統計数理研究所彙報
第32巻 第2号 (1984)

エントロピ・情報・統計

統計数理研究所 松縄 規

(1984年9月 受付)

要 旨

統計の基本的概念に結びつく量，エントロピ，について，その熱力学的背景，Boltzmann-Planck の原理の周辺および Planck の熱放射分布の導出などを通じて検討し，その物理的，統計的意義について考察する．Planck の誘導過程で使われた熱力学的基準を抽出すると現代の統計学にとっても有益である豊富な内容が含まれており，例えば AIC や Rao-Blackwell の定理ともある意味で密接に関係するものがあることを指摘する．Kullback-Leibler 情報量とエントロピとの関係についても若干の検討を行い，統計に於けるエントロピと情報量の意味と役割りについて述べる．

1. 序

統計の理論および応用の分野でエントロピと情報量の利用は殆ど不可欠になりつつある．Boltzmann の仕事を踏まえた Akaike (1973) による情報量基準の発見とその後の理論及び応用面での成功はその端的な現れであろう．しかしそれ以前そして現在でも統計学に於けるエントロピや情報量は，「何故これらの量が統計の広範囲の問題に共通に利用できる自然な量であるのか」という疑問に十分納得のゆく説明を与えることを阻み続けている様に見える．その原因の一つとして，統計学もこれらの量を主として Shannon 流の情報理論に納まる量として導入し，しかも依然としてその影響が続いていることが考えられる．この為，これらの量の理解に対し部分的には公理的，数学的な面が強調される反面，本質的なところであいまいさが残り，それが結局は応用範囲を狭めてきたと言えよう．

上述の疑問に十分説得力を持って答えるには様々な観点からの検討と統一的な考察が必要であり，筆者の力の到底及ぶところではない．そこで本稿では情報量基準の物理的な一側面に限定し，エントロピという量を中心に答えを見出すことを目的とし，その歴史的背景を踏まえながら筆者なりの観点から考察することにする．本稿で扱うエントロピは物理的意味の明確な Clausius の熱力学的エントロピ，Boltzmann のエントロピ，Gibbs のエントロピであり，情報量は Kullback-Leibler のそれである．これらの量の間には後述する様に密接な関係がある．

結論として，情報量基準はエントロピを通じて明確な物理的な裏付けを持っており，統計学という限られた分野からの意味づけに留まらず，より広い分野の共通する現象の数理的解明に有効な原理であるという事の一端が，従来とは違った観点から明らかにされたと言える．

本稿の以下の構成は次の通りである．次節でエントロピのルーツである熱力学的分野の事柄を簡単に述べる．第3節ではその名称と表式が有名であるがその内容が現今の多くの文献で必ずしも明確でない Boltzmann-Planck の原理の意味を考え，更に Boltzmann, Gibbs のエントロピについて Jaynes を参照して考察する．第4節では Planck の熱放射理論の物理的，統計的

意味を検討し，Einsteinの仕事との関連，情報量基準との関係について言及する．又，関連事項として，Rao-Blackwell定理やKolmogorovによる熱伝導方程式に基づく最小分散不偏推定について触れる．最後に第5節でK-L情報量のエントロピを通じての統計的，物理的意味について若干の考察を行う．

尚，エントロピに関する多くの教科書や啓蒙書があるが，それらの中に，先述の本稿の目的に沿うものが筆者の見た限りではなかった。多くは部分的に関連があるけれども総合的に見ると混乱を招く危険の方が強いのでそれらの書物を参考文献として掲げることはしていない．その代りなるべく関係する原典に直接当り，本稿ができる限り自己充足的になる様に，取り上げた話題に則して基礎知識も筆者流の解釈で記述した．熱力学および物理学の関連する用語や基本的事項一般については主に谷下（1981）と池田（1975）に依った．

2. 熱力学的エントロピ

エントロピのルーツはS. Carnot（1824）の仕事にある事はよく知られている．「一様な流体を動作物質としてもち高低二つの熱源の間でのみ熱交換を行う熱機関を準静的変化によって作動させるカルノーの可逆サイクルの熱効率は使用する動作物質の種類によらず両熱源の温度のみの関数として決る」というCarnotの原理の発見は，熱素論の立場で論じられていたにもかかわらずその思考実験の中に革命的な意味が含まれていたことがやがて明らかにされた．Clausius（1850）は1834年のClapeyronの論文によりCarnotの結果を知り，彼等の熱素論からの矛盾をJouleの実験に基づく熱の分子運動説により正し，今日では熱力学の第1および第2法則の名で知られている基本原理に理想気体の場合に到達した．第2法則は翌年W. Thomsonにより，一般の物質系に対しても成立するような形で拡張された．巨視的熱力学に於ける状態量である物体の内部エネルギがClausiusにより，絶対温度目盛がThomsonにより発見された（1848年）ことも，結局はCarnotの仕事がそれらの源泉である．

Clausius（1865）はエントロピ（entropy）と呼ぶ量を熱力学の理論の中へ導入した．これにより当初は言葉のみで述べられていた熱力学第2法則が数量的に表現されるようになった．このエントロピ生誕にまつわる事柄を十分に検討することは重要な意義を持つと思われるが，ここではその余裕がないので直ちに彼の熱力学的エントロピの定義に進むことにしよう：

一つの熱力学系があってそれが状態1にあるものとする．この系が準静的変化（常に熱平衡を保ち途中の状態を可逆に進み得る変化）で状態2へ変化し，この間外部から可逆的（reversibly）に $\delta Q_{\rm rev}$ の熱量を受け取るものとする．このとき量

$$\delta Q_{\rm rev}/T \tag{2.1}$$

が与えられた系に関連する関数 S の完全微分であるような積分因子 $1/T$ が存在するとき，関数 S をこの系のエントロピと名付ける．このとき

$$dS=\delta Q_{\rm rev}/T \quad \text{または} \quad \int_1^2 \frac{\delta Q_{\rm rev}}{T}=S(2)-S(1) \tag{2.2}$$

