalities

Keiiti Isii

This paper gives a brief review of Tchebycheff type inequalities which appeared in the mathematical or statistical papers. Fourty-five types of inequalities are quoted in § 1, with short expositions and some examples. In § 2 some noteworthy points are discussed and some suggestions on future developments are given. The results quoted in § 1 are conveniently tabulated at the end of the paper.

Institute of Statistical Mathematics

§ 0. はしがき

この小文で述べようとすることは，いわゆるチェビシェフ型不等式についての総括的な紹介と，若干の批判とである．統計量の母集団分布が完全にはわかっていないが，いくらかの手がかり（特にいくつかのモーメントなど）が知られているとき，その統計量がある領域に落ちる確率の上限又は下限を求めようというのがチェビシェフ型不等式の問題である．殊に，近年の統計学では，分布に何ら具体的な型（正規分布等）を想定しないで，与えられたデータから何らかの統計的決定を行うノンパラメトリックな方法が重要視されるようになってきたが，その際に，決定の信頼度や危険度についての有力な手がかりを与えるものの一つがチェビシェフ型の不等式であろう．分布の型を想定しない場合の推論の信頼度を与えるには，この種のものが唯一最善のものであることはよく知られている通りである．また，いくつかのモーメントの知識にもとづいて，分布函数に適当な曲線をあてはめる場合，その近似がどのくらいの“妥当性”をもっているかということを知る手がかりも，この種の不等式によるほかはない．

さて，諸結果の紹介に入る前に問題を定式化しておこう．いま，考えている確率変数の分布はある分布の族 $\mathfrak{F}$ に属することだけがわかっているとするとき，この確率変数が与えられた領域 E に落ちる確率 $P(E)$ の値については，どの程度のことが云えるであろうか．すなわち，$\sup_{\mathfrak{F}} P(E)$ と $\inf_{\mathfrak{F}} P(E)$ とはどう与えられるかという問題である．最も極端な場合として，その確率変数の分布函数が完全にわかっているときは，$\mathfrak{F}$ が唯一つの分布だけから成る場合であり，$P(E)$ の値も一意に決定される．これとは逆に，確率変数の分布については，全く何もわかっていない時には，$\mathfrak{F}$ はあらゆる分布の集合であり，$P(E)$ の値も 0 と 1 の間ということしか云えない．これらの trivial な両極端の間に，分布について若干の information が与えられているという中間的な場合が問題であり，それが，われわれの扱う対象となる．たとえば，平均値 μ と分散 σ^2 とがわかっているときに，μ を中心とするある開区間の外に確率変数が落ちる確率の上限を与える式

$$P\{|X-\mu|\geqq k\}\leqq \min\left(\frac{\sigma^2}{k^2},\ 1\right)$$

はよく知られた Bienaymé-Tchebycheff の不等式であるが，このときには，$\mathfrak{F}$ は，平均値 μ，分散 σ^2 なる分布全体の集合である．また単峰分布（unimodal 分布）のモードのまわりの一次及び二次の絶対モーメントを用いて与えた Royden の不等式（(3.10) 参照）の場合には，$\mathfrak{F}$ は，与えられた一次及び二次のモーメントを持ち，原点をモードとする単峰分布の全体であり，E は原点を中心とする区間である．なお，Guttman の不等式（(4.1) 参照）のように，E が統計量の値によって定まるランダムな集合である場合もあることを注意しておこう．

Bienaymé は前掲の不等式の原型を 1853 年に与えたが，この種の不等式がはじめて系統的に研究された

のは，分布函数 $F(x)$ の最初の $2n$ 個のモーメントがきまっているときに，任意の x に対して $F(x)$ の値の上限及び下限を求めることであった．これは，Tchebycheff (1874) によってはじめて証明なしで与えられ，Markoff (1884) と Stieltjes (1884) によってほとんど同時に，しかも独立に，解かれた．Stieltjes は後に (1894) 別の証明を与えたが，モーメントに基くそれ以後の結果の主なもの Wald [47]，Royden [39] 等の方法はその流れをくむものであり，更に Mallows [29] も，その着想は明らかにこれに示唆されている．

モーメントのみの知識によって得られる不等式は，数学的には，統一的な方法で扱いやすい一つの体系をなすものではあるが，実用という見地から見ると，上限と下限の差が "広すぎる" という点に大きな不満がある．そのため，分布についてモーメント以外に，さらに何らかの知識を想定することにより不等式の効率をよくしようという試みがいろいろなされた．チェビシェフ型不等式は一般に

$$A \leqq P(E) \leqq B$$

のような形で与えられるが，A(又は B) を，もっと大きい（又は小さい）値で置きかえることができないものを最良不等式と呼ぶことにすれば，モーメントのみの知識にもとづいた最良不等式では，等号が成りたつのは，有限個の点の上にだけ質量をもつデイスクリートな分布になる．これは極端に滑らかでない分布の一種であるから，$\mathfrak{F}$ を滑らか(何らかの意味で)な分布の集合にして滑らかでないものを排除しておけば，もっと鋭い不等式が得られるであろうという見地から，分布函数の形に或る種の制約(殊に滑らかさ)を置くことが考えられる．その結果，単峰分布であるとか，$F(x)$ の高階微分可能性とその大きさに対する制約とか，あるいは更にそれらに加えて有界変域の確率変数であるとかいう条件を課したものが得られる．このような場合の最も包括的でかつ興味ある結果は Mallows [29] であろう．ところで分布函数の形に上のような "幾何学的な" 条件をつけることによって，不等式は大巾に改良されるのであるが，そこで問題になるのは，実際に適用する際に，どの程度までこれらの仮定を確かめることができるだろうかということで，実用上の立場からはいろいろの問題点が残っている．

一方，やはり極端に滑らかでない分布を排除するための一つの試みとして，分布函数の型に直接条件をつけるのではなく，統計量の性質に条件を置く方法がある（もっともこれも結局は分布函数に対する制約には違いないのだが，適用の場合から考えて一応区別した)．前述したように，モーメントのみによる不等式では，等号が成り立つのは，たいてい有限個（しかも少数）の点のみの上の分布である．ところが，たとえば，n 個の独立な確率変数の和は，十分大きい n に対してはあまり少数のディスクリートな値だけをとることはできないから，不等式は改良される筈である．この方向の結果でよく知られているのは Guttman [16]，Kolmogorov [23, 24] 等の結果であろう．このような制約のつけ方は，実用の立場からもしばしば要求されるもので，前者よりずっと自然なものといえる．

チェビシェフ型の不等式に属するものとしては，その他にもいろいろある．モーメントを一般化して，ある函数の期待値が与えられている場合もわりあい古くから扱われている．また一次元ばかりでなく，多次元への拡張も多く研究されているが，一次元の場合にくらべると，若干の断片的な事柄が得られているに過ぎない．きわめて最近，無限次元すなわち確率過程への拡張が Whittle [48] によってなされているが，重複対数の法則などと関連して考えると興味あることであろう．

以上で問題の概観とその分類とをごく大ざっぱに述べた．次節以下でこれらについての諸結果を綜合報告風に紹介したい．文献はできるだけ多く目を通したが，入手できないものもあるので，他の文献から間接的に引用したものもあることをお断わりしておきたい．なお，一つ一つについて証明の概略や解説を加えていると老大な紙数になってしまうので，特に重要なものを除いては，よく知られたものや他の文献に紹介されているものは，単に結果の羅列にとどめた．また 1954 年頃までの結果の多くは Godwin [14] にとりあげられているので，おもに 1955 年以後の結果に解説の重点を置いた．1954 年以前のもので直接入手できなかった文献の多くは Godwin [14] から引用したことをお断わりしておきたい．

参照の便宜のために，§ 3 に，結果の一覧表を与えておく．§ 1 における小項目の括弧つき番号は，一覧表における番号に対応している．

§ 1. 各種の結果とその解説

1.1 モーメントにもとづく不等式

1.1.1 モーメントによる不等式といえば誰でも直ちに想起するのは，Bienaymé または Tchebycheff の名を冠して呼ばれている不等式――平均値からのへだたりを分散を用いて確率的に評価する不等式――であろう．これはまた，非負確率変数の分布函数に対する Markoffの名で呼ばれている不等式から直ちに導かれるものであることもよく知られている．はしがきでも触れたように若干個のモーメントの知識によって分

布函数の大きさを評価しようとすることの，最初の系統的な研究は，n 個のモーメントによって分布函数の上限及び下限を与えようとする試みであり，Tchebycheff にはじまり，その弟子 Markoff に引きつがれ，更に Stieltjes 等も研究して，非常にきれいな完全な結果を与えたものであった．それ以後，その結果を一般化しようとする試みや，やや異なった条件下で同種の問題を扱おうという試みは，実用上の見地からも，また純数学的な見地からも，あきることなく続けられ，後述するような多くの著しい結果が得られている．モーメントのみの知識を前提としたこれらの立場は，数学的な面からは，§0 で述べた集合 $\mathfrak{F}$ が凸集合をなすという点にとり扱いの上での著しい利点があり，得られる結果も多くは秩序立った美しいものである．

一方実用的な面からは，前提とする information が少ないため，あまり鋭い結果が得られないという点に不満がある．

1.1.2 さて，この場合のおもな不等式とその解説を与えよう．最初に記号について述べておく．$F(x)$ を分布函数とするとき $\mu_j(\xi)$, $\nu_j(\xi)$ はそれぞれ点 ξ のまわりの j 次のモーメント及び絶対モーメントを表わすものとする．すなわち

$$\mu_j(\xi) = \int (x-\xi)^j dF(x),$$

$$\nu^j(\xi) = \int |x-\xi|^j dF(x).$$

特に $\mu_1(0)$ を単に μ とかき，また，$\mu_j(\mu)$, $\nu_j(0)$, $\mu_j(0)$を単にそれぞれ μ_j, ν_j, m_jとかくことがある．

さて，問題の第一は，

(1.1) $m_1, \cdots, m_{2n}$ が与えられているとき，$E=(-\infty, x]$ に対する $P(E)$ の値．これは分布函数 $F(x)$ の値をモーメントによって評価することであり，Tchebycheff はこれを証明なしで与えたということである．Markoff の方法と Stieltjes の第一の証明は，連分数の理論とある種の直交多項式の理論とを巧みに結びつけたもので，与えられた $2n$ 個のモーメントをもち $n+1$ 個の点の上にだけ分布しているある種の分布を構成し，これらの点のうち E に属するところでは 1，属さない点では 0 をとる非負多項式の期待値を考えることによって $P(E)$ の上限を得た．下限も同様である．後に Stieltjes は，点 x を含む適当な $n+1$ 個の点の上にのみ集中している分布をとると，与えられたモーメントをもつどの分布函数も，これと $2n$ 回か $2n+1$ 回交わることを示し，これを用いて同じ結果を導いた．解は一般に次のようにして求めることができる．今，固定した x に対して，次の行列式によって定義される z の多項式を，

$$Q(z) = \begin{vmatrix} 1 & z & z^2 \cdots\cdots z^{n+1} \\ 1 & x & x^2 \cdots\cdots x^{n+1} \\ m_0 & m_1 & m_2 \cdots\cdots m_{n+1} \\ m_1 & m_2 & \\ & \cdots\cdots & \\ m_{n-1} & m_n & \cdots\cdots m_{2n} \end{vmatrix}$$

とおいて $Q(z)=0$ の根（これは一般にはすべて実数値の単根になる．重根をもつのはモーメントが退化して分布が一意的に決ってしまうときだけである）を大きさの順に並べて $\alpha_0<\alpha_1<\cdots\cdots<\alpha_n$ とする（そのうちの一つは x に等しい）．

$$p_i = \prod_{j \neq i} \left(\frac{m-\alpha_j}{\alpha_i-\alpha_j}\right), \quad i=0,1,\cdots\cdots,n.$$

なる量を考える．ただし記号 m は，右辺の積を展開した式の m^k を m_k で置きかえることを表わす．すると $F(x)$ の上限と下限 $U(x)$ と $L(x)$ は次によって与えられる．

$$L(x) = \sum_{\alpha_i<x} p_i, \quad U(x) = \sum_{\alpha_i \leq x} p_i.$$

なお，$Q(z)$ において z^{n+1} の係数が 0 に等しくなったときには根の数が一つ減るが，結果は同じである．具体例を示そう．

（例） $m_0=1, m_1, m_2$ が与えられた時の $L(x), U(x)$ を求める．$m_2=0$ のときは，分布が 1 点に退化していて直接求められるから一般性を失わずに $m_1=0, m_2=1$ としてよいであろう．このとき $Q(z)$ は

$$Q(z) = (x-z)(1+xz)$$

となるから，$x<0$ なら，$\alpha_0=x$，$\alpha_1=-\dfrac{1}{x}$．

また $x>0$ なら，$\alpha_0=-\dfrac{1}{x}$，$\alpha_1=x$ となる（$x=0$ のときはその極限を考えればよい）．

容易に，$p_0=\dfrac{1}{1+\alpha_0{}^2}$，$p_1=\dfrac{\alpha^2}{1+\alpha_0{}^2}$ を得るから，

$$L(x) = \begin{cases} 0 & x \leqq 0 \\ \dfrac{x^2}{1+x^2} & x>0 \end{cases} \qquad U(x) = \begin{cases} \dfrac{1}{1+x^2} & x \leqq 0 \\ 1 & x<0 \end{cases}$$

を得る．片側検定に用いられる Uspensky の不等式はこれに含まれている．なお，4 次までのモーメントによる場合の解は Mallows [29], Zelen [50] に与えられている．

(1.2) 同じく $m_1, \cdots, m_{2n}$ が与えられているとき，Eとしてもっと一般の集合をとったらどうなるか．これについての一般的な結果とその具体例は筆者 [21] に与えられている．まず与えられたモーメント $m_0=1$, $m_1, m_2, \cdots, m_{2n}$ は，実際に，ある分布のモーメントとして実現されうるものでなければ意味がない（たとえば，$m_2<m_1{}^2$ などでは困る）から，そのような $m_1, \cdots, m_{2n}$ について考える．そのための条件は Ham-

burger のモーメント問題としてよく知られている通り，次のようなものである (Shohat & Tamarkin [44]).

$$\Delta_r = \begin{vmatrix} m_0\ m_1 \cdots\cdots m_r \\ m_1\ m_2 \cdots\cdots m_{r+1} \\ \cdots\cdots\cdots \\ m_r\ m_{r+1} \cdots\cdot m_{2r} \end{vmatrix} \qquad r=0,1,\cdots,n$$