と表現される．積分分母 T は熱力学的には系の絶対温度を表す．

上の定義は直観的には次の様な背景があってなされたと説明できる．いま気密で摩擦のないピストンを持つシリンダ内に均一の圧力 P，体積 V，内部エネルギ U の理想気体があって，この系に外部から熱量を可逆的に $\delta Q_{\rm rev}$ 加えて準静的変化をさせるものとしよう．気体の内部エネルギ増加が dU，体積の増加を dV とすると，可逆性は熱力学の第1法則（エネルギ保存則）から

$$\delta Q_{\rm rev}=dU+PdV \tag{2.3}$$

と書いて保証される．特に変化が断熱的に行われるとすると

(2.4) $$dU+PdV=0$$

なる微分方程式を得る．ところがこれは微分方程式の初等的理論から完全微分方程式ではない．そこでこれを完全微分方程式にする様な積分因子即ち(2.3)を完全微分とするような因子を探してみよう．(2.4)の一般解が $\mathscr{C}(U, V)=C$(C: 任意定数)と求まったとする．熱力学的には C をパラメータとする断熱曲線が得られたことになる．そこでこの解曲線を全微分すると

(2.5) $$\frac{\partial \mathscr{C}}{\partial U}dU+\frac{\partial \mathscr{C}}{\partial V}dV=0$$

$\partial\mathscr{C}/\partial U=1/\mathscr{I}$ と置くと，(2.4), (2.5)から

(2.6) $$P=-dU/dV=\left(\frac{\partial \mathscr{C}}{\partial U}\right)^{-1}\frac{\partial \mathscr{C}}{\partial V}=\mathscr{I}\frac{\partial \mathscr{C}}{\partial V}$$

これを(2.3)に代入して

(2.7) $$\delta Q_{\mathrm{rev}}=\mathscr{I}\left(\frac{\partial \mathscr{C}}{\partial U}dU+\frac{\partial \mathscr{C}}{\partial V}dV\right)=\mathscr{I}\cdot d\mathscr{C}$$

と表現できる．即ち $\mathscr{I}^{-1}$ は(2.3)の積分因子であり，$\delta Q_{\mathrm{rev}}/\mathscr{I}$ は完全微分となる．数学的に $\mathscr{I}$ は一意に決らないが，熱力学に於ては第2法則から $\mathscr{I}=T$, $\mathscr{C}=S$ が要請される．以上によりエントロピの定義式(2.2)が或る程度自然であることが分るであろう．

式(2.2)の様に S が表現できるということは，物理的には S が始めの状態1と終りの状態2だけで決ってしまう，いわゆる状態量であることを意味している．これはまた次の様に表現できる．系が可逆サイクルをなすとき次の Clausius 積分は

(2.8) $$\oint\frac{\delta Q_{\mathrm{rev}}}{T}=\oint dS=0$$

である．$\oint$ は一サイクルについての積分を表す．不可逆サイクルのときは

(2.9) $$\oint\frac{\delta Q}{T}<0 \qquad \text{(Clausius の不等式)}$$

が成立する．ここに δQ は，温度 T の熱源から系が吸収する熱量である．一方，不等式

(2.10) $$\oint\frac{\delta Q}{T}>0$$

を成立させる様なサイクルは実現不可能である．(2.10)がもともとの熱力学第2法則の一つの数量的表現である．この第2法則から一般に $\delta Q_{\mathrm{rev}}>\delta Q$ であることが証明される．従って(2.2)と組合せて，不可逆変化に対するエントロピ変化は

(2.11) $$dS>\delta Q/T$$

となることが分る．この不等式から孤立系を想定した自然界の熱的変化の方向を知ることができる：断熱変化を考えるから

(2.12) $$dS\geqq 0$$

である．等号は変化が可逆の場合である．もしこの系がエントロピの極大値を持つならば(2.12)により系はこの極大値に対応する状態に向ってエントロピを増大し続けることになる．エントロピが極大に達するとこの系はもはや熱的変化のない熱力学的平衡状態に達する．この古典熱力学に於るエントロピ極大原理が，Boltzmann, Gibbs 等により統計的観点からも大いに使われたことは周知の通りである．

以上簡単に触れた古典熱力学はいくつかの熱力学的基本関係式を基礎にした決定論的な性格を持っている．しかし Carnot の考え出した準静的変化や Clausius の不可逆的変化の尺度としてのエントロピの役割りから，結果論ではあるが，エントロピは統計的に解釈されざるを得な

い運命にあった．その大きな原因は熱の本性が分子の運動であるという認識であって，Carnot 自身が後年熱素論から運動論へと転向をした事実や（cf. 広重，1973），Clausius がエントロピ導入以前（1857）に気体運動論の基礎を論じ，次いで Maxwell が分子の速度分布を与えた事にこの統計的解釈への兆しが見られる．しかし，何と言っても Boltzmann によって熱力学第2法則が統計的内容を持つものであることが明らかにされた事は熱の理論にとって革命的な意義を持つものであった．そこで次節では彼によって基本的に確立された統計的エントロピの原理および関連事項を考察することにする．

3. Boltzmann-Planck の原理，Gibbs-Jaynes エントロピ

熱力学に於て単原子分子の理想気体のエントロピ変化は次の様に表される：気体の圧力，体積，温度，エントロピをそれぞれ P[Pa]，V[m^3]，T[K]，S[J/K] とし，この気体1 kmolが可逆変化する時に外部から Q_{rev}[J] の熱量を供給されるとする（熱力学エントロピが次元を持つ事を考慮して，本節のみ各物理量に国際単位の次元を付け，kmol も便宜上次元と見做す）．この時微小な状態変化に対し気体の熱力学的エントロピの増加は次の様に表現される．

(3.1) $$dS = dQ_{\mathrm{rev}}/T = (C_V dT + PdV)/T \qquad [\mathrm{J/kmol\ K}],$$

ここに C_V は 1 kmol の気体の定容比熱を表す．この理想気体に対して $PV = R_0 T$[J/kmol](R_0 = 8314.3[J/kmol K]：一般ガス定数）であるから

(3.2) $$dS = C_V dT/T + R_0 dV/V \qquad [\mathrm{J/kmol\ K}].$$

右辺の変量はいずれも状態量であり，積分可能である．従って熱力学的エントロピは

(3.3) $$S = \int dS = R_0 \ln(VT^{3/2}) + C_1 \qquad [\mathrm{J/kmol\ K}]$$

と表現できる．但し上で単原子分子の理想気体では比熱は温度によらず一定で $C_V = (3/2)R_0$ であることを使った．C_1 は積分定数である．

Boltzmann（1872）は気体分子運動論の観点から熱力学第2法則をいわゆる H-定理として表現した：

$$\frac{dE}{dt} \leqq 0,\ E(t) = \int\cdots\int f(\boldsymbol{x},\ \boldsymbol{u};\ t)\ln f(\boldsymbol{x},\ \boldsymbol{u};\ t)\boldsymbol{dxdu},$$

ここに $f(\boldsymbol{x},\ \boldsymbol{u};\ t)\boldsymbol{dxdu}$ は時刻 t で座標と速度がそれぞれ $[\boldsymbol{x},\ \boldsymbol{x}+\boldsymbol{dx}]$，$[\boldsymbol{u},\ \boldsymbol{u}+\boldsymbol{du}]$ の範囲にあるような，単位体積中の気体分子の数を表す関数である．これに関連して彼は，単原子分子の理想気体の場合には次の関係が成立することを導いた：

(3.4) $$N_0\left[\ln\left\{V\left(\frac{4\pi e}{3M/N_0}\right)^{3/2}\right\} + \frac{3}{2}\right] = -H_B \qquad [\mathrm{O/kmol}],$$

ここに [O] は無次元単位，V[m^3] は 1[kmol]($=M$[kg]) の理想気体の占める体積，N_0 をその中の全分子数（Avogadoro 数 6.0225×10^{26}[O/kmol]），e[J] を気体分子1個の平均運動エネルギを表す．H_B は当初 $E^*(=E-N_0\ln N_0)$ と記され，後には記号 H を使うようになった．(cf. Boltzmann, 1896)．

そこで熱力学的エントロピ S と Boltzmann の H_B を結びつけよう．単原子分子の理想気体では $e=(3/2)(R_0/N_0)T$[J] だから，(3.4) 式から

(3.5) $$\ln(VT^{3/2}) = -H_B/N_0 - 3/2\cdot\{\ln 2\pi R_0/MN_0 + 1\} \equiv -H_B/N_0 - C_2 \quad [\mathrm{O/kmol}].$$

これと (3.3) とから

(3.6) $$S = -(R_0/N_0)H_B + (C_1 - C_2R_0) \equiv -kH_B + C_3 \qquad [\mathrm{J/kmol\ K}],$$

ここに $k = R_0/N_0 = 1.3806\times10^{-23}$[J/K] でいわゆるボルツマン定数である．上式によりマクロ

な量である熱力学的エントロピ S と，ミクロな観点からの量 H_B が比例関係で結ばれていることが分る．

ところでよく知られているように，Boltzmann (1877) は今日 Boltzmann-Planck の原理として知られている

(3.7) $$S=k\ln W$$

の原形を引き出す考察を行った．W は統計的重率と呼ばれるもので，統計力学でいうミクロカノニカル集合に於てエネルギ範囲が E と $E+\varDelta E(\varDelta E\ll E)$ の間にある微視的状態の数である．Boltzmann 自身は上の表現を明示しなかった．本稿では (3.6) を利用して (3.7) 式に接近する．

Boltzmann (1872) の第1節の仮定のように，始めの時刻に於てあらゆる速度の方向が等確率でかつ速度分布が一様であると仮定する．時刻 t で気体の一粒子の速度 $\boldsymbol{u}=(u_x, u_y, u_z)$[m/s] が区間 $[u_x, u_x+du_x]$, $[u_y, u_y+du_y]$, $[u_z, u_z+du_z]$ の間にある確率を

(3.8) $$f_1(u_x, u_y, u_z;t)du_xdu_ydu_z\equiv f_1dV \qquad [\mathrm{O}]$$

としよう．ここで確率とはその粒子が上の状態にある時間をそれが運動する全時間で割ったものを意味する．

今 1[kmol] の単原子分子の理想気体の中に存在する N_0 個の分子の運動エネルギを L[J] とすると，粒子一個の運動エネルギの各成分は L を越えない．従って粒子の速度 $u=\sqrt{u_x^2+u_y^2+u_z^2}$ に対応する三次元空間の各点は

(3.9) $$V=[2(2LN_0/M)^{1/2}]^3 \qquad [\mathrm{m^3/s^3}]$$

の立方体の中に分布することになる．この立方体を r 個の等しい体積をもつ細胞（セル）に分割するものとし，それらのセルに 1, 2, …, r と番号付けをし，それぞれに $n_1, n_2, \cdots, n_r$ 個の分子（運動点）が配分されるものとする．全ての気体分子が区別可能であるとすれば，先に述べた速度分布の一様性の仮定から，気体分子が上の様にセルに配分される確率は