とおくと，$m_0,\cdots,m_{2n}$ が分布のモーメントとして実現されうるためには ① ある k に対して $\Delta_0>0, \Delta_1>0,\cdots\cdots, \Delta_{k-1}>0, \Delta_k=\Delta_{k+1}=\cdots\cdot=\Delta_n=0$ か ② $\Delta_r>0,\ r=0,1,\cdots,n$ かでなければならない．① のときには分布は一意的に決定され，しかもちょうど k ケの点の上にだけ分布している．② の場合には無限に多くの分布が存在する．① の場合を退化した場合と呼ぶことにしよう．与えられたモーメントの列が退化しているならば，その分布は有限個の方程式をとくことによって直接求められるから，$P(E)$ の値（一意的!）もそれによってきまる．退化していないときには，$P(E)$ の上限と下限は，理論的には次のように決定される．「E の定義函数を $\chi_E(x)$[1] とし，高々 $2n$ 次の多項式 $f(x)$ で，すべての x に対して $f(x)\geqq\chi_E(x)$ なるものの全体を $\mathfrak{F}_1$，$f(x)\leqq\chi_E(x)$ なるものの全体を $\mathfrak{F}_2$ とする．また，記号的に，$f(m)$ は $f(x)$ の展開式の x^k を m_k で置きかえたものを表わすとする．すると，$L(E)=\sup\limits_{f\in\mathfrak{F}_2} f(m)$，$U(E)=\inf\limits_{f\in\mathfrak{F}_1} f(m)$ とおけば，$L(E)$，$U(E)$ はそれぞれ，$P(E)$ の下限，上限に一致する」というのが一般的な定理である．更に，E が閉集合ならば，$U(E)$ は m_{2n-1} までを満足する分布によって実際に達せられる．ただこのときその分布によって m_{2n} は必ずしも満足されないが，$U(E)$ にいくらでも近い $P(E)$ をとる分布で m_{2n} まで満足するものが存在する．E が開集合ならば，$L(E)$ の方について同様なことがいえる．しかし，この定理だけでは，実際にどのようにして $L(E)$，$U(E)$ を求めるかという点に問題が残るが，次の定理がこれに答える：「E が閉集合のとき上と同じ条件のもとで，実際に $U(E)=f_0(m)$ となる $f_0\in\mathfrak{F}_1$ が存在する．しかも，実際に $P(E)=U(E)$ となる分布 F_0 は，$f_0(x)=\chi_E(x)$ なる点 x の集合の上にのみ分布する．更に，もし f_0 がちょうど $2n$ 次なら，F_0 のモーメントは m_{2n} までの全部を満足する」これによって，E が与えられれば $f_0(x)=\chi_E(x)$ となる点の数がきまるから，計算によってその分布を求めればよい．(1.1) のように $E=(-\infty, x]$ の場合は，高々 $n+1$ 個の点の上の分布であり，E が有界区間の場合は高々 $n+2$ 個の点である．

(例 1) m_4 まで与えられたとき $P\{|X-m_1|\geqq k\}$ の上限を求めるはに次の手順による．一般性を失わずに $\mu=0$，$\mu_2=1$，$\mu_3\geqq 0$ としよう．高々4次の多項式 $f(x)$ で常に $f(x)\geqq\chi_E(x)$ なるものは，明らかに高々4つの点で $\chi_E(x)$ と一致する．故に $F_0(x)$ は高々4点の上にのみ分布している．一方，非退化の条件から，少なくとも3ケの点の上に分布していなければならないから，結局3点分布か4点分布に限られる．3点分布の場合には，その3点 x_1, x_2, x_3 のうち一つは E の境界であるからその自由度は 2，それと確率 p_1, p_2, p_3 とを併せて自由度5となり，$\mu_0(=1), \mu, \mu_2, \mu_3, \mu_4$ の5つの条件でとけばよい．また，4点分布のときは，図形的な考察から容易にわかるように，x_1, x_2, x_3, x_4 のうち2点は E の境界，残りの2点は一方によって他方がきまるから自由度 1，p_1, p_2, p_3, p_4 とあわせてやはり自由度5となる．あとは計算をすればよいわけである．一応この場合の結果を掲げておく．

［I］ $\mu_3>0$

（i） $-\dfrac{1}{k}\leqq\beta<0,\ \alpha\beta\leqq-1,\ 2\beta^2-\alpha\beta-k\leqq 0$ のとき，ただし，

$$\alpha=\frac{X+\sqrt{X^2-4Y}}{2},$$

$$\beta=\frac{X-\sqrt{X^2-4Y}}{2},$$

$$X=\frac{\mu_3 k^2-(\mu_4-1)k-\mu_3}{k^2-\mu_3 k-1},$$

$$Y=\frac{-k^2+\mu_3 k+\mu_4-\mu_3^2}{k^2-\mu_3 k-1}$$

とする．この場合 $U(E)$ は，

$$U(E)=1-\frac{1+\alpha k}{(\alpha-\beta)(k-\beta)}$$

で与えられる．

（ii） α, β の定義は（i）と同じとして $-k<\beta\leqq-\dfrac{1}{k}$，$-\beta<\alpha<k$，$\alpha\beta\geqq-1$ のときは

$$U(E)=\frac{1+\alpha\beta}{(k-\alpha)(k-\beta)},$$

これを μ_3, μ_4, k を使って直接かけば

$$U(E)=\frac{\mu_4-\mu_3^2-1}{k^4-2\mu_3 k^3+(\mu_4-3)k^2+2\mu_3 k+(\mu_4-\mu_3)^2}$$

（iii） $-k<\beta<0,\ \alpha\beta\geqq-1,\ 2\beta^2-\alpha\beta-k^2\geqq 0$，ただし

$$\alpha=\frac{X+\sqrt{X^2-4Y}}{2},\quad \beta=\frac{X-\sqrt{X^2-4Y}}{2},$$

$$X=\frac{\mu_3 k^2+(\mu_4-1)k-\mu_3}{k^2+\mu_3 k-1},$$

---

1) $\chi_E(x)=\begin{cases}1 & x\in E\\ 0 & x\notin E\end{cases}$

$$Y=\frac{-k^2-\mu_3 k+\mu_4-\mu_3{}^2}{k^2+\mu_3 k-1},$$

のときは，

$$U(E)=1-\frac{-1+k\alpha}{(\alpha-\beta)(k+\beta)}$$

(iv)　$\mu_3 \leqq \min\left\{\frac{(\mu_4-1)k}{k^2-1}, \sqrt{(k_2-\mu_4)(k^2-1)}\right\}$

のときは

$$U(E)=\frac{\mu_4-1}{k^4-2k^2+\mu_4}$$

(v)　t の方程式

$$2(k^2-1)t^3+3\mu_3 t^2-(\mu_4+k^3-k^2-k)t-\mu_3 k=0$$

が，$-\frac{k}{2}<\beta<0$, $\mu_3 \geqq (1-k^2)\beta$, $\mu_3 \geqq (1-k^2)\alpha$, $\mu_3 \leqq k\alpha\beta+(\alpha+\beta)+k$, $\mu_3 \leqq -k\alpha\beta+\alpha+\beta+k$, ただし $\alpha=2\beta-\frac{k^2}{\beta^2}$，を満足する根 β をもつならば

$$U(E)=1-\frac{(k^2-1)\alpha+\mu_3}{(k^2-\beta^2)(\alpha-\beta)}$$

(vi)　その他の場合は　$U(E)=1$,

(II)　$\mu_3=0$

(vii)　$k^2 \geqq \mu_4$ ならば

$$U(E)=\frac{\mu_4-1}{k^4-2k^2+\mu_4}$$

(これは (iv) に含まれる)

(viii)　$1 \leqq k^2 \leqq \mu_4$ ならば

$$U(E)=\frac{1}{k_2},$$

(ix)　$k \leqq 1$ ならば，$U(E)=1$.

以上の結果を眺めると，次の諸点が観察される．(viii) の場合には，f_0 は4次でなく2次となり，このときの分布 F_0 は μ_3 までは満足するが，μ_4 を満足しないものの例である．また，k を $k^2>\mu_4$ なる如く固定して．$|\mu_3|$ を充分小にすれば上の (iv) の場合になる．これは，k を固定したとき $|\mu_3|$ が充分小さければ，$U(E)$ は μ_3 に無関係になることを示している．ところが逆に，$\mu_3(\neq 0)$ を固定して，k を充分大にすれば，(ii) になり，k がいくら大きくてもわずかではあるが μ_3 に影響される．

(例 2)　μ, μ_2 がわかっていて，E が μ を含む任意の開区間である時の $L(E)$．これは Selberg [42] が最初に与えた結果であるが，上述の原理を適用すれば直ちに出てくる．例によって $\mu=0$ とし，$E=(-\alpha,\beta)$, $\beta \geqq \alpha \geqq 0$ とする．F_0 は $-\alpha, \beta, \frac{\beta-\alpha}{2}$ の3点分布であるか又は，$-\alpha$, $\gamma\left(0 \leqq \gamma \leqq \frac{\beta-\alpha}{2}\right)$ の2点分布であるから，方程式は

$$\begin{cases} p_1+p_2+p_3=1 \\ -p_1\alpha+p_2\beta+p_3\frac{\beta-\alpha}{2}=0 \\ p_1\alpha^2+p_2\beta^2+p_3\left(\frac{\beta-\alpha}{2}\right)^2=\mu_2 \end{cases}$$

又は

$$\begin{cases} p_1+p_2=1 \\ -p_1\alpha+p_2\gamma=0 \\ p_1\alpha^2+p_2\gamma^2=\mu_2 \end{cases}$$

となる．従って，

$$\begin{cases} p_1=\frac{2\mu_2+\beta(\beta-\alpha)}{(\alpha+\beta)^2} \\ p_2=\frac{2\mu_2-\alpha(\beta-\alpha)}{(\alpha+\beta)^2} \\ p_3=\frac{4\alpha\beta-4\mu_2}{(\alpha+\beta)^2} \end{cases}$$

又は

$$\begin{cases} \gamma=\frac{\mu_2}{\alpha} \\ p_1=\frac{\gamma}{\gamma+\alpha}=\frac{\mu_2}{\alpha^2+\mu_2} \\ p_2=\frac{\alpha}{\gamma+\alpha}=\frac{\alpha^2}{\alpha^2+\mu_2}. \end{cases}$$

$p_i \geqq 0$ 及び $0 \leqq \gamma \leqq \frac{\beta-\alpha}{2}$ の条件により，次の結果を得る．

$\alpha(\beta-\alpha) \geqq 2\mu_2$　　ならば $p(-\alpha,\beta) \geqq \frac{\alpha^2}{\alpha^2+\mu_2}$

$2\alpha\beta \geqq 2\mu_2 \geqq \alpha(\beta-\alpha)$ ならば $p(-\alpha,\beta) \geqq \frac{4\alpha\beta-4\mu_2}{(\alpha+\beta)^2}$

$\mu_2 \geqq \alpha\beta$　　ならば $p(-\alpha,\beta) \geqq 0$.

(1.3)　$m_1,\cdots,m_n$ が与えられていて，かつ確率変数の変域がある有界区間 $[a,b]$ であるとき（換言すれば，分布函数 $F(x)$ が，$F(x)=0$ $(x<a)$, $F(x)=1$ $(x>b)$ のとき)．このとき $F(x)$ の値の上限と下限についての結果と証明は，Shohat & Tamarkin [44] に詳述されている．上限又は下限を実際にとる分布は次のような点から成るディスクリート分布である．n が偶数ならば，x 及び他の $\frac{n}{2}$ 個以下の点，又は a, x, b 及び他の $\frac{n}{2}-1$ 個以下の点．また n が奇数ならば，a, x と他の $\frac{n-1}{2}$ 個以下の点，または x, b と他の $\frac{n-1}{2}$ 個以下の点．

(1.4)　任意個数の絶対モーメント $\nu_{i1},\cdots,\nu_{in}$ がわかっているときの $E=[-d,d]$ に対する $P(E)$ の値．これは Wald [47] によって与えられた．絶対モーメントが必らずしも最初の n 個でなく，番号がとびとびでもよい点に一般性がある．Wald は分布を原点に関して折り返すことにより，正値確率変数に変換しておいて，その分布函数の値を評価した．その際の方

法は，適当なディスクリート分布を構成して，他の分布函数との交点の数による特徴づけによるもので，基本的には，(1.1) における Stieltjes の方法を踏襲したものであるが，任意のモーメントという点で若干の技巧が必要である．また，この結果は，Wald と全く別の方法 ((1.2) における筆者 [21] の方法) でも導かれる．Wald の結果によると次のようになる．F_0 は，n が偶数ならば，x と他の $\frac{n}{2}$ 個以下の点とから成るディスクリート分布であるか又は，$0, x$ と $\frac{n}{2}-1$ 個以下の点 (この場合 F_0 の i_n 次モーメントだけは ν_{i_n} より小さくなる場合があることは (1.2) と同様である) とから成るものである．また n が奇数ならば，F_0 は 0, x と $\frac{n-1}{2}$ 個以下の点とから成る．

なお，Wald [47] は，与えられた数値 $\nu_{i_1},\cdots,\nu_{i_n}$ に対して，それらを $i_1,\cdots,i_n$ 次のモーメントとして持つ分布の存在条件を与えている．しかし，実際にこれらの数値が与えられたとき，手っとり早い方法は，上述したようなディスクリート分布を構成してみるのがよい (有限個の方程式をとけばよい)．もし，そのような分布が構成できなければ (たとえば，方程式をといた結果，確率が負になったりすれば)，与えられたモーメントをもつ分布は存在しないのである．具体的な例についてみよう．

(例) ν_n, ν_{2n} が与えられたとき上述したような分布は (i) x に確率 p，他の一点 a に確率 $1-p$ をもつか，または，(ii) x に確率 p，0 に $1-p$，かつそのときの $2n$ 次の絶対モーメントは $\leqq \nu_{2n}$，のどちらかである．

(i) の場合．方程式は $px^n+(1-p)a^n=\nu_n$，$px^{2n}+(1-p)a^{2n}=\nu_{2n}$，$a$ を消去して，$(\nu_n-px^n)^2=(1-p)\times(\nu_{2n}-px^{2n})$．ゆえに $p=\dfrac{\nu_{2n}-\nu_n^2}{\nu_{2n}-\nu_n^2+(x^n-\nu_n)^2}$，$a^n=\dfrac{\nu_n x^n-\nu_{2n}}{x^n-\nu_n}$ を得る．ここで，$0\leqq p\leqq 1$ の条件より $\nu_{2n}\geqq\nu_n^2$ を得る．$a\geqq 0$ としてよいから，$x^n\leqq\nu_n$ か，または $\nu_n x^n\geqq\nu_{2n}$ である．前者の場合は $a\geqq x$，後者のときは $a\leqq x$ に対応する．