(3.10) $$P=\frac{N_0!}{n_1!n_2!\cdots n_r!}\left(\frac{1}{r}\right)^{N_0}\equiv W\left(\frac{1}{r}\right)^{N_0}$$

で与えられる．但し $n_1, n_2, \cdots, n_r$ は次の二条件

(3.11) $$n_1+n_2+\cdots+n_r=N_0 \qquad [1/\mathrm{kmol}],$$

(3.12) $$\sum_{i=1}^{r}n_i\frac{1}{2}\frac{M}{N_0}(u_x^2+u_y^2+u_z^2)f_1dV=\frac{3}{2}R_0T \qquad [\mathrm{J/kmol}]$$

を満さねばならない．Stirling の公式を適用して (3.10) から

(3.13) $$\ln P\fallingdotseq N_0\ln N_0-\sum_{i=1}^{r}n_i\ln n_i-N_0\ln r+\mathrm{const.}$$

i 番目のセルに入る気体の分子数は

$$n_i=N_0f_1(u_{xi}, u_{yi}, u_{zi};t)du_{xi}du_{yi}du_{zi}\equiv N_0f_1dV$$

で与えられるから

(3.13)′ $$\ln W\fallingdotseq N_0\ln N_0-N_0\sum_{i=1}^{r}f_1\ln(N_0f_1dV)dV+C_4.$$

r を十分大きく取り，$dV\to 0$ とする．$dV\ln dV\to 0$ に注意すると

(3.14) $$\ln W\fallingdotseq N_0\ln N_0-N_0\int f_1\ln f_1dV+C_5=-H_B+C_5.$$

よって，(3.6) とから

(3.15) $$S_B\equiv k\ln W+(C_3+C_5) \qquad [\mathrm{J/kmol\ K}]$$

を得る．これは定数項を除けば (3.7) と一致する．

ところで (3.15) 式の誘導過程を振返って見ても，$C_3+C_5=0$ とすることは一般には出来ない．これらの定数項は絶対定数でないから，二つの異なる状態に於るエントロピを考えて，そ

の差を取ったとしても消えない．このことは(3.7)式と古典統計力学に沿って得られた(3.15)式の間には大きな飛躍があることを示唆している．実際(3.7)式は統計力学に於るエントロピの定義式とされている．歴史的に見てもこれはPlanckによって初めて採用され，誘導された表現であることは良く知られている．現今の多くの文献で(3.7)式をBoltzmannの原理として取り上げているが，その誘導は殆どなされておらず，あってもかなり便宜的な説明で済まされているように見受ける．この式の意味を理解する為にもPlanck (1930)に倣って証明することは意義があると思うので以下にその概略を述べよう．

次の統計力学に於る仮説から出発する：微視的状態全体はそれぞれに孤立しているものの集団を形成している．物理的対象の巨視的状態の各々はその対象の確定した数 W（統計的重率）の微視的状態を含んでいる．

Planckが指摘したようにBoltzmannの基本的思想の一つは，自然界の現象として我々が認識する巨視的な世界に於て不可逆的変化の尺度を表す熱力学的エントロピ S の本性は，微視的世界に於る統計的重率 W の評価によって理解できると考えたこと，即ち

(3.16) $$S=g(W) \qquad [\mathrm{J/K}]$$

で表現される関数 g を求めようとしたと解釈できる．上で S は連続的な変化が可能な量であるのに対し，W は整数であるから $g(W)$ は不連続にしか変化できない．しかし一般に熱力学的エントロピや温度はわずか数個の分子などに存在しないし，現実に我々が考察する物理対象に対しては W は十分に大きく，上記の不連続性による誤差は全く無視できるものと仮定してよい．

さて，N_1 個および N_2 個の分子からなる二つの独立な物理対象を考えよう．それぞれの熱力学的エントロピを S_1, S_2，統計的重率を W_1, W_2 で表す．この二つをまとめて一つの物理的対象物と見做せるものとすると，その熱力学的エントロピ S_{12} は

(3.17) $$S_1+S_2=S_{12} \qquad [\mathrm{J/K}]$$

を満さねばならない(熱力学的エントロピの示量性による)．一方，対象1の各々の微視状態に対象2の微視的状態の各々が組合せ得るから，両者を合併した後の統計的重率 W は次式の関係を満す：

(3.18) $$W_1\cdot W_2=W.$$

(3.16)～(3.18)から g についての関数方程式

(3.19) $$g(W_1)+g(W_2)=g(W_1W_2)(=g(\mathrm{W})) \qquad [\mathrm{J/K}]$$

を得る．W_1 について偏微分して

$$\frac{\partial g(W_1)}{\partial W_1}=\frac{\partial g(W_1W_2)}{\partial W_1}\cdot W_2=\frac{\partial g(W_1W_2)}{\partial(W_1W_2)}\cdot\frac{\partial(W_1W_2)}{\partial W_1}.$$

更に W_2 について偏微分すると

$$0=\frac{\partial^2 g(W_1W_2)}{\partial(W_1W_2)^2}\cdot\frac{\partial(W_1W_2)}{\partial W_2}W_2+\frac{\partial g(W_1W_2)}{\partial(W_1W_2)}=\frac{\partial^2 g(W_1W_2)}{\partial(W_1W_2)^2}W_1W_2+\frac{\partial g(W_1W_2)}{\partial(W_1W_2)}$$

即ち

(3.20) $$W\frac{d^2g(W)}{dW^2}+\frac{dg(W)}{dW}=0.$$

これを解いて

$$g(W)=a\ln W+b \qquad [\mathrm{J/K}].$$

上式は任意の物理的対象に対して成立すべきであるから，(3.19)とから，$b=0$ でなければならない．従って

(3.21) $$g(W)=a\ln W \qquad [\mathrm{J/K}]$$

となる．定数 a については，古典熱力学の範囲ではその単位は [J/K] で，非常に小さな値を持つ普遍定数であるという事が言えるに過ぎないが，この式を理想気体に適用する事により統計力学の立場では a は実は，Boltzmann 定数 k[J/K] に等しく取られる．以上が (3.7) 式の誘導の大略である．なお，(3.7) による S は常に非負で，物理的対象も理想気体に限る必要はないことも分っている．

ここまで来ると (3.6) 式に対しても，もっと進んだ一般的に成立する表現があるのではないかという期待，そして (3.6), (3.7) 式と熱力学第 2 法則がどの様に結びついているかという疑問が湧いてくる．この事については Jaynes (1965) の指摘が役立つ．その重要な結果は，しかしながら，統計や工科系の文献の中で引用された事は殆どないと思われる．後述の議論にも必要となるので以下に概略を紹介する：(a) Boltzmann のエントロピ S_B (後述) は単原子分子の理想流体に対しては熱力学的エントロピと正しく対応する．しかし (b) 分子間引力を考慮しなければならない一般の場合に於ては Gibbs のエントロピ S_G (後述) が熱力学的エントロピと正しく対応する．従って (c) Boltzmann 原理 (3.7) に於ても，一般に S_G を使うのが正しい．また (d) Boltzmann の H-定理は稀薄気体に対しても熱力学第 2 法則の一般的証明にはならない．これに対して (e) Gibbs の H_G (後述) は時間に関して不変である事が，通常の理解と違って，分子間引力を考慮した場合でも熱力学的第 2 法則の簡単な証明を与える．更に (f) 熱力学第 2 法則は任意の巨視的な物理的対象の時間変化の過程が現象的に可逆であるための一般的条件の内の特別な場合に該当する．以上の事などを Jaynes は指摘した．

(3.6) に関連する (a), (b) は次の様に説明される．N 個の粒子からなる単原子分子の流体を考える．Gibbs のカノニカルアンサンブルを表現する Liouville 関数を $f_N(\boldsymbol{x}, \boldsymbol{p}\,;\, t)$ とする．$\boldsymbol{x}=(x_1, \cdots, x_N)$, $\boldsymbol{p}=(p_1, \cdots, p_N)$ はそれぞれ粒子の位置および広義運動量を表す．f_N が粒子系の全位相空間に於る確率密度関数を意味しているものとする．Gibbs および Boltzmann の H をそれぞれ

(3.22) $$H_G=\int f_N \ln f_N d\tau, \qquad (d\tau=d^3x_1, \cdots d^3p_N),$$

(3.23) $$H_B=N\int f_1 \ln f_1 d\tau_1, \qquad (d\tau_1=d^3x_1 d^3p_1)$$

で定義する．ここに f_1 は一粒子確率密度関数で，f_N の周辺密度関数

$$f_1(x_1, p_1\,;\, t)=\int f_N d\tau_{-1}, \qquad (d\tau_{-1}=d^3x_2\cdots d^3p_N)$$

として与えられる．(3.4) の H_B, (3.8) の f_1 はここで考えているものの特別な場合である．

(3.24) $$H_B \leqq H_G$$

であることが容易に示される．等号は $f_N(\boldsymbol{x}, \boldsymbol{p}\,;\, t)=\prod_{i=1}^{N} f_1(x_i, p_i\,;\, t)$ のとき成立する．

さて，f_N が特に次のカノニカル分布 g_N で与えられる熱平衡状態で熱力学エントロピはどう表現できるかは非常に興味ある問題である．結果に本質的影響はないので，Jaynes は外力の場はないとして g_N として次の表現を用いた：

(3.25) $$g_N=\left(\frac{\beta N}{2\pi M}\right)^{3N/2} Q^{-1} \exp\{-\beta H\}$$

である．ここに $\beta=1/kT \quad$ [1/J],

$$Q=Q(\beta,\ V)=V\int_V \exp(-\beta\phi)d^3x_2\cdots d^3x_N \qquad [\mathrm{O}],$$

$$H=\sum_{i=1}^{N} p_i^2 N/2M+\phi(x_1, \cdots, x_N), \qquad [\mathrm{J}],$$

M, T, V は系内の粒子の総質量 [kg], 温度 [K] および体積 [m³] を，ϕ は位置エネルギ [J] で

粒子座標系に関する対称関数で与えられるものとする．(3.25)より

(3.26) $$g_1 = g_1(x,\ p) = (\beta N/2\pi M)^{3/2} V^{-1} \exp\{-\beta N p^2/2M\}.$$

よって，若干の計算により

(3.27) $$H_G - H_B = \int g_N \ln\{\prod_{i=1}^{N} g_1(x_i,\ p_i)/g_N\} d\tau = \ln Q - N \ln V + \beta\langle\phi\rangle,$$

(3.28) $$d(H_G - H_B) = \beta d\langle\phi\rangle + \beta[\langle P\rangle - P_0] dV,$$

ここに〈・〉はカノニカル集合での平均で

(3.29) $$\langle\phi\rangle = -\partial \ln Q/\partial\beta [\mathrm{J}],\ \beta\langle P\rangle = \partial \ln Q/\partial V \quad [1/\mathrm{m}^3]$$

と表現できる．P および P_0 はそれぞれ Gibbs および Boltzmann の粒子系に対する圧力 [Pa] で，後者は同じ温度と密度を持つ単原子分子の理想気体の圧力 $P_0 = NkT/V$[Pa] に一致する．