(ii) の場合．$px^n=\nu_n$，$px^{2n}\leqq\nu_{2n}$ をといて，$p=\dfrac{\nu_n}{x^n}$ を得る．$0\leqq p\leqq 1$ の条件から $x^n\geqq\nu_n$．従ってまた，$\nu_n^2\leqq\nu_{2n}$ を得る．結局，(i), (ii) をまとめて，ν_n, ν_{2n} のみたすべき条件は $\nu_{2n}\geqq\nu_n^2$ であり，そのとき，

(a) $x^n\leqq\nu_n$ ならば，

$$0\leqq F(x)\leqq\frac{\nu_{2n}-\nu_n^2}{\nu_{2n}-\nu_n^2+(x^n-\nu_n)^2}.$$

(b) $\nu_n\leqq x^n\leqq\dfrac{\nu_{2n}}{\nu_n}$ ならば，

$$1-\frac{\nu_n}{x^n}\leqq F(x)\leqq 1.$$

(c) $x^n\geqq\dfrac{\nu_{2n}}{\nu_n}$ ならば，

$$\frac{x^n-\nu_n}{\nu_{2n}-\nu_n^2+(x^n-\nu_n)^2}\leqq F(x)\leqq 1.$$

これらの不等式は Cantelli の得たものと同じである．

(1.5) μ, μ_2, μ_4 が与えられ，$E=I^{\smile}-I$ のときの $P(E)$ の下限．ここに I はある区間を表わし，$-I=\{-x; x\in I\}$ である．これは Guttman [14] の与えたもので，

$$P(\mu-k_1\sigma, \mu-k_2\sigma)+P(\mu+k_2\sigma, \mu+k_1\sigma)\geqq 1-\frac{1}{\lambda^2}.$$

ここに，$k_1=(1+\lambda\alpha)^{\frac{1}{2}}$，$k_2=(1-\lambda\alpha)^{\frac{1}{2}}$ または，

$$k_1=(1+\alpha(\lambda^2+1)/\sqrt{\lambda^2-1})^{\frac{1}{2}},$$
$$k_2=(1-\alpha\sqrt{\lambda^2-1})^{\frac{1}{2}} \quad \text{または}$$
$$k_1=(1+\alpha\sqrt{\lambda^2-1})^{\frac{1}{2}},$$
$$k_2=(1-\alpha(\lambda^2+1)/\sqrt{\lambda^2-1})^{\frac{1}{2}}.$$

ただし，$\sigma^2=\mu_2$，$\mu_4=(\alpha^2-1)\mu_2$ とおいた．これが最良不等式であるためには，次のことがそれぞれ必要である．$\alpha\lambda\leqq 1$，$\alpha\sqrt{\lambda^2-1}\leqq 1$ であるか $\alpha\leqq\sqrt{\lambda^2-1}$．

(1.6) Lurquin [28] は，変域有界の場合に，μ, μ_2 が与えられたとして

$$P(\mu-t\sqrt{\mu_2}, \mu+t\sqrt{\mu_2})\leqq 1-\frac{\mu_2(1-t^2)}{z^2}$$

を与えた．

ここに，z は，μ と変域内の点との距離の最大値である．これは最良不等式でないが，修正して次の最良不等式にすることができる．

$$P(\mu-t\sqrt{\mu_2}, \mu+t\sqrt{\mu_2})\leqq 1-\frac{\mu_2(1-t^2)}{z^2-t^2\mu_2}\quad (t>1).$$

1.2 一般の函数の期待値が与えられたときの不等式．

1.2.1 前節においてモーメントが与えられた場合について一通り述べたが，モーメントというものは分布函数の集合の上で定義された一種の汎函数である．そこで，これを拡張して一般にいくつかの汎函数が与えられたときに，それにもとづいて確率を評価しようとする試みが von Mises 等によってなされ，最近 Blum [8] 等が一般的な形で研究している．ある意味では，この方向は，数学的には極めて自然的で，また秩序立ったものである．なぜならば，ある区間の確率の値そのものも，分布函数全体の集合の上で定義された一つの汎函数であり，また，分布函数は適当な可算個の汎函数の値によって一意的に決定されることは明らかだから，適当な有限個の汎函数の値によって他

の汎函数の値がどの程度まで規制されるかを見ることは，数学的に極めて興味深いことである．

モーメント以外の汎函数を用いる場合でも，多くは，汎函数として適当な函数の期待値を考える．

1.2.2　主な結果は次の通りである．

(2.1)　Markoff の不等式と本質的には同じものであるが，やや一般な形で得られるものに次の不等式がある．$g(x)$ が偶函数で，$x \geqq d \geqq 0$ において単調非減少であるとすれば

$$P(-d, d) \geqq 1 - \frac{E[g(x)]}{g(d)}.$$

ここに $E[g(x)]$ は $g(x)$ の期待値を表わす．これは全く明白な不等式であろう．$g(x)=|x|$ とおけば，いわゆる Markoff の不等式を得る．

(2.2)　二つの函数の期待値がわかっているときについて，von Mises [32] が次の結果を示している．$g(t), h(t)$ を $0 \leqq t \leqq a (\leqq \infty)$ において定義された単調増加な函数で $g(0)=h(0)=0$, かつ x-y 平面上の曲線 $\Gamma=\{g(t), h(t); 0 \leqq t \leqq a\}$ は下に凸なるものとする．$g=E[g(x)], h=E[h(x)], C=(g, h), Q=(p, q)=(g(a), h(a))$, とし，直線 OC が Γ と $(g(t_1), h(t_1))$ で交わり，直線 QC は Γ と $(g(t_2), h(t_2))$ で交わるものとする．更に $M=(g(d), h(d))$ とし，(g', h') を直線 CM と Γ との交点とする．そうすると，$E=[0, d]$ に対し，

$$d \leqq t_2 \text{ ならば } 0 \leqq P(E) \leqq 1 - \frac{g-g(d)}{g'-g(d)};$$

$$t_2 \leqq d \leqq t_1 \text{ ならば}$$

$$\frac{(q-p)(h-h(d)-(g-g(d))(h-q)}{ph(d)-qg(d)} \leqq P(E) \leqq 1 - \frac{gh(d)-hg(d)}{ph(d)-qg(d)};$$

$$t_1 \leqq d \text{ ならば } \frac{g-g(d)}{g'-g(d)} \leqq P(E) \leqq 1.$$

しかもこれらは最良不等式であり，等号の起るのは高々3点分布のときである．

(2.3)　有界変域の分布に対して，函数列の期待値から，分布函数の値を評価することを，Blum [8] がやっている．$\{f_n(y)\}$ を連続函数の列とし，$F(y)$ を $[-A, A]$ 上の分布函数とする．

$a_n=\int_{-A}^{A} f_n(y)dF(y)$ とおくとき，数列 $\{a_n\}$ によって分布函数 $F(y)$ が一意に決定されるための必要かつ充分な条件は，$\{f_n(y)\}$ が，$[-A, A]$ における連続函数全体の作る Banach 空間で稠密なことである．この条件が満足されているとき，ディスクリートな分布函数 $F_n(y)$ を作って，$F_n(y)$ が $F(y)$ に法則収束するようにできる．

1.3　分布函数の滑らかさに条件のある場合の不等式．

1.3.1　前にも触れたように，モーメントだけの知識では，よほど多数のモーメントを知ることができない限り，得られる不等式は非常に鈍いものであり，また，予備知識として与えられるモーメントも，実際にはせいぜい三次か四次くらいまでであり，パラメトリックの場合とくらべて効率の極めて低い検定しかすることができない．そこで，分布函数が正規分布であるとか一様分布であるとかいうほど具体的にはわからないにしても，ごく大ざっぱに，対称な分布であるとか，単峰分布であるとかいうような何らかの幾何学的な制約をおくことができる場合には，もっとずっと能率のよい不等式が得られるだろうと考えられる．この主張を裏づける根拠として次のことを考えるとよい．1.1 において，モーメントのみの知識を用いて得た不等式において等号の成立するのは多くの場合，有限個の値のみをとるディスクリートな分布であり，それは極端に滑らかでない分布である．そこでこのような分布を排除すれば不等式の上限や下限はもっと改良されるに違いないと推測される．この見地から，いろいろの条件のもとで，チェビシェフ型の不等式を得るための研究がなされている．Gauss の不等式と Selberg の不等式を除けば，多くは非常に複雑な結果で使いにくい．理論的な研究として最も系統立っていて興味があるのは Mallows の研究であろう．しかも，最も重要な定理の一般の場合に対する証明が未解決であり，今後に多くの問題を残している．

1.3.2　おもな結果を紹介しよう．

(3.1)　年代の順序は逆になるが，この方向へのかなり体系的な興味ある研究として，Mallows [29] を挙げねばならない．

分布函数を，その“滑らかさ”の程度に応じて分類する．分布函数 $F(x)$ が滑らかさの条件 (k, λ) を満足するとは，

1)　$F(x)$ の $k+1$ 階の導函数が存在して連続，

2)　$k+2$ 個の数 $\beta_0, \beta_1, \cdots, \beta_{k+1}$ が存在して
　$\beta_i < x < \beta_{i+1}$ で $0 < (-1)^i F^{(k+1)}(x) < \lambda$,
　$i=0, 1, \cdots, k.$

ただし，$F(\beta_0)=0$, $F(\beta_{k+1})=1$ とする．($\beta_0=-\infty$, $\beta_1=+\infty$ も考える)．

を満足する場合をいう．そのとき，次のような函数 $L(x)$ と $U(x)$ を定めることが目的である：

$F(x)$ が　(i) $\int_{-\infty}^{\infty} x^s dF(x)=m_s \quad s=0, 1, \cdots, n,$

(ii)　滑らかさの条件 (k, λ) の両者を満足するならば常に $L(x) < F(x) < U(x)$ となるような，しかも最良の $L(x)$ と $U(x)$.

たとえば，滑らかさの条件が $(0, \infty)$ ならば，(1.1)の場合になる．また $(1, \infty)$ は，単峰分布の場合である．上の問題をモーメント問題 (n, k, λ) と呼ぶ．Mallows は $(2n, k, \lambda)$ について興味ある研究をしているが，ここにまとまりよく紹介するにはあまりに詳細，多岐にわたっている．しかも証明の完成しているのは，比較的簡単な若干の場合であり，一般の場合については，結果の正しいことが推測されているだけで，証明は完成されていない．基本的な方法はやはり，(1.1) に対する Stieltjes の第二の方法に示唆されているようで，ある種の型の分布（それは，$(2n, k, \lambda)$ の解の極限と考えられる．これを極値分布と呼ぼう）を構成すると，$(2n, k, \lambda)$ の任意の解はこれと $2n$ 回または $2n+1$ 回交わる．このような極値分布のグラフの全体は，平面上のある領域を帯上に埋めるが，各 x_0 に対して直線 $x=x_0$ によるこの帯状領域の切口が区間 $(L(x_0), U(x_0))$ を与えると考えられる．Mallows はまた，分布の Stieltjes 変換を有効に用いて理論を展開し，実際に $(2n, 1, \infty)$, $(2n, 2, \infty)$, $(2n, 0, \lambda)$, $(2n, 1, \lambda)$ 等について一般的な解き方を与え，また，$(2, 0, \infty)$, $(2, 1, \infty)$, $(2, 2, \infty)$, $(2, 0, \lambda)$, $(2, 1, \lambda)$, $(4, 0, \infty)$, $(4, 1, \infty)$ の各場合について数値的な結果を示している．これらのうち $(2, 0, \infty)$ は，1.1. で扱ったものと同じであるから，$(2, 1, \infty)$ の場合について，解き方の手順を示してみよう．$(2n, k, \infty)$ において $n=1$, $k=1$ の場合であることを考慮して，和が k になる $n+1$ 個の非負整数の順列はこの場合 $(0, 1)$, $(1, 0)$ の二種となる．そこで $(1, 0)$ に対して次のような分布函数 $E(x)$ を作る：2点 α_0, α_1 $(\alpha_0<\alpha_1)$ があって，α_0 で $E(x)$ が単純不連続性をもち，また，α_1 で $E'(x)$ が単純不連性をもつ．しかも，$E(x)$ は点 α_0 に確率 p をもち，区間 (α_0, α_1) に確率 $1-p$ が一様に分布しているようなものである．このような分布に，与えられたモーメントの条件（常に $m_1=0$, $m_2=1$ とする）をあてはめることによって

$$\alpha_1=\frac{-(3+\alpha_0^2)}{2\alpha_0},\quad p=\frac{3-\alpha_0^2}{3(1+\alpha_0^2)}$$

を得る．従って，$-\sqrt{3}<\alpha_0<0$ でなければならない．そのとき $E(x)$ は，

$$E(x)=\begin{cases}0 & x\leqq\alpha_0\\ 1-\dfrac{4\alpha_0^2(3+d_0^2)}{9(1+\alpha_0^2)^2}-\dfrac{8\alpha_0^3}{9(1+\alpha_0^2)^2}x & \alpha_0<x\leqq\alpha_1\\ 1 & \alpha_1<x\end{cases}$$

となる．$(0, 1)$ に対してはこれと対称な結果になる．そこで $-\sqrt{3}<\alpha_0<0$ なる範囲を α_0 が動くとき及びそれと対称なときに $E(x)$ のグラフが埋めつくす帯状領域を考えると，その下側のふちが $L(x)$，上側のふちが $U(x)$ を与えるのである．かくして，$L(x)$ は，$x>0$ ならば一部は点 $(\alpha_0, 1-p)$ の軌跡，又一部は $E(x)$ のグラフの $\alpha_0<x\leqq\alpha_1$ に対応する部分の包絡線である．これを考察することにより，結局，

$$L(x)=\begin{cases}\dfrac{4x^2}{3(1+x^2)} & 0\leqq x\leqq\sqrt{\dfrac{5}{3}}\\ 1-\dfrac{4}{9}\dfrac{1}{1+x^2} & \sqrt{\dfrac{5}{3}}<x\end{cases}$$

$$U(x)=1\qquad 0\leqq x.$$

$x\leqq 0$ のときはこれと対称な結果を与える．

以上が $(2, 1, \infty)$ の場合の具体的な解なのであるが，$(2, 1, \infty)$ という条件の内容的な意味は，容易にわかるように，μ, μ_2 の与えられた単峰分布の分布函数を評価することである（ただしモードの位置は指定しない）．実際，$L(x)$ における係数 $\frac{4}{9}$ は，直ちに単峰分布に対する Gauss の不等式を想起させる．ただし，ここに与えた不等式は片側検定に対応するものであるが，Gauss のは両側検定に対応し，しかもモードの位置も指定されている．