次に次式によって Gibbs および Boltzmann のエントロピを定義する：

(3.30) $$S_i = -kH_i\ (i = G,\ B) \qquad [\mathrm{J/K}].$$

S_B は先に与えた(3.6)式と本質的に同じである．これは(3.6)式に於る定数 C_3 は，エントロピの基準を固定すれば，絶対定数と見做せるからである．(3.28)式を可逆な道筋で状態1から状態2まで積分すると

(3.31) $$(S_G - S_B)_2 - (S_G - S_B)_1 = \int_1^2 \{d\langle\phi\rangle + [\langle P\rangle - P_0]\}/T dV \qquad [\mathrm{J/K}].$$

(3.23), (3.30)から

(3.32) $$S_B = Nk\{(3/2)\ln(2\pi MkT/N) + \ln V + 3/2\} \qquad [\mathrm{J/K}].$$

これは(3.4)式に対応するが，そこでは粒子の速度を変数として取り上げたので上式と表現に於て若干の差が出る．(3.32)より

(3.33) $$\left(\frac{\partial S_B}{\partial V}\right)_T dV = \frac{Nk}{V} dV = \frac{P_0 dV}{T} \qquad [\mathrm{J/K}],$$

(3.34) $$\left(\frac{\partial S_B}{\partial T}\right)_V dT = \frac{3}{2} \cdot \frac{Nk}{T} dT \equiv \frac{d\langle K\rangle}{T} \qquad [\mathrm{J/K}],$$

ここに $\langle K\rangle = (3/2)NkT$[J] は粒子系の全運動エネルギを表す．よって(3.31)と同じ可逆な道筋で

(3.35) $$\begin{aligned}(S_B)_2 - (S_B)_1 &= \int_1^2 \left\{\left(\frac{\partial S_B}{\partial V}\right)_T dV + \left(\frac{\partial S_B}{\partial T}\right)_V dT\right\} \\ &= \int_1^2 (P_0 dV + d\langle K\rangle)/T \qquad [\mathrm{J/K}].\end{aligned}$$

この結果と(3.31)から

(3.36) $$\begin{aligned}(S_G)_2 - (S_G)_1 &= \int_1^2 \{\langle P\rangle dV + d\langle K + \phi\rangle\}/T, \\ &= \int_1^2 dQ_{\mathrm{rev}}/T \qquad [\mathrm{J/K}]\end{aligned}$$

が従う．ここに

(3.37) $$dQ_{\mathrm{rev}} = \langle P\rangle dV + d\langle K + \phi\rangle \qquad [\mathrm{J}]$$

は1→2の可逆な状態変化の間の，任意の二つの瞬間の間に分子集団の集りが可逆的に吸収する熱量を表す．

(3.35), (3.36)式が Jaynes の指摘(a)，(b)の数式による表現である．従って(3.6)式を一般的な状況で拡張するには，エントロピおよび H 関数として S_G と H_G をそれぞれ用いなければならない．この事に関連する前述の他の指摘の検討も興味深いが，かなりの準備を要するので省略する．

4. Planck の熱放射分布とエントロピ

今世紀の最初の年に M. Planck（1900 a, 1900 b, 1901）はその当時工業面からの要請とも関連して，空洞放射に於るエネルギ密度の分布を見出す問題で，放射場のエネルギが振動数に対してどのように分布しているかを研究した．その結果，彼は今日 Plank の放射法則と言われるエネルギ分布則を与えたことは余りにも有名である．彼の式は振動数の殆ど全範囲に渡って実験結果によく合うことが知られている．この法則は現在では光子の統計に関する分布として簡単に導かれ，プランク流のモデル化も導出方法も必要としない．しかし量子論誕生の端緒となった仕事の主要な武器の一つがエントロピであることや，分布則の導出が統計学に於ける分布の密度関数の推定と思想的にも類似するものがあることなどは非常に興味深い．彼の仕事の中には，現代の統計家の問題の取組み方の原点と言ってもよいものが見出される様に思える．以下で，その事を彼の式の誘導の概略と共に見る事にしよう．

Planck（1900 b, 1901）は放射場を振動数 ν の N 個の線形な振動子（共鳴子）がたえずその振幅や位相を不規則に変動させているものと考え，照射を受け振動する共鳴子の一個当りの熱力学的エントロピ S をその平均振動エネルギ U の関数として与えることにより問題の解決を図った．系の全エネルギを $U_N=NU$，全エントロピを $S_N=NS$ とする．S_N は全エネルギ U_N を放射場を構成する共鳴子に配分することから起る無秩序さに関係している．Planck（1900 b, 1901）は S_N を特定化する為に全エネルギの分配に関し次の仮定を置いた：連続量である U_N を離散的な整数個の有限な等しい部分（エネルギ要素 ε）からなるとする．これにより

$$U_N=P\cdot\varepsilon(=NU). \tag{4.1}$$

但し P は U_N/ε の整数部分で一般に大きな数である．ε が満すべき条件は分布則を得る過程で決定される．彼の上述のモデルが量子仮説を生み出したわけだが，気体分子運動論で Boltzmann（1877）が根本思想として全エネルギの有限な離散化を考えながら，結局は $\varepsilon\to 0$ の無限小を考えて連続化してしまったのと違って，Planck は ε を最後迄有限に留めた事で成功した点は結果的に物理学に革命をもたらしたことになる．この事を可能にした要因の一つとして Planck が十分信頼するに足る実験データの提供を受けて理論を作り上げる機会に恵まれたことも数えられよう．