その他の場合について結果だけ引用しよう．

$(2, 0, \lambda)$ の場合，これは μ, μ_2 が与えられ，しかも密度函数が有界 $(<\lambda)$ であるような分布のみを考えていることに相当する．（このときも $\mu=0$, $\mu_2=1$ としていることもちろんである．以下の諸例も同様）．このときは必然的に $\lambda>\frac{1}{2\sqrt{3}}$ が必要で，$L(x)$ は $x\geqq-\frac{1}{2\lambda}$ に対しては次の方程式の一根である．

$$\left(x+\frac{1}{2\lambda}\right)L^2-\lambda\left\{\left(x+\frac{1}{\lambda}\right)^2+1-\frac{1}{3\lambda^2}\right\}L+\lambda\left(x+\frac{1}{2\lambda}\right)^2=0.$$

$U(x)$ はこれと対称な結果になる．$x\to\infty$ のとき $L(x)$ の漸近展開は

$$L(x)=1-\frac{\sigma^2}{x^2}+\frac{\sigma^2}{x^4}\left(1-\frac{1}{3\lambda^2}\right)+\frac{\sigma^4}{\lambda x^5}+\cdots\cdot,$$

ここに $\sigma^2=1-\frac{1}{12\lambda^2}$．ここで $\lambda\to\infty$ としてみると，(1.1) で述べた例と同じ結果

$$L(x)=1-\frac{1}{x^2}+\frac{1}{x^4}-\frac{1}{x^6}+\cdots\cdots$$

に帰する．

$(4, 0, \infty)$ の場合，これは μ_4 までが与えられ，他に何も条件がない場合と同じである．（従って，1.1 に述べた方法でもできる）．モーメントの満足すべき一般の条件から（$\mu=0$, $\mu_2=1$ として），

$$\varDelta=\mu_4-\mu_3^2-1>0$$

が満足されねばならない．結果は次のようになる．

$\delta_1=\frac{1}{2}(\mu_3-\sqrt{\mu_3{}^2+4})$, $\delta_2=\frac{1}{2}(\mu_3+\sqrt{\mu_3{}^2+4})$ とおくとき,

（i） $x\leqq\delta_1$ ならば, $L(x)=0$,

$$U(x)=\frac{\varDelta}{(x^2-\mu_3x-1)^2+\varDelta(1+x^2)}.$$

（ii） $\delta_1<x<\delta_2$ ならば, $L(x)=\frac{1+\beta x}{(x-\alpha)(\beta-\alpha)}$,

$$U(x)=1-\frac{1+\alpha x}{(\beta-x)(\beta-\alpha)},$$

但し, α,β は, z の方程式 $Q_0(x)z^2-(\mu_3Q_0(x)+x\varDelta)z-Q_0(x)-\varDelta=0$ $(Q_0(x)=-x^2+\mu_3x+1)$ の二根 $(\alpha<\beta)$.

(iii) $x\geqq\delta_2$ ならば,

$$L(x)=1-\frac{\varDelta}{(x^2-\mu_3x-1)^2+\varDelta(1+x^2)},\quad U(x)=1.$$

$(2,1,\lambda)$ の場合. すなわち, m, m_2 の与えられた単峰分布で, かつ, 密度函数の導函数が λ でおさえられている場合である. $L(x)$ と $U(x)$ は次のように与えられる. 一般の $(2n,1,\lambda)$ の場合の原理をも示唆するために, 手順を少し詳しく述べる. 便宜上 $m_0=1$ とおいて, z の函数

$$I(z)=\frac{m_0}{z}+\frac{m_1}{z^2}+\frac{m_2}{z^3}$$

を考え, $e^{\frac{1}{\lambda}I'(z)}$ を $\frac{1}{z}$ の昇べき順に展開して, $\frac{1}{z^5}$ 以下の項を省いたものを

$$J(z)=1-\frac{1}{\lambda}\left(\frac{1}{z^2}+\frac{2\mu_1{}^*}{z^3}+\frac{3\mu_2{}^*}{z^4}\right)$$

とすれば, $\mu_1{}^*,\mu_2{}^*$ は m_0,m_1,m_2,λ によって定まる. いま, 3次の多項式 $S(z),Q(z)$ で, 方程式 $Q(z)=0$ が重根 β をもち, かつ β と $S(z)=0$ の3根 $\gamma<\delta<\varepsilon$ との間に, $\gamma<\beta<\delta$ 又は $\delta<\beta<\varepsilon$ がなり立ち, さらに, $\frac{S(z)}{Q(z)}$ を $\frac{1}{z}$ の昇べき順に形式的展開をしたとき $\frac{1}{z^4}$ の項までが $J(z)$ と一致するようなものが存在するかどうかを調べる. すると

$$\frac{S(z)}{Q(z)}=1-\frac{1}{\lambda}\frac{2\beta z+3(\varphi^2-\beta^2)}{(z-\beta)^2(2\beta z+3\varphi^2+\beta^2)},\quad \varphi^2=1-\frac{1}{6\lambda}$$

の形になることがわかる. 今述べた β と, $\gamma,\delta,\varepsilon$ の間の大小関係がなり立つためには,

$$-\varphi\sqrt{3}<\beta<\varphi\sqrt{3}$$

でなければならないこともわかる. そこで, 次のような分布函数 $E(x)$ を構成する. $0<\beta<\varphi\sqrt{3}$ なる β に対しては, α を $Q(z)=0$ のもう一つの根 $\alpha=-\frac{3\varphi^2+\beta^2}{2\beta}$ とするとき,

$$E(x)=\begin{cases}0 & x\leqq\gamma\\ \frac{\lambda}{2}(x-\gamma)^2 & \gamma\leqq x\leqq\alpha\\ \lambda(\alpha-\gamma)x-\frac{1}{2}\lambda(\alpha^2-\gamma^2) & \alpha\leqq x\leqq\delta\\ 1-\lambda(\varepsilon-\beta)^2+\frac{1}{2}\lambda(x+\varepsilon-2\beta)^2 & \delta\leqq x\leqq\beta\\ 1-\frac{\lambda}{2}(x-\varepsilon)^2 & \beta\leqq x\leqq\varepsilon\\ 1 & \varepsilon\leqq x\end{cases}$$

とする. $-\varphi\sqrt{3}<\beta<0$ のときは, これと対称な関係になるように作る. さて β が可能な全範囲を動くとき, $E(x)$ のグラフは, ある帯状領域を掃過するであろう（ある部分は $E(x)$ の包絡線であり, ある部分は $E(x)$ のグラフの “かど” の軌跡となる). この領域の下側のふちが $L(x)$, 上側のふちが $U(x)$ を与える. 計算はあまり簡単ではないが, 数値的な結果は出てくる. $|x|$ が充分大のときの $L(x)$ の漸近展開は,

$$L(x)=1-\frac{4\varphi^2}{9x^2}+\frac{4\varphi^2(8-9\varphi^2)}{27x^4}-\frac{4\varphi^2(256-600\varphi^2+327\varphi^4)}{243x^6}+\cdots\cdots$$

となり, 前述 $(2,1,\infty)$ の場合とくらべると漸近的には $\frac{4}{9}$ が $\frac{4}{9}\varphi^2\left(\varphi^2=1-\frac{1}{6\lambda}\right)$ となった分だけ能率がよくなっている.

Mallows は, なお, $(2n,0,\lambda)$, $(2n,1,\infty)$, $(2n,2,\infty)$ 等について, 一般的な手順を与えているが, それらをいちいち引用することは, この報告の Capacity を越えるので, このくらいにしておく.

さて, Mallows [29] 以前の結果のうちには, 上の諸結果に含まれてしまうものも多いのであるが, 文献上の独立した価値を考えて, 一応独立の項目を設けて引用することにした.

(3.2) ν_r が与えられ, かつ分布函数の変動の仕方にある種の条件をつけたもの. von Mises [31] の結果は, 次のようなものである. r 次の絶対モーメント ν_r が与えられた場合, (1.4) における Wald と同様に, 非負確率変数に変換できるから, $x\geqq0$ なる範囲の分布として $P([0,x])$ を考える.

いま, $x_0<x_1$ なるある x_0 に対して, 分布函数 $F(x)$ のグラフが, $x<x_0$ の範囲では, 点 $(x_1,F(x_1))$ を通るある直線より下にあるという条件を満たすならば,

$$F(x_1)\geqq1-\frac{\nu_r}{\zeta^r},$$

ここに ζ は, $g(x)=x^{r+1}$ に対して $g'(\zeta)=\frac{g(\zeta)-g(x_0)}{\zeta-x_1}$ を満足する実数, すなわち $(r+1)\zeta^r(\zeta-x_1)=\zeta^{r+1}-x_0{}^{r+1}$ を満足するものである.

更に，τ を $\dfrac{g(\tau)-g(x_0)}{\tau-x_0}=(r+1)\nu_r$ を満足するように定めると，もし $\nu_r>x_0^r$, $\zeta<\tau$ ならば

$$F(x_1)\geqq\frac{x_1-x_0}{\tau-x_0}.$$

von Mises の条件が満足されている場合の例は，$f(x)=F'(x)$ が存在して $x>x_0$ で単調減少（すなわち，$F(x)$ が $x>x_0$ で凹函数）のときである．

(3.3) 次に，$F(x)$ の高階導函数についての条件（ある区間での $F(x)$ の強い意味の単調性）があるときについての van Dantzig [13] の結果は次の通り．X を非負確率変数，$F(x)$ をその分布函数とし，$R(x)=1-F(x-0)$ とおく．$0\leqq a<b<\infty$ なる区間 $[a,b]$ に x が属するとき，$(-1)^jR^{(j)}(x)\geqq0\ (0\leqq j\leqq h)$ であって，$j=h$ のときには左辺は非増加であるとする．

$h^*=\left[\dfrac{(b-x_0)F'(x_0)}{P([x_0,b])}\right]$（[ ] は Gauss の記号）とおき，更に，$H=\min(h,h^*)$, $r=x_0+\dfrac{F'(x_0)}{HP([x_0,b])}$,

$$\alpha_k'=\int_a^b(x^k-a_k)dF(x),\quad \Gamma_{H,k}(\alpha)=\max_{0\leqq\rho\leqq1}\frac{\rho^k(1-\rho)^H}{k\int_{\alpha\rho}^1\xi^{k-1}(1-\xi)^Hd\xi}$$

とおく．すると

$$F(x_0)\geqq F(b)-\Gamma_{H,k}\left(\frac{\alpha}{r}\right)\alpha_k'/x_0^k$$

である．

上の不等式は，$F'(x_0)$ などを含むので，実用的な面で不利である．そこで，これよりあらいが $F'(x_0)$ を含まない不等式は，

(3.4)

$$F(x_0)\geqq F(b)-\left\{\Phi_{H,k}\left(\frac{a}{x_0},\int_a^bx^kdF(x)\Big/x_0^kP([a,b])\right)\right\}^HP([a,b]),$$

ただし，$\Phi(\alpha,c)$ は，

$$\Phi^H\Big/\gamma_{H,k}-\sum_{j=1}^H(-1)^{j-1}\frac{H!\,k!}{(H-j)!(k+j)!}\times\Phi^{H-j}\left(\frac{1-\Phi}{1\quad\alpha}\right)^jI_{k+1,j}(\alpha)=C,$$

$$I_{l,H}(\alpha)=\sum_{j=1}^{H-1}\binom{l+H-1}{l+j}\alpha^{l+j}(1-\alpha)^{H-1-j},$$

$$\gamma_{H,k}=\Gamma_{H,k}(0)=\frac{(H+k)!\,k^kH^k}{H!\,k!\,(H+k)^{H+k}}$$

を満足するものである．

これと同種のものとして，

(3.5)

$$F(x_0)\geqq F(b)-\gamma_{H,k}\times\int_a^bx^kdF(x)\Big/x_0^k(H\lambda)-P([a,b])\psi\Big/1+\lambda^{-1},$$

ここに，

$$\lambda=\frac{k^kH^H}{(k+H)^{k+H}}\frac{(a/x_0)^{k+H}}{(1-a/x_0)^H},$$

$$\psi=\frac{I_{k+1,H}(a/x_0)}{(a/x_0)^{H+k}}.$$

特に $h=1$ のときは，(3.4) と (3.5) とは同一の結果を与える．

(3.5) で，$h=1$, $k=2r$, $b=\infty$ とおき，$\int_0^ax^kdF(x)$ を無視すれば次の不等式を得る．

(3.6)

$$F(x_0)\geqq1-\theta-\frac{\nu_{2r}}{(1+1/2r)^{2r}(1+\lambda)x_0^{2r}},$$

ただし，

$$\theta=\frac{F(a)}{1+\lambda^{-1}},\quad \lambda=\left(\frac{a}{x_0}\frac{3r}{2r+1}\right)^{2r}\Big/(2r+1)\left(\frac{x_0}{a}-1\right)$$

(3.7), (3.5) で，特に $a=0$, $h=1$ とすれば，

$$F(x_0)\geqq1-\frac{\nu_k}{x_0^k}\left(\frac{k}{k+1}\right)^k.$$

これは，Meidell によって最初に証明された結果であるが，そっくりこれに対応するディスクリートな場合の結果が Meidell [30] に与えられているようである．すなわち，

(3.8) 整数値 i を確率 p_i でとる確率変数 X を考え，区間 $|x|\leqq c\left(1+\dfrac{1}{n}\right)$ において，これらの確率の値は $i=0$ で一つの極大値をとるとすれば

$$P(|X|\leqq c)\leqq1-\frac{\nu_n}{c^n}\left(\frac{n}{n+1}\right)^n$$

(3.7) 及び (3.8) を見ると直ちに，$x=0$ にモードをもつ分布に対するよく知られた不等式

$$P(|X|\leqq d)\geqq1-\frac{4\nu_2}{9d^2}\ \text{が想起される．}$$

実際，これは (3.7) の特別な場合であり，Gauss により早く (1821) 与えられたのと本質的に同じである．

(3.9) 上は，原点にモードのある場合であったが，必ずしも原点と限らないで，モードの場所を指定しない場合には，当然，いくらか能率の低い結果しか得られない．Sellerg [41] によると，上の $\dfrac{4}{9}$ という係数が，それよりやや大きい数値で置きかえられて，

$$P(|X|\leqq x)\geqq1-\frac{\theta\nu_2}{x^2}$$

ただし，θ は方程式 $\theta^3-9\theta^2+3\theta+1=0$ の一根で，$\theta=0.565376\cdots\cdots$ である．この結果は最良ではない．同様な問題について，Selberg [43] で更に論じら

れているようである．

(3.10) 単峰分布に対する (3.7), (3.8) 等の結果は，ただ一個の絶対モーメントを用いるものであった．一次及び二次の絶対モーメントを用いた場合は Royden が与えている．結果は，

$$L(x)=\inf P(|X|\leqq x),\quad U(x)=\sup P(|X|\leqq x)$$

とおくと，

$$U(x)=\begin{cases}1-\dfrac{(2\nu_1-x)^2}{3\nu_2-2x\nu_1}{}^{*)}, & 0\leqq x\leqq 2\nu_1,\\ 1 & 2\nu_1\leqq x\end{cases}$$

$$L(x)=\begin{cases}\dfrac{x}{2\nu_1} & ,\ 0\leqq x\leqq \nu_1,\\ 1-\dfrac{\nu_1}{2x} & ,\ \nu_1\leqq x\leqq\dfrac{3\nu_2}{4\nu_1},\\ 1-\dfrac{4\nu_1^2}{3\nu_2}+\dfrac{8\nu_1^3x}{9\nu_2^2} & ,\ \dfrac{3\nu_2}{4\nu_1}\leqq x\leqq\dfrac{\nu_2}{\nu_1},\\ 1-\dfrac{\Lambda-1}{3t^2-4t+\Lambda} & ,\ \dfrac{\nu_2}{\nu_1}\leqq x,\end{cases}$$

ここに，$\Lambda=\dfrac{3\nu_2}{4\nu_1^2}\ (\geqq 1)$，$x=4\nu_1\dfrac{t^3-t^2}{3t^2-4t+\Lambda}$．

(3.11) 与えられた点にモードをもち，かつ一個のモーメント ν_{2r} がわかっているときの不等式 (Smith [45, 46])．

X が非負確率変数で，c にモードをもつものとすれば，

$$F(x)\geqq 1-\frac{\nu_{2r}-c^{2r}\left(1-\dfrac{2rc}{x(2r+1)}\right)}{(x/\theta)^{2r}-c^{2r}(1-2rc/x(2r+1))}$$

が，$x<\dfrac{c}{F(c)}$ に対してなり立つ．ここに，θ は

$$x=\frac{2r\{x^{2r+1}-(c\theta^{2r+1}\}}{(2r+1)\theta\{x^{2r}-(c\theta)^{2r}\}}$$

によって定義される．

(3.12) 非負確率変数で，密度函数が，原点を端とするある区間で単調な場合 (Narumi [33])．

$f(x)=F'(x)$ が $0\leqq x\leqq b$ で単調非減少であるとすれば，$\nu_n<b^n<(n+1)\nu_n$ のとき，b_1 を

$\dfrac{x(b^n-x^n)}{b-x}=(n+1)\nu_n-b^n$ の正根として，

$b_1\leqq x\leqq b$ に対し，

$$\frac{x}{b}-\frac{b-x}{b^n-x^n}\,\frac{(n+1)\nu_n-b^n}{b}\leqq F(x)\leqq\frac{x}{b},$$

$b\leqq x$ に対し

$$1-\frac{(n+1)\nu_n-b^n}{(n+1)x^n-b^n}\leqq F(x)\leqq 1,$$

また，$b^n=(n+1)\nu_n$ のときは，上の場合の極限として，

$0\leqq x\leqq b$ に対し　$F(x)=\dfrac{x}{b}$

$b\leqq x$ 　に対し　$F(x)=1$.

また，$b^n<(n+1)\nu^n$ という場合は起りえない．

(3.13) 上と逆に，$f(x)$ が $0\leqq x\leqq b$ で非増加の場合（ただし，$b^n>\nu_n$）．

$0\leqq x\leqq\dfrac{nb}{n+1}$ に対し，

$$\frac{(n+1)x}{nb}\left(1-\frac{\nu_n}{b^n}\right)\leqq F(x)\leqq 1.$$

$\dfrac{nb}{n+1}\leqq x\leqq b$ に対し，$1-\dfrac{\nu_n}{b^n}\leqq F(x)\leqq 1.$

$b\leqq x$ 　に対し，$1-\dfrac{\nu_n}{x^n}\leqq F(x)\leqq 1.$

1.4. 統計量の stochastic な性質に条件のある場合．

1.4.1 次に，前節と同じく，モーメントによる条件以外に何らかの条件をつけ加えて，不等式の能率をあげようとする一つの例として，統計量の性質に制限をつける場合がある．たとえば，互いに独立な n 個の統計量の和であるという場合などがそれである．もっとも，統計量に条件をつけるといっても，それは結局，分布函数に対する条件となって反映される（それが前節と異なる点は，分布函数に対しては implicit な条件であるということにすぎない）のであるから，前節と本質的には区別がないといえるのであるが，前節においては分布函数の形に explicit な幾何学的仮定をおいたのに対して，こんどは，統計量の性質に対しては explicit であるが，幾何学的には implicit な条件という点で，実用の便宜の上からも区別するのが妥当であろう．

この場合にも，1.1 とくらべると，ずっと鋭い不等式が得られると考えられる．その理由の一つは，やはり，独立な確率変数の和などはあまり少数個の値のみをとりうる確率変数ではありえず，一方 (1.1) で与えられた不等式において等号に達するのは，多くの場合，少数個の値のみをとりうる分布であるという点である．

得られている結果の多くは，独立確率変数列の和または相加平均の場合についてであるから，特に，大数の法則，重複対数の法則，中心極限定理などとの関連において考察すると興味深い．大数の弱法則や中心極限定理と同じ次元に立って見たものが，Guttman, Robbins, Bernstein, Hoeffding 等の結果であり，大数の強法則や重複対数の法則と共通の観点に立つのが，Kolmogoroff や Hájek & Rényi の結果である．

1.4.2 特に断わらない限り，"確率変数の和" という時には，独立な確率変数の場合とし，μ_r 等は個々

---

*) Royden は，原論文で，誤って $3\nu_2-2x\nu_1$ の 2 を落している．また，Godwin [14] は Royden の結果を引用する際，$P(|X|\leqq x)$ としないで $P(0\leqq X\leqq x)$ としているのは，引用の誤りであろう．

の変数に対応するモーメントとする.

まず，最も有名なものの一つは，

(4.1) Guttman [16] の不等式. $X_1, X_2, \cdots, X_n$ は同一分布（平均値 μ, 分散 μ_2）に従うとし，

$\bar{X}=\frac{1}{n}(X_1+\cdots X_n)$ とおけば，

$$P\left(|\bar{X}-\mu|\leqq\left\{\frac{s^2}{n-1}+\mu_2\sqrt{\frac{2(\lambda^2-1)}{n(n-1)}}\right\}^{\frac{1}{2}}\right)\geqq 1-\frac{1}{\lambda^2}.$$

ここに，$s^2=\frac{1}{n}\sum(X_i-\bar{X})^2$. 確率を考える区間が，ランダムな要素をもっていること，すなわち，標本値に依存することがその特徴である.

(4.2) 平均値と分散が与えられたとき，平均値 μ を中心とする区間に対する不等式 (Robbins [37]), $\mu=0$, $\mu_2=\frac{1}{n}$ なるあらゆる分布に対する$P\left(|\bar{X}|\leqq\frac{t}{n}\right)$ の下限を $1-\phi_n(t)$ とすれば，$n>1$, $t>\sqrt{n}$ に対し，$\phi_n(t)<\frac{1}{t^2}$ が成り立つ. しかも，n を任意に固定して $t\to\infty$ ならしめたとき，$t^2\phi_n(t)\to 1$ である.

(4.3) モーメントに対するある種の制約のもとでの，独立変数の和の分布に対する不等式 (Bernstein [3]). X_i のモーメントを $\{\mu_r^{(i)}\}$ とするとき，$|\mu_r^{(i)}|\leqq\frac{1}{2}\times H^{r-2}r!$ (H は定数) がすべての i と $r\geqq 2$ のすべてに対して満足されているとする. $\mu^{(i)}=0$, $B_n=\sum_{i=1}^{n}\mu_2^{(i)}$ とすれば

$$P\left(|\bar{X}|\leqq\frac{\lambda}{n}\right)\geqq 1-2e^{\frac{-\lambda^2}{2B_n+2H\lambda}}$$

特に，各 X_i が有界変数 ($|X_i|\leqq M$) ならば，上で $H=\frac{M}{3}$ にとることができる.

さらに，独立でない確率変数の和の場合にも，逐次の条件つき確率を考えることにより，同様な結果が得られている (Bernstein [4]).

(4.4) 各変数の $\mu_2^{(i)}$ と $\nu_3^{(i)}$ に対する条件のもとにおける，長さだけが与えられた一般の区間に対する不等式 (Offord [34]). 各 X_i に対応する $\mu_2^{(i)}$, $\nu_3^{(i)}$ のすべてが，

$$\frac{\sqrt{\mu_2^{(i)}}}{\sqrt[3]{\nu_3^{(i)}}}\geqq 2\sqrt[3]{k}$$

を満足するものとすれば，

$$\sup_t P\left(|\bar{X}-t|\leqq\frac{x}{n}\right)\leqq\frac{6\log n}{k^3\sqrt{n}}\left(\log n+\frac{kx}{\min_i\sqrt{\mu_2^{(i)}}}\right),$$

(4.5) 正規分布との差の評価. 上において，n が大きいとき，$\bar{X}$ の分布は漸近的に正規分布に従うから，その漸近度を知りたくなる. Berry [5] は，次の評価式を与えた.

$$\left|P\left(\bar{X}\leqq\frac{\lambda\sqrt{B_n}}{n}\right)-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{k}e^{-\frac{t^2}{2}}dt\right|<\frac{1.88}{\sqrt{B_n}}\max_i\frac{\nu_3^{(i)}}{\mu_2^{(i)}},$$

ただし，$\mu^{(i)}=0$ と仮定し，$B_n=\sum_{i=1}^{n}\mu_2^{(i)}$ とおいた. Berry の論文には後に指摘されているように，誤りがあるようであるが，それが修正された上でも，1.88 という定数は正しいとのことである.

(4.6) 上と同じ問題に対するもう一つの結果. (Bergström [2]).

$$\left|P\left(\bar{X}\leqq\frac{\lambda\sqrt{B_n}}{n}\right)-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{k}e^{-\frac{t^2}{2}}dt\right|<\sum_{i=1}^{n}\frac{\nu_3^{(i)}}{B_n^{\frac{3}{2}}}.$$

この両者は，ある場合には前者の方が精度がよく，又別の場合には逆であるから，どちらが精度がよいと一概にはいえない.

(4.7) $X_1,\cdots,X_n$ が同一の対称かつ単峰で変域 $[-a,a]$ をもつ分布に従う場合 (Birnbaum [7]).

$$\Psi_n(t)=\frac{2}{n}\sum_{\frac{n(t+1)}{2}<k\leq n}(-1)^k\binom{n}{k}\left\{\frac{n}{2}(t+1)-k\right\}^n$$

とおくと，

$$P(|\bar{X}|\leqq x)\geqq 1-\Psi_n\left(\frac{x}{a}\right).$$

(4.8) Hoeffding [18] は，n 個の独立変数にもとづく統計量の期待値について次のような一般的な考察を試みている. 各 X_i の分布函数を $F_i(x)$ とし，$g_{ij}(x)$ ($1\leqq i\leqq k$, $1\leqq j\leqq n$) を与えられた函数，c_{ij} を与えられた実数とする. さらに，n 変数の函数 $K(x_1,\cdots,x_n)$ を，期待値の存在するような函数とする. $F(x_1,\cdots,x_n)=F_1(x_1)\cdots F_n(x_n)$ によって n 次元空間の上の分布が定義されるが，

$$\int g_{ij}dF_j=c_{ij},\quad 1\leqq i\leqq k,\quad 1\leqq j\leqq n$$

なる F_j を成分にもつような F の集合を C とするとき，$\sup_{F\in C}E[K(x)]$ 及び $\inf_{F\in C}E[K(x)]$ を求めるには，適当な正則条件下においては，各 F_i が階段函数のときを考えればよいということが示される. これを使って，Hoeffding & Schrikhande [19] は，k 個のモーメントがわかっている同一分布に従う n 個の独立変数の和の分布函数の上限及び下限について，次の結果を与えている. すなわち，$n=2$ に対しては，高々 $2k+2$ 個の値をとるディスクリートな分布によって上限又は下限は達せられるか，またはいくらでも近づくことができる. 平均値の与えられた非負確率変数について，$n=2$ のときに結果が与えてある.