さて，Planck の場合全エネルギの分配の仕方の総数 W は組合せ理論から見れば次の事を意味する：P 個の同種の（区別のつかない）エネルギボールを，N 個の異なる（区別のある）箱に分配する方法の数に等しい．但し一個もボールが入らない箱があってもよい．このことから

$$W=\binom{N+P-1}{P} \tag{4.2}$$

と与えられる．これは量子統計に於る Bose-Einsteins 統計の原形と見做されよう．Planck は上の W を構成するそれぞれの配分の仕方を Boltzmann（1877）の考察したものからヒントを得て同じく Komplexion と呼んだが，考え方はまるで異なっている．Planck の Komplexion は正整数 P を N 個の非負整数に順序も考慮して分割する仕方の事であり，組合せ論に於る Composition に対応している．各点の座標要素が非負である N 次元の整数格子点分布に関する確率の計算に於て，各格子点，即ち Komplexion に等確率が付与されている時，P の個数に関してある特定な条件が満されている領域内に含まれる格子点の数を数えあげる問題がしばしば起る．その際我々は上の W の計算が必要になる．それは Planck の考えた確率の計算方法と一致するものであり，彼の考えは統計学の分布論の立場から見れば極く自然と言える．

Planck は上の設定で，Boltzmann-Planck の原理と，P, N を十分大きいとして，Stirling の公式を用い，放射場の全エントロピの近似式を次の様に計算した：

$$S_N = k \ln W = k\{(N+P)\ln(N+P) - N \ln N - P \ln N\}.$$

$S_N = NS$, $P/N = U/\varepsilon$ に注意して，共鳴子一個当りのエントロピ S はその平均振動エネルギ U の関数として

(4.3) $$S = k\{(1+U/\varepsilon)\ln(1+U/\varepsilon) - (U/\varepsilon)\cdot\ln(U/\varepsilon)\} \equiv \varphi(U/\varepsilon)$$

の如く変数 U/ε と普遍定数を含む関数形で与えられた．この段階では放射場が必ずしも熱平衡状態にあるとは仮定する必要はない．

Planck は更に，絶対温度 T で熱平衡状態にある放射の理論的基礎を Wien (1893)のいわゆる変位則に置き，今考えている問題に適する形に書き直す事を試みた．それは，振動数が ν と $\nu + d\nu$ の間にある放射のエネルギの空間的密度を $u_\nu(\nu, T)d\nu$ とすると

(4.4) $$u_\nu = (\nu/c)^3 \varphi_0(T/\nu),$$

ここに c は光速度で，関数 φ_0 の定数は c に無関係である．u_ν と U の関係は Maxwell の古典電磁気学に基づいて Planck (1900) は既に次の関係を得ていた：

(4.5) $$u_\nu = 8\pi\nu^2 U/c^3.$$

共鳴子一個当りの平均エネルギは，従って

(4.6) $$U = \nu\varphi_1(T/\nu)$$

の形で表わせる．温度について見れば

(4.7) $$T = \nu\varphi_2(U/\nu)$$

と表現できる．

そこで Planck は熱力学の基本関係式の一つ

(4.8) $$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_V$$

を利用することを考えた．ここに V は放射場の体積である．

この式の導入の背景には次のようなことが考え得る．U と V を独立変数とし $S = S(U, V)$ とする．今我々は放射場が定温熱源である空間の壁に接触し，その体積も一定である熱的平衡状態を考えればよいから，熱力学の理論により，Helmholtz の自由エネルギに F について

(4.9) $$dF \equiv dU - TdS = 0$$

が成立しなければならない．これより (4.8) を得る．従って U に関する推定方程式とも見做せる (4.8) は，Boltzmann に従って Wien (1896) が考えた U, V が一定の場合のエントロピ極大化原理に基づくものではなく，Helmholtz の自由エネルギ極小化原理に基づいていることが分る．統計学との類推から見て Planck は U をパラメータとして，修正された熱力学的最尤方程式とでも言うべき (4.8) を用いたので，Wien とは違って，偏りのない U の推定関数を得ることになったとも言えよう．

(4.7), (4.8) より

$$\frac{dS}{dU} = \frac{1}{\nu}\varphi_3(T/\nu), \ (\varphi_3(T/\nu) = 1/\varphi_2(T/\nu)),$$

だからこれを積分して，熱平衡状態のエントロピ

(4.10) $$S_{eq} = \varphi_{eq}(U/\nu), \ \left(\frac{d\varphi_{eq}}{dU} = \frac{1}{\nu}\varphi_3(U/\nu)\right)$$

を得る．これが Wien (1893) の変位則に対する Planck の表現である．これにより体積一定の透熱的な放射場で振動する共鳴子一個当りのエントロピは，熱平衡状態では変数として U/ν のみを持ち，その他は普遍定数だけを含む関数 φ_{eq} で表現される．従って φ_{eq} の型を決める事が問題になる．Planck はここで (4.3) で与えたエントロピ S の表現を φ_{eq} の決定に用いた．即ち (4.10) を (4.3) に適用するなら

$$\varepsilon = h \cdot \nu \tag{4.11}$$

なる比例関係が成立しなければならない事を指摘した．これが Planck のいわゆる量子仮説である．h が周知の Planck 定数で，彼は全放射エネルギの空間密度の理論式と熱放射に関して当時得られていた実験値を突合せて，Boltzmann 定数 k と一緒にその数値を決定した．

なお (4.11) と (4.3) から逆に放射の熱平衡条件を与える Wien-Planck の変位則 (4.10) が得られる．よって条件 (4.11) は Planck のモデルに於て，放射が熱平衡状態にある為の必要十分条件と見做せる．