(4.9) 部分和の変動をも考える場合. いままでのはすべて，ある n に対する横断的な場合を扱っている

が，過程的——すなわち，部分和の変動をも考えにいれる場合（大数の強法則や，重複対数の法則と，ある意味で同じ立場である）について最も知られているのは Kolmogoroff [23, 24] の不等式であろう．

$\mu^{(i)}=0$ とし，$S_r=X_1+\cdots+X_r$ $(r=1,\cdots,n)$ とおけば，$E(S_r)=0$, $B_n=E(S_n^2)=\sum_{i=1}^{n}\mu_2^{(i)}$ である．このとき，$S_1,\cdots,S_n$ の少くとも一つが絶対値において R を越える確率は，

$$P(\mathrm{Max}(|S_1|,\cdots,|S_n|)\geqq R)\leqq\frac{B_n}{R^2}.$$

さらに，各 X_i が有界変域（すべてのについて $|X_i|\leqq M$ とする）のときには，m を整数，$R>M$ として

$$P(\mathrm{Max}(|S_1|,\cdots,|S_n|)\geqq mR)\leqq\frac{D}{(R-M)^{2m}}.$$

また，同じく有界変域のとき，$S_1,\cdots,S_n$ が一様に R をこえない確率は，

$$P(\max(|S_1|,\cdots,|S_n|)\leqq R)\leqq\frac{4(M+R)^2}{D^2}$$

で与えられる．

(4.10) Kolmogoroff の結果の拡張．各 S_i に対して R が異なる場合への拡張を Hájek & Rényi [17] が与えている．確率変数列 $\{X_i\}$ $(i=1,2,\cdots)$ に対し，$\mu^{(i)}=0$ を仮定し，$\{c_k\}(k=1,2,\cdots)$ を非増加な正数の数列とすれば，任意の $\varepsilon>0$ と，任意の二整数 $n, m\,(n<m)$ に対し，

$$P(\max_{n\leqq k\leqq m}c_k|S_k|\geqq\varepsilon)\leqq\frac{1}{\varepsilon^2}\times\left(c_n^2\sum_{k=1}^{n}\mu_2^{(k)}+\sum_{k=n+1}^{m}c_k^2\mu_2^{(k)}\right).$$

が成り立つ．

明らかに，$n=1$, $c_k=1$ $(k=1,2,\cdots)$ とすれば，Kolmogoroff の結果に帰する．

1.5 多次元の場合．

1.5.1 今まで述べたのはほとんど一次元の分布に関する事柄であった．これらを多次元の場合に拡張することは，数学的興味の上からも，実用的見地からも必要である．

高次のモーメントを用いるものは，一次元の場合とくらべてはるかに複雑で困難となるので，具体的な結果が得られているものの多くは，二次までのモーメントを用いるものである．一次元とくらべて問題になる点をいくつか指摘しておこう．

まず，領域の設定のことである．確率を考える領域は，一次元の場合には，区間または区間の合併が極めて自然であり，実用の面からもそれでほぼ充分であった．ところが一次元における区間は多次元においては何に相当すると考えるのが妥当であろうか．第一に考えられるのは，多次元でも区間，すなわち，座標面に平行な面で囲まれた領域を考えることであろう．しかし一方，多次元における楕円体も，一次元での区間に相当する．このことから，直方体でなく，楕円体を考えるということもまた妥当と考えられる．実際，今までに得られている結果を見ると（おもに2次元であるが）長方形領域と楕円領域とがどちらも盛んに扱われて，いろいろの結果が生み出されている．どちらを選ぶべきかということは，実用の面から考えてきめるべきであろう．数学的な面から見れば，座標軸は分布に対して何ら固有の意味をもつものではなく，むしろ，座標軸と無関係にきまる（いい換えれば，座標変換で不変の）ある種の楕円領域などの方が，秩序立ったきれいな結果を与えるのではないかと考えられる．一方長方形領域は，領域設定の簡易さと，標本がこの領域内にあるかどうかの判別の容易さなどに利点をもっている．

次に注意すべき点は，多次元の変数は，複数の変数を成分としてもつのであるから，単なるモーメントばかりでなく，変数の間の確率的な関係について何らかの前提（たとえば独立性など）をおくことのできる場合が実際問題には多いということである．

以下に紹介する諸結果のうち，(5.1)—(5.5) は長方形領域に関するもの，(5.6)—(5.9) は楕円領域に関するものであり，(5.9) と (5.10) は各成分間の独立性を仮定している．

1.5.2

以下各変数の平均値は特に断わらない限り0とし，

$$\mu_{r\cdots t}=E(X_1^r\cdots X_n^t)$$

と書くことにする．特に $n=2$ のときは，$\mu_{20}=\sigma_1^2$, $\mu_{02}=\sigma_2^2$, $\mu_{11}=\rho\sigma_1\sigma_2$ とかくのが便利である．

(5.1) 長方形領域に関したもので古いものは，Berge [1] によって与えられた次の不等式であろう．

$$P(|X_1|<k\sigma_1,|X_2|<k\sigma_2)\geqq1-\frac{1+\sqrt{1-\rho^2}}{k^2}$$

しかも，これは最良不等式である．

(5.2) 上記の結果を，長方形の二辺の長さの比が $\sigma_1:\sigma_2$ に等しくない場合に拡張して Lal [25] の与えたのは，

$$P(|X_1|<k_1\sigma_1,|X_2|<k_2\sigma_2)\geqq1-K_1,$$

ここに，

$$K_1=\frac{1}{2k_1^2k_2^2}(k_1^2+k_2^2+\sqrt{(k_1^2+k_2^2)^2-4\rho^2k_1^2k_2^2}$$

これも最良不等式である．

(5.3) Lal は更に高次元への拡張を試み，

$$P(|X_i|<k\sigma_i,\ i=1,2,\cdots,n)\geqq1-K_2$$

ただし，$K_2=\dfrac{n\{1-(n-1)rt_1\}}{\{1-(n-1)t_1\}\{1+t_1\}^{n-1}},$

$$t_1 = \frac{(l+m)-\sqrt{l(l+2m)}}{m(n-1)},$$

$l = n(1-r)$, $m = (n-1)r$,

r は $X_1, \cdots, X_n$ のうちの任意の二つの変数の相関係数.

また, k が同じでないときにも,

$$P(|X_i|<k_i\sigma_i, i=1, \cdots, n) \geqq 1-K_3,$$

ただし, $K_3 = \dfrac{A-2tB}{(1-\overline{n-1}\,t)(1+t)^{n-1}}$,

$$A=\sum_{i=1}^{n}\frac{1}{k_i^2},\quad B=\sum_{i,j}\frac{r_{ij}}{k_ik_j},$$

(r_{ij}) は分散行列.

という結果を出している. t は任意であるから, K_3 を最小ならしめるためには, t として, $2(n-1)^2Bt^2-\{n(n-1)A+2(n-2)B\}t+2B=0$ の根をとればよい.

ところで, Lal は, 此の結果を最良不等式であると主張しているが, $n>2$ のときにはそれは誤りである.

(5.4). Olkin & Pratt [35] は, Lal の結果が最良不等式でないことを指摘して, 実際に上記の確率の下限はどのように与えられるかを研究している. その結果,

$$P(|X_i|<k_i, i=1, \cdots, n) \geqq 1-\min(1, \mathrm{tr}\, VB^{-1}).$$

ただし, V は分散行列, B は正の定符号対称行列で, かつ B の対角線要素は,

$$b_{ii} = k_i^2 \ (i=1, \cdots, n),$$

となるようなもの, t_r は行列のシュプールを表わす. 特に, 最良不等式となるのは $B^{-1}VB^{-1}$ が対角線要素以外は0となるときであり, 等号は, 高々 $2n+1$ の点から成るディスクリート分布によって達せられる. そのような B は実は一意にきまるが, 実際にそれを求めるには, 厄介な行列方程式をとかねばならないので, 実用的見地からは, 多少効率は落ちるが, Lal の結果も便利であろうと思われる. (5.1)—(5.4) において用いられた方法は, 基本的には (5.1) における Berge の方法を踏襲して拡張したものである. すなわち, 正値二次形式で, 考えている長方形 (または, n 次元直方体) の補集合の定義函数より小でないものを考え, その期待値をとることによって不等式を得る方法である.

(5.5) 以上の諸結果は, 平均値 (原点) を中心とする長方形の場合であったが, 中心が原点をはずれている長方形に対しては, 筆者 [21] の結果がある. 両軸に平行な辺をもち, 対角線上に原点があり, さらに, x 軸に平行な辺と y 軸に平行な辺の長さの比は $\sigma_1:\sigma_2$ に等しいような長方形 R についてだけ考える.

一般性を失うことなく, $\sigma_1=\sigma_2=1$ とし, 原点を通る対角線上の二頂点 $(-\alpha, -\alpha)$, (β, β) について $\beta \geqq \alpha$ と仮定してさしつかえない. すると,

(i) $\beta-\alpha \geqq \sqrt{2}\lambda$ かつ $2\alpha^2>1-\rho$ ならば

$$P(X\in R) \geqq \frac{1+\rho}{\lambda^2+(1+\rho)},$$

ただし,

$$\lambda = \frac{\sqrt{2}\,\alpha(1+\rho)+\sqrt{2(1-\rho^2)(\alpha^2+\rho)}}{2\alpha^2-(1-\rho)},$$

(ii): (i) の場合でなくて, $\alpha\beta \geqq 1$ かつ $2(\alpha\beta-1)^2 \geqq 2(1-\rho^2)+(1-\rho)(\beta-\alpha)$ ならば,

$$P(X\in R) \geqq \frac{4\alpha\beta-4-\sqrt{16(1-\rho^2)+8(1-\rho)(\beta-\alpha)^2}}{(\alpha+\beta)^2}$$

(iii) それ以外のすべての場合は, trivial な結果

$$P(X\in R) \geqq 0$$

となる. いずれも最良不等式で, 等号は, 高々5点から成るディスクリート分布で達せられる.

原点を中心としないさらに一般の長方形については, 2次元の場合でさえ, 非常に複雑で, 見やすい結果は得られていない. 上のような, かなり特殊な場合でさえも, 原点に中心を置く Berge 等の結果とくらべると, 著しく複雑になっているのであるから, もっと一般の場合などは, はるかに複雑な結果になると思われる. そもそも, 座標軸などというものは, 便宜上のものであって, 平面上の質量分布とは何ら本質的な関係がないものであるから, 座標軸方向の長方形を考えるということにはどうしても無理があるのであって, あまり見通しのよい結果が得られないのは当然であろう. それに比べると, 座標変換によって不変な領域, たとえば集中楕円のようなものを考える方が, ずっと自然であると思われる. そのような場合についても多くの仕事がなされている. すなわち,

(5.6) Pearson [36] によると, $\rho \neq 0$ のとき,

$$P\{(x_1, x_2)\in E\} \geqq 1-\frac{I_s}{\lambda_0^{2s}},$$

ここに E は楕円

$$\frac{\theta_{11}x_1^2}{\sigma_1^2}-\frac{2\theta_{12}x_1x_2}{\sigma_1\sigma_2}+\frac{\theta_{22}x_2^2}{\sigma_2^2}=\lambda_0^2(1-\rho^2)$$

の内部, I_s は,

$$I_s=\iint\frac{1}{1-\rho^2}\left(\theta_{11}\frac{x_1^2}{\sigma_1^2}-2\theta_{12}\frac{x_1x_2}{\sigma_1\sigma_2}+\theta_{22}\frac{x_2^2}{\sigma_2^2}\right)dF(x_1, x_2)$$

である. I_s は, 分布のモーメントを用いて簡単に表わされる. この結果は, 最良であって, 等号は, 楕円の周上とその中心以外では0であるような分布によって達せられる.

(5.7) n 次元楕円体面 $\mu_{20\cdots0}x_1^2+2\mu_{110\cdots0}x_1x_2+\cdots=1$ を E とする. $\lambda E=\{\lambda x; x\in E\}$ とすると,

$$P((X_1, \cdots, X_n)\in\lambda E) \geqq 1-\frac{n}{\lambda^n}$$

が成り立つ (Chapelon [11]).

(5.8) 分布函数のある種の単調性を仮定したとき

(Leser [26]).

座標軸に平行な面をもち，原点を中心とする楕円体

$$\sum\left(\frac{x_j}{\lambda_j\sigma_j}\right)^2=n \qquad (\sigma_j^2=E(x_j^2))$$

の内部を E で表わし，$P((X_1,\cdots,X_n)\in\overline{E})=P$ とおく．σ_0^2, λ_0^2 を，

$$\frac{n}{\sigma_0^2}=\sum\frac{1}{\sigma_i^2},\quad \frac{n}{\lambda_0^2}=\sum\frac{1}{\lambda_i^2}$$

によって定義し，曲面 $\lambda_0^2\sigma_2^2\sum\left(\frac{x_i^2}{\lambda_i\sigma_i^2}\right)=R_0^2$ の上の平均確率密度を $A(R_0)$ とする．$A(R)$ は $R\leqq k\sigma_0\times\sqrt{n}$ に対して R の非増加函数とする．すると，

$k\leqq 1$ ならば，

$\lambda_0\leqq 1$ に対して $\qquad P\geqq 0$,

$\lambda_0\geqq 1$ に対して $\qquad P\geqq 1-\frac{1}{\lambda_0^2}$.

$1\leqq k\leqq\sqrt{1+\frac{2}{n}}$ ならば，

$\lambda_0\leqq\left(\frac{2}{n+2}\right)^{\frac{1}{n}}k$ に対して

$$P\geqq\frac{n+2}{2}\left(1-\frac{1}{k^2}\right)\left(\frac{\lambda_0}{k}\right)^n,$$

$\left(\frac{2}{n+2}\right)^{\frac{1}{n}}k\leqq\lambda_0\leqq k$ に対して $P\geqq 1-\frac{1}{k^2}$,

$k\leqq\lambda_0$ に対して $\qquad P\geqq 1-\frac{1}{\lambda_0^2}$.

$k\geqq\sqrt{1+\frac{2}{n}}$ ならば，

$\lambda_0\leqq\left(\frac{2}{n+2}\right)^{\frac{1}{n}}\left(\frac{n+2}{n}\right)^{\frac{1}{2}}$ に対して $P\geqq\left(\frac{n}{n+2}\right)^{\frac{n}{2}}\lambda_0^n$,

$\left(\frac{2}{n+2}\right)^{\frac{1}{n}}\left(\frac{n+2}{n}\right)^{\frac{1}{2}}\leqq\lambda_0\leqq\left(\frac{2}{n+2}\right)^{\frac{1}{n}}k$ に対して

$$P\geqq 1-\left(\frac{2}{n+2}\right)^{\frac{2}{n}}\frac{1}{\lambda_0^2},$$

$\left(\frac{2}{n+2}\right)^{\frac{1}{n}}k\leqq\lambda_0\leqq k$ に対して $\quad P\geqq 1-\frac{1}{k^2}$,

$\lambda_0\geqq k$ に対して $\qquad P\geqq 1-\frac{1}{\lambda_0^2}$.

(5.9) 独立変数の場合．Birnbaum, Raymond & Zuckerman [6] は，X_1 と Y_1 とが独立の場合に，次の結果を与えている．$\frac{\sigma_1^2}{s^2}\leqq\frac{\sigma_2^2}{t^2}$ なる任意の正数 s, t に対し，

$$P\left(\frac{X_1^2}{s_1^2}+\frac{X_2^2}{t^2}\leqq 1\right)\geqq 1-L(s,t).$$

ここに，

$$L(s,t)=\begin{cases}1\quad; & \frac{\sigma_1^2}{s^2}+\frac{\sigma_2^2}{t^2}\geqq 1\\ \frac{\sigma_1^2}{s^2}+\frac{\sigma_2^2}{t^2}-\frac{\sigma_1^2}{s^2}\frac{1-(\sigma_1^2/s^2+\sigma_2^2/t^2)}{1-\sigma_1^2/s^2}\ ; & \\ \qquad \frac{\sigma_1^2}{s^2}+\frac{\sigma_2^2}{t^2}\leqq 1\leqq\frac{1}{2}\left(\frac{\sigma_1^2}{s^2}+\frac{2\sigma_2^2}{t^2}+\sqrt{\frac{\sigma_1^4}{s^4}+\frac{4\sigma_2^4}{t^4}}\right) & \\ \frac{\sigma_1^2}{s^2}+\frac{\sigma_2^2}{t^2}-\frac{\sigma_1^2\sigma_2^2}{s^2t^2}\ ; & \\ \qquad \frac{1}{2}\left(\frac{\sigma_1^2}{s^2}+\frac{2\sigma_2^2}{t^2}+\sqrt{\frac{\sigma_1^4}{s^4}+\frac{4\sigma_2^2}{t^4}}\right)\leqq 1. & \end{cases}$$

これらは最良不等式である．これは次の一般的な定理の系として直ちに導かれる．

W, Z が互いに独立で，非負値のみをとる確率変数であるとし，$E(W)=\lambda$, $E(Z)=\mu$, $\lambda\leqq\mu$ とすれば，任意の $t>0$ に対して

$$P(W+Z\geqq t)\leqq M(t),$$

ここに

$$M(t)=\begin{cases}1\ , & t\leqq\lambda+\mu\ ,\\ \frac{\lambda+\mu}{t}-\frac{\lambda}{t}\cdot\frac{t-(\lambda+\mu)}{t-\lambda}=\frac{\mu}{t-\lambda}, & \\ \qquad \lambda+\mu\leqq t\leqq\frac{1}{2}(\lambda+2\mu+\sqrt{\lambda^2+4\mu^2}), & \\ \frac{\lambda+\mu}{t}-\frac{\lambda\mu}{t^2}, & \frac{1}{2}(\lambda+2\mu+\sqrt{\lambda^2+4\mu^2})\leqq t\end{cases}$$

これも最良不等式である．

Birnbaum, Raymond & Zuckerman は，これを，まず，W と Z が有限個の点のみをとると仮定して，W も Z も2個の点のみをとる場合に還元できることを示し，実際に2個の点のとき，等号の成立するような分布を構成している．一般の W, Z に対しては，有限個の値をとる場合でいくらでも近似できるからさしつかえない．

(5.10) 独立試行の場合の観測頻度．Romanovski [38] は，n 個の互いに独立な試行を行ったときの観測頻度と真の頻度とのへだたりについて，次の不等式を与えた．起りうる結果は，s 個の排反事象 $E_1,\cdots,E_s$ からなり，$P(E_i)=p_i$, E_i の観測頻度を q_i とすれば，

$$P(\sum(q_i-p_i)^2<\varepsilon^2)>1-\frac{1-\frac{1}{s}}{n\varepsilon^2}.$$

さらに，二組の観測値に対しては

$$P(\sum(q_i'-q_i'')^2<\varepsilon^2)>1-\frac{1}{\varepsilon^2}\left(1-\frac{1}{s}\right)\left(\frac{1}{n'}+\frac{1}{n''}\right),$$

ただし，第一の組の試行数を n', 観測頻度を q' とし，他方のを n'', q'' とした．

1.6 確率過程への拡張.

1.6.1 n 次元の場合に $P(|X_i|\leqq k_i,\ i=1,\cdots,n)$

の下限を評価することから，さらに $n\to\infty$ としたときの結果が示唆されることに着目して，Whittle [48] は，チェビシエフ型不等式の確率過程への拡張を試みている．その際分散行列は当然函数 $v(s,t)=E[X(s)X(t)]$ によって置きかえられる（$E[X(t)]\equiv 0$ としてよい）．従って，問題は，$v(s,t)$ $(0\leqq s,t\leqq 1)$ と函数 $\alpha(t)$，が与えられたとき，$P(|X(t)|\leqq\alpha(t))$ の下限を評価することである．後に §2 でも触れるがこの場合には，見本過程の性質にかなり強い仮定を置かない限り，意味のある結果は得られがたい．実際 Whittle も，具体的に不等式を与える際には，見本過程の微分可能性等のかなり強い条件を課している．

1.6.2 Whittle の結果を簡単に紹介しておこう．

(6.1) Whittle の方法は，n 次元から $n\to\infty$ とした状態にヒントを得て，ある種の双一次汎函数を構成することにより，次の結果を得ている．

函数 $x(t)$，$y(t)$ $(0\leqq t\leqq 1)$ の対称な双一次汎函数 $S(x,y)$ で，$x\neq 0$ なら $S(x,x)>0$ なるものを考える．また，t,s の函数 $B=B(t,s)$ が存在し，$y_s(t)=B(t,s)$ とおくときすべての x に対して $S(x,y_s)=x(s)$ となるとする．すると，$\{X(t)\}(0\leqq t\leqq 1)$ を確率変数族とするとき，

$$P(|X(t)|<B(t,t),0\leqq t\leqq 1)\geqq 1-E[S(X,X)]$$

が成立する．

具体的な例として，$E[X(t)^2]<\infty$ $(0\leqq t\leqq 1)$ かつ，$X'(t)$ が存在して $E[X'(t)^2]<\infty$ $(0\leqq t\leqq 1)$ ならば，

$$S(x,y)=\frac{x(0)y(0)+x(1)y(1)}{2\alpha^2}+\frac{1}{2\theta\alpha^2}\times\int_0^1[\theta^2x(t)y(t)+x'(t)y'(t)]dt$$

とおくと，$B(t,s)=\alpha^2e^{-\theta|s-t|}$ となって，

$$P(|X(t)|<\alpha)\geqq 1-E[S(X,X)].$$

もっと高階の導函数の存在まで仮定すれば，さらに別の結果が得られる．

1.7 その他の結果．

1.7.1 以上，一通りチェビシエフ型不等式の種々の型について述べてみたが，以上のどの分類にも入れなかったものとして，平均範囲 (mean range) を用いるもの二つを補足しておきたい．

1.7.2 Winsten の不等式

(7.1) mean range による不等式 (Winsten [49]).

w_n を，大きさ n の標本に対する mean range とし t を与えられた正数とする．また整数 m を，

$$\sum_{i=1}^{m}\left\{1-\left(\frac{i}{m}\right)^n-\left(1-\frac{i}{m}\right)^n\right\}\leqq\frac{1}{t}<\sum_{i=1}^{m}\left\{1-\left(\frac{i}{m+1}\right)^n-\left(1-\frac{i}{m+1}\right)^n\right\}$$

によって定める．さらに p を

$$\frac{1}{t}=\sum_{i=1}^{m}\left\{1-(ip)^n-(1-ip)^n\right\}$$

なるように定める．そうすると，

$$\sup_x P(x\leqq X\leqq x+tw_n)\geqq p$$

等号は $m+1$ 個の点からなる分布によって成立つ．

(7.2) さらに，密度函数が原点に関して対称で，かつ，原点をモードとするならば，

$$P(|X|\leqq tw_n)\geqq 1-4t\left\{\frac{1+y^{n+1}-(1-y)^{n+1}}{n+1}-y^{n+1}-y(1-y)^n\right\}$$

ただし $\dfrac{1}{2t}=1-(1-y)^n-y^n$.

§ 2. 問題点と今後の方向．

前 § において，大体，今までに知られているおもな結果について述べた．その説明の際にも，折にふれて多少の批判とか，問題となるような点に言及したつもりであるが，言い落したことなどもあるので，多少重複点もあるかと思うが，特に注意すべき点や興味ある点などをここにとりまとめて述べ，あわせて今後の発展の方向なども考えてみたい．

まず，これらを統計の問題に用いることをたてまえとする以上，いうまでもなく実用性ということが重視されなければならない．実際それだからこそ，モーメントだけによる不等式の鈍さという実用上の難点を克服しようとして，種々の条件を分布に課して少しでも不等式を鋭くしようという試みが行われた筈であった．ところが，その結果，皮肉にも，別の実用上の難点が生じることになる．たとえば，1.2 における如く，分布にある種の“滑らかさ”を仮定して得られた結果は，なるほど 1.1 における結果よりもかなり能率の良いものではあるが，実際にそれを適用しようとする際に，はたして今とり扱かおうとする対称がその滑らかさの条件を満足しているかどうかということを判定することは多くの場合あまり容易ではない．もちろん，統計の多くの問題におけると同様に，厳密な意味でその条件が満足されていることはありえないし，またその必要もないであろう．しかし近似的にでもその条件を確かめることができる場合というのは，せいぜい単峰であるとか，対称であるとかいう程度の単純な条件であって，Mallows における滑らかさの条件の如きは，実際にあたって確かめることは極めて難かしいのではなかろうか．要するに，あまりくどくどした条件を知っているくらいなら，極端ないい方をするならば，その分布を実際に知っているのと大差がなくなってしまうおそれがある．数学的に興味のある条件であ

ればあるほど，実用性から遠くなるという矛盾をもつ点,大いに考えさせられる．この意味で Mallows [29] などは，その分岐点にあるのではなかろうか．もちろん，筆者は，これらの仕事がその故に価値をもたないなどとは決して思ってはいない．むしろ純数学的な見地からは大いに興味をそそられる問題なのであって，これから先充分に研究されるべき事柄だと思う．

次に，1.6 で述べた確率過程への拡張のことについて二，三の点を指摘したい．1.6 においてもちょっと触れたが，第一に明らかなことは，有限次元から無限しかも連続の濃度にまで拡張することは非常な質的な飛躍であって，いきおい種々の附帯条件を加えなければ意味のある結果は得られないということである．しかも問題にしている事象が，“すべての t” についての条件を含むのであるから，このことだけでも，その確率過程に，可分性などの条件を置かなければ意味がないであろう．$v(s,t)$ は，過程の 2 時点における横断的な状況を示すものであるから,この他に,たてのもっと強い関連（いいかえれば，見本過程についての性質）を強制するような条件を置くことがどうしても必要になってくる．実際，Whittle の結果も，それを具体的に表現するときには，見本可程の可微分性等の相当強い条件を課している．しかし時間的なつながりにあまり強い条件をおくことは，その確率過程の性質を相当はっきり規定してしまうことになり，ちょうど有限次元において各変数の間に相関係数よりはるかに強い函数的関係を置くのと同様に，一般性からは遠いものとなり，チェビシエフ型不等式本来の目的から離れてしまうおそれがある．この点注意すべきであろう．

また，確率過程の場合には，確率の上限や下限を正確に与えることが非常にむつかしい．ややあらい不等式は得られても最良不等式を得ることは，その確率過程の構造を実際に知っている場合以外は不可能に近い．Whittle の結果も，やはり最良不等式ではなく，“ともかくこれより小さい” といった程度の不等式に過ぎない．

けれども，Whittle の研究は，一つの新しい試みとして大へん興味深いものがあり，これを改良すれば，まだまだ面白い結果が得られるものと思われる．これについて，二三の可能性にふれてみる．

第一に，確率を考える事象にもう少し融通性を与えること，たとえば，“すべての t” の代りに，“ほとんどすべての（ルベック測度の意味で）” とか，あるいはもっとゆるく,“全区間で稠密な t の集合の上で” などとすると，他の条件をゆるめても，よりよい結果が得られる可能性がある．

第二に，確率過程について，単に相関函数ばかりでなく，ある程度の stochastic な性質を仮定すると，どのような結果が出てくるかを考えることも興味がある．

さて，確率過程以外の場合についても，今後に残された仕事の可能性は非常に多い．たとえば，一次元の場合に得られている結果の殆んどすべては，多次元に拡張される余地を残している．殊に，$2n$ 個のモーメントにもとづく分布函数の評価についてのほぼ完成された結果（1.1 参照）も 2 次元以上に対しては殆んど何もやられていない．

一次元の範囲でもいろいろの事が残っている．さしあたっては，Mallows [29] における未証明の予想定理の証明に興味がひかれる．さらに，Mallows の結果をもっと能率のよいものにするための試みとして，たとえば，分布の変域を有界区間にすると，ずっとよい結果が得られるのではないかと考えられる．また，高階導函数の大きさを定数 λ でおさえているが，むしろ実用性の面からいえば，定数の代りに函数 $\lambda(x)$ を用いる方が自然であると思われる．なぜならば，実際には，$\lambda(x)$ は遠くでは充分小さくなると考えるのは多くの場合に妥当であるし，しかもそうすれば，得られる結果の精度もぐっとよくなるのではないかと考えられるからである．

最後に，1.1 で与えた結果にしても，また Mallows の結果にしても，比較的きれいに出来上っているのは分布函数の値の評価，すなわち，区間 $(-\infty, x]$ に対する確率の評価である．これを任意の区間で置きかえた場合は，これにくらべると，あまり統一的な結果が与えられていない．筆者 [21] は，その一つの試みであるが，さらに Mallows の場合についても，任意区間に対する結果を得ることができれば，ますます有益である．$(-\infty, x]$ に対する評価は片側検定にしか用いられないので，任意区間に対する評価は，実際上の要求からも，ぜひ必要なことであろう．

チェビシエフ型の不等式についての興味ある事柄はまだいろいろあると思うが，おもな点は述べたと思うので，このくらいにとどめ，引用した諸結果を § 3 にまとめて一覧表として掲げ，また参考文献のうち重要なものをあげておく．

終りに，この報告を書くことを勧められ，種々便宜をはかられた統計数理研究所林知己夫氏，および文献の複写や整理に協力して下さった統計数理研究所尾崎睿子・蔭山富江両氏に厚く感謝したい．

§ 3. 諸結果一覧表

左の欄の数字は，§2 における小項目（括孤つき）の番号．中央の欄は与えられているモーメントその他の条件．右の欄は，確率を評価すべき領域あるいは評価される対象．$F(x)$ は分布函数を表わす．m_n は 0 のまわりの n 次モーメント，μ は平均値（$=m_1$）．μ_n は μ のまわりの n 次モーメント．ν_n は 0 のまわりの n 次絶対モーメント．非負とあるのは，非負変域の確率変数の分布を示す．

1.1. モーメントによる不等式

(1. 1)	最初の $2n$ 個のモーメント（$m_1, m_2, \cdots, m_{2n}$）	$(-\infty, x]$ （*i.e.* 分布函数の値）
	（例）μ, μ_2	explicit な結果
(1. 2)	$m_1, \cdots, m_{2n}$	任意の閉集合又は開集合．
	（例）μ, μ_2, μ_3, μ_4	μ を中心とする区間
	（例）μ, μ_2	任意の区間
(1. 3)	$m_1, \cdots, m_n$；有界変域 $[a, b]$	区間 $[a, x]$ （x は任意，a は変域の左端）
(1. 4)	任意の絶対モーメント $\nu_i, \cdots, \nu_{in}$	原点を中心とする区間
(1. 4)′	$m_{i1}, \cdots, m_{in}$；非負変域（これは本質的には (1.4) と同等）	$(-\infty, x]$
	（例）ν_n, ν_{2n}	$(-\infty, x)$
(1. 5)	μ, μ_2, μ_4	μ に関して対称の位置にある二つの区間の合併
(1. 6)	μ, μ_2；有界変域	μ を中心とする区間

1.2. 一般の函数の期待値による不等式

(2. 1)	ある種の函数 $g(x)$ の期待値 $E[g(x)]$	有界区間
(2. 2)	二つの函数の期待値	有界区間
(2. 3)	函数列 $\{f_n(x)\}$ の期待値；有界変域	$(-\infty, x]$

1.3. 分布函数の滑らかさに条件のあるときの不等式

(3. 1)	$m_1, \cdots, m_{2n}$；滑らかさの条件 (k, λ)	$(-\infty, x]$ （*i.e.* 分布函数の値）
	（例）$(2, 1, \infty)$ （*i.e.* μ, μ_2；単峰分布）	$(-\infty, x]$
	（例）$(2, 0, \lambda)$ （*i.e.* μ, μ_2；密度函数 $<\lambda$）	$(-\infty, x]$
	（例）$(4, 0, \infty)$ （*i.e.* μ, μ_2, μ_3, μ_4）	$(-\infty, x]$
	（例）$(2, 1, \lambda)$ （*i.e.* μ, μ_2；単峰；密度函数 $<\lambda$）	$(-\infty, x]$
(3. 2)	ν_r；$P([0, x])$ のグラフがある区間である直線より下．	原点を中心とする区間
(3. 3)	$F^{(j)}(x)$ $(0 \leqq j \leqq h)$ の符号；$\int_a^b x^n dF(x)$；非負	$[0, x]$
(3. 4)	(3.3) をゆるめたもの．	$[0, x]$
(3. 5)	同上	$[0, x]$
(3. 6)	ν_{2r}；(a, ∞) における $F'(x)$ の単調性；非負	$[0, x]$
(3. 7)	ν_k；$[0, \infty)$ における $F'(x)$ の単調性；非負	$[0, x]$
(3. 8)	ディスクリート；ν_n；各点の確率の単調性	0 を中心とする区間
(3. 9)	ν_2；単峰（モードの位置指定せず）	0 を中心とする区間
(3.10)	ν_1, ν_2；原点をモードとする単峰	0 を中心とする区間
(3.11)	ν_{2r}；モード指定された単峰；非負	$[0, x]$
(3.12)	ν_n；非負；ある a に対し $[0, a)$ で $F'(x)$ 単調非減少	$[0, x]$
(3.13)	ν_n；非負；ある a に対し $[0, a)$ で $F'(x)$ 単調非増加	$[0, x]$

1.4. 独立変数列に関する不等式

($\mu_n^{(i)}, \nu_n^{(i)}$ は X_i の n 次モーメント（平均値のまわり）及び絶対モーメント．$S_k = X_1 + \cdots + X_k$)

(4. 1)	μ, μ_2; 標本分散; 同一分布	μ を中心とする区間（$\bar{X}$ について）
(4. 2)	μ, μ_2	μ を中心とする区間（$\bar{X}$ について）
(4. 3)	すべてのモーメントの大きさに対する制約	μ を中心とする区間（$\bar{X}$ について）
(4. 4)	$\mu_2^{(i)}, \nu_3^{(i)} (1 \leqq i \leqq n)$. $\sqrt[3]{\mu_2^{(i)}}/\sqrt[3]{\nu_3^{(i)}}$ に対する制約	長さ与えられた区間のすべてについての上限
(4. 5)	$\mu^{(i)}, \mu_2^{(i)}, \nu_3^{(i)}$ $(1 \leqq i \leqq n)$	正規分布函数との差の評価
(4. 6)	同上	同上
(4. 7)	対称; 単峰; 有界変域; 同一分布	μ を中心とする区間（$\bar{X}$ について）
(4. 8)	各 X_i の若干個の函数の期待値 特に $\mu^{(i)}, \cdots, \mu_k^{(i)}$ $(1 \leqq i \leqq n)$	函数 $K(x_1, \cdots, x_n)$ の期待値 $(-\infty, x]$ （$\bar{X}$ について）
(4. 9)	$\sum_{i=1}^{n} \mu_2^{(i)}$	Max $(\lvert S_1 \rvert, \cdots, \lvert S_n \rvert) \geqq \varepsilon$
(4.10)	$\mu_2^{(i)}$ $(1 \leqq i \leqq n)$	Max $(c_1 \lvert S_1 \rvert, \cdots, c_n \lvert S_n \rvert) \geqq \varepsilon$

1.5. 多 次 元 分 布

(平均値はすべて 0, $\mu_{r \cdots t} = E(X_1^r \cdots X_n^t)$. 特に $\mu_{20} = \sigma_1^2$, $\mu_{02} = \sigma_2^2$, $\mu_{11} = \rho\sigma_1\sigma_2$)

(5. 1)	σ_1^2, σ_2^2, ρ	長方形（辺の長さの比が σ_1/σ_2, 中心原点）
(5. 2)	σ_1^2, σ_2^2, ρ	長方形（中心原点）
(5. 3)	分散行列	n 次元区間（中心原点）
(5. 4)	分散行列	n 次元区間（中心原点）
(5. 5)	σ_1^2, σ_2^2, ρ	長方形（中心が原点でないとき）
(5. 6)	σ_1^2, σ_2^2, ρ	ある種の楕円（中心原点）
(5. 7)	分散行列	ある種の n 次元楕円体（中心原点）
(5. 8)	分散行列; 分布函数のある種の単調性	ある種の楕円体
(5. 9)	σ_1^2, σ_2^2, ρ; 独立確率変数	任意の楕円（原点中心）
(5.10)	独立試行列の観測頻度	観測頻度と母集団頻度（または，二組の観測頻度）とのある種の距離の大きさ

1.6. 確率過程の見本過程の上界に対する不等式

(6. 1)	共分散函数	見本過程がある函数で majorize される確率

1.7. 平均範囲を用いる不等式

(7. 1)	大きさ n のサンプルに対する平均範囲	一定の長さの区間に対する確率の上限の下限
(7. 2)	平均範囲；対称；単峰	モードを中心とする区間

統計数理研究所

文　献

本文中に引用したことがらに関係のあるものだけにとどめた．その他のものについては，Savage [40] に多くの文献が載つている．なお，以下のうちには，直接入手できなくて，他の文献から引用したものもあることを諒解されたい．

[1] Berge, P. O. A note on a form of Tchebycheff's theorem for two variables, *Biometrika*, 29 (1937), 405-6.

[2] Bergström, H. On the central limit theorem in the case of not equally distributed random variables, *Skandinavisk Aktuarietidskrift*, 32 (1949), 37-62.

[3] Bernstein, S. Sur une modification de l'inégalité de Thchebichef, *Ann. Sc. Instit. Sav. Ukraine*, Sect. Math. 1 (1924). (Russian, French Summary).

[4] Bernstein, S. Sur quelques modifications de l'inégalité de Tchebycheff, *Comptes Rendus (Doklady) de l'Académie Des Sciences De*

l'URSS, 17 (1937), 279-82.

[5] Berry, A. C. The accuracy of the Gaussian approximation to the sum of independent variates, *Transactions of the American Mathematical Society*, 49 (1941), 122-36.

[6] Birnbaum, Z. W., Raymond J. and Zuckerman, H. S. A generalization of Tchebyshev's inequality to two dimensions, *Annals of Mathematical Statistics*, 18 (1947), 70-9.

[7] Birnbaum, Z. W. On random variables with comparable peakedness, *Annals of Mathematical Statistics*, 19 (1948), 76-81.

[8] Blum, J. R. A note on estimating distribution functions, *Proceedings of the American Mathematical Society*, 6 (1955), 953-7.

[9] Camp, B. H. Generalization to N dimensions of inequalities of the Tchebycheff type, *Annals of Mathematical Statistics*, 19 (1948), 568-74.

[10] Cantelli, F.P. Intorno ad uh teorema fundamentale della teoria del rischio, *Bolletino dell' Associazione degl. Attuari Italiani*, Milan, (1910), 1-23.

[11] Chapelon, J.M. Sur l'inégalité fondamentale du calcul des probabilités, *Bulletin de la Société Mathématique de France*, Paris, 65 (1937), 100-8.

[12] Craig, C. C. On the Tchebycheff inequality of Bernstein, *Annals of Mathematical Statistics*, 4 (1933), 94-102.

[13] Van Dantzig, D. Une nouvelle généralisation de l'inégalité de Bienaymé, *Annales de l'Institut Heuri Poincaré*, 12 (1951) 31-43.

[14] Godwin, H. J. On generalization of Tchebychef's inequality, *Journal of American Statistical Association*, 50 (1955), 923-45.

[15] Guttman, L. An inequality for kurtosis, *Annals of Mathematical Statistics*, 19 (1948), 277-8.

[16] Guttman, L. A distribution-free confidence interval for the mean, *Annals of Mathematical Statistics*, 19 (1948), 410-3.

[17] Hájek, J. & Rényi, A. Generalization of an inequality of Kolmogorov, *Acta Math. Acad. Sci. Hungar.*, 6 (1955), 281-3.

[18] Hoeffding, W. The extreme of the expected value of a function of independent random variables, *Annals of Mathematical Statistics*, 26 (1955), 268-75.

[19] Hoeffding, W. & Shrikhande, S. S. Bounds for the distribution function of a sum of independent, identically distributed random variables, *Annals of Mathematical Statistics*, 26 (1955), 439-49.

[20] Hsu, P. L. The approximate distributions of the mean and variance of a sample of independent variables, *Annals of Mathematical Statistics*, 16 (1945), 1-29.

[21] Isii, K. On a method for generalizations of Tchebycheff's inequality, *Annals of the Institute of Statistical Mathematics*, 10(1959), 65-88.

[22] Khamis, S. H. On the reduced moment problem, *Aanals of Mathematical Statistics*, 25 (1954), 113-22.

[23] Kolmogoroff, A. N. Über die Summen durch den Zufall bestimmter unabhängiger Grössen, *Mathematische Annalen*, 99 (1928), 309-19.

[24] Kolmogoroff, A. N. Bemerkungen zu meiner Arbeit 'Über die Summen Zufalligen Grössen', *Mathematische Annalen*, 102 (1929), 484-8.

[25] Lal, D. N. A note on a form of Tchebycheff's inequality for two or more variables, *Sankhyā*, 15 (1955), 317-20.

[26] Leser, C. E. V. Inequalities for multivariate frequency distributions, *Biometrika*, 32 (1942), 284-93.

[27] Lurquin, C. Sur une inégalité fondamentale de probabilité, *Comptes Rendus*, Paris, 187 (1928), 868-70.

[28] Lurquin, C. Sur un théoreme de limite pour la probabilité au sens be Bienaymé-Tchebycheff, *Bulleten Acad. Bruxelles*, (5) 14 (1928), 641-58.

[29] Mallows, C. L. Generalizations of Tchebycheff's inequalties. *Journal of the Royal Statistical Society*, Ser. B. 18 (1956), 139-68; discussion 168-76.

[30] Meidell, B. Sur la probabilité des erreurs, *Comptes Rendus*, Paris 176 (1923), 280-2.

[31] Von Mises, R. Sur inégalité pour les moments d'une distribution quasiconvexe, *Bulletin des Sciences Mathématiques* (2), 62 (1938), 68-71.

[32] Von Mises, R. The limits of a distribution function if two expected values are given, *Annals of Mathematical Statistics*, 10 (1939), 99-104.

[33] Narumi, S. On further inequalities with possible application to problems in the theory of probability, *Biometrika*, 15 (1923), 245-53.

[34] Offord, A. C. An inequality for sums of independent random variables, *Proceedings of the London Mathematical Society* (2), 48 (1945), 467-77.

[35] Olkin, I. & Pratt, J. W. A multivariate Tchebycheff inequality, *Annals of Mathematical Statistics*, 29 (1958), 201-11.

[36] Pearson, K. On generalized Tchebycheff theorems in the mathematical theory of statistics, *Biometrika*, 12 (1919), 284-96.

[37] Robbins, H. Some remarks on the inequality of Tchebychef, *Courant Anniversary Volume*, New York: Interscience Publishers, Inc.,

(1948), 345–50.

[38] Romanovski, V.I. On inductive conclusions in statistics, *Comptes Rendus (Doklady) de l'Académie des Sciences de l'URSS* (N.S.), 27 (1940), 419–21.

[39] Royden, H. L. Bounds on a distribution function when its first *n* moments are given, *Annals of Mathematical Statistics*, 24 (1953), 361–76.

[40] Savage, I. R. Bibliography of Nonparametric Statistics and Related Topics, *Journal of the American Statistical Association*, 48 (1953), 844–906.

[41] Selberg, H. L. Uber eine Ungleichung der Mathematischen Statistik, *Skandinavisk Aktuarietidskrift*, 23 (1940), 114–20.

[42] Selberg, H. L. Zwei Ungleichugen zur Ergänzung des Tchebycheffschen Lemmas, *Skandinavisk Aktuarietidskrift*, 23 (1940), 121–5.

[43] Selberg, H. L. On an inequality in mathematical statistics, *Norsk Matematisk Tidsskrift*, 24 (1942), 1–12, (Norwegian).

[44] Shohat, J. A. and Tamarkin, J. D. The problem of moments, *American Mathematical Society*, New York, 1943.

[45] Smith, C. D. On generalized Tchebycheff inequalities in Mathematical Statistics, *American Journal of Mathematics*, 52 (1930), 109–26.

[46] Smith, C. D. On Tchebycheff approximation for decreasing functions, *Annals of Mathematical Statistics*, 10 (1939), 190–2.

[47] Wald, A. Limits of a distribution function determined by absolute moments and inequalities satisfied by absolute moments, *Transactions of the American Mathematical Society*, 46 (1939), 280–306.

[48] Whittle, P. Continuous Generalizations of Tchebichev's inequality, Теория вероятностей и ее применения, 3 (1958), 386–94.

[49] Winsten, C.B. Inequalities in terms of mean range, *Biometrika*, 33 (1946), 283–95.

[50] Zelen, M. Bounds an a distribution function that are functions of moments to order four, *Journal of Research of the National Bureau of Standards*, 53 (1954), 377–81.