さて (4.11) を使って熱平衡状態にある共鳴子一個当りのエントロピの具体形は

$$S_{eq} = k\{(1+U/h\nu)\ln(1+U/h\nu) - (U/h\nu)\ln(U/h\nu)\} \tag{4.12}$$

となる．これと (4.8) から U に関する熱力学的推定方程式が次の様に与えられる：

$$\frac{1}{T} = \frac{k}{h\nu}\ln\left(1+\frac{h\nu}{U}\right). \tag{4.13}$$

U について解けば，共鳴子一個当りの平均の振動エネルギは

$$U = \frac{h\nu}{e^{h\nu/kT}-1} \tag{4.14}$$

と求まる．よって放射の空間エネルギ密度は (4.5) により

$$u_\nu = \frac{8\pi h\nu^3}{c^3} \cdot \frac{1}{e^{h\nu/kT}-1} \tag{4.15}$$

である．これが有名な Planck (1901) の熱放射分布である．

ここで少し関連する事柄を述べておこう．(4.13) は Planck (1900 b) が Wien (1896) の放射分布を試行錯誤とも見える方法で (4.15) の形式に改良した時に使った次の式を積分した形になっている．

$$\frac{d^2S}{dU^2} = \frac{\alpha}{U(\beta+U)} = -\frac{\alpha}{\beta}\left(\frac{1}{\beta+U} - \frac{1}{U}\right). \tag{4.16}$$

(4.13) と比べると，実は，$\alpha = -k$, $\beta = \varepsilon$ である．(4.15) を導いた時の Planck (1901)に於ては d^2S/dU^2 は表面に現れなかった．しかし Planck (1900 b)は (4.15) に対応する結果を見出す舞台裏でこの二次導関数の逆数とエネルギ U の間に (i) U の小さな値に対しては U に比例し，(ii) 大きな U の値に対しては U^2 に比例するという重要な関係に気づいて，Wien の放射公式の外挿を図り，その結果として (4.16) にたどり着いたようである．もっと正確には，Planck の (4.16) 式には，その後，次の様な意味も含まれている事が分った：(i) Planck の指摘の様に Wien の放射公式に対して

$$1\Big/\frac{d^2S}{dU^2} = -(\nu h/k)U \tag{4.17}$$

という形が成立しており，一方 (ii) エネルギ等配分を仮定すると古典電磁力学に基づく Rayleigh-Jeans の式に対して

$$1\Big/\frac{d^2S}{dU^2} = -(1/k)U^2 \tag{4.18}$$

という形が成立している．そして (4..16) はこれら二式の最も簡単な内挿公式として，両者を加えその逆数を取ったものとしても出て来る事が明らかになった (cf. Planck, 1930；天野, 1948).

ところで Einstein (1906) は彼の Brown 運動の研究で展開した'熱のゆらぎ'の考察手法を平衡状態にある熱放射のエネルギのゆらぎの平均 $\overline{\Delta u_\nu}$ を求める為に応用した．やはり Boltzmann の原理に基礎を置くその結果は，Planck の放射公式 (4.15) を経験事実の表現と見做した上で，計算された．それを使うと体積 V の中にある放射エネルギを u_ν, その時間的平均を

$\bar{u}_\nu$とすると振動数がνと$\nu+d\nu$の間にある単色放射線のエネルギの時間的ゆらぎの二乗平均は

$$\overline{(\varDelta u_\nu)^2}=\overline{(u_\nu-\bar{u}_\nu)^2}=\nu h u_\nu+\frac{c^3}{8\pi\nu^2 d\nu}\frac{u_\nu^2}{V} \tag{4.19}$$

で与えられた．よく知られている様に第1項は光の粒子性を，第2項は波動性を表わしていることをEinsteinは明らかにしたわけである．

そこで我々はPlanckの式(4.16)および(4.17), (4.18)とEinsteinの結果を比べて見よう．振動数がνと$\nu+d\nu$の間にある共鳴子の数を$n(\nu,\ d\nu)$とすると

$$u_\nu/U=n(\nu,\ d\nu)=8\pi V\nu^2 d\nu/c^3 \tag{4.20}$$

であるから（cf. 池田，1975），(4.19)式より

$$\overline{(\varDelta u_\nu)^2}=n(\nu,\ d\nu)\{\nu hU+U^2/d\nu\}.$$

Planckの熱放射分布は，$V=1$, $d\nu=1$とした場合の表現であることに注意すれば，

$$\begin{aligned}\overline{(\varDelta u_\nu)^2}/kn(\nu,\ d\nu)&=(\nu hU+U^2)/k\\&=-k\{(4.17)\text{の右辺}+(4.18)\text{の右辺}\}\end{aligned} \tag{4.21}$$

となる．結局，Planckの(4.16)式あるいは関連する(4.17), (4.18)式の中に熱平衡状態にある放射エネルギの，平衡状態からの熱的なゆらぎが陰にではあるが見事に捉えられていたと言えよう．この事は$u_\nu\propto U$から予想されることであるが，従来指摘されていなかったように見受ける．

ところで(4.12), (4.20)を使えば，放射場の熱平衡状態に於るエントロピ密度$\mathscr{E}^*$と放射エネルギの空間密度u_νの関係は次のようになる：

$$\begin{aligned}\mathscr{E}^*&=\mathscr{E}^*(u_\nu/\nu)\\&=nk\{(1+u_\nu/nh\nu)\ln(1+u_\nu/nh\nu)-(u_\nu/nh\nu)\ln(u_\nu/nh\nu)\}.\end{aligned} \tag{4.22}$$

(4.9)の条件は従って

$$J_\nu\equiv\frac{u_\nu}{kT}-\frac{1}{k}\mathscr{E}^*\left(\frac{u_\nu}{\nu}\right)\to\text{極小}, \tag{4.23}$$

と同値である．(4.20)の2番目の関係に注意すると，この基準と同等な$dJ_\nu/du_\nu=0$, $d^2J_\nu/du_\nu^2>0$によって，当然の事ながら，(4.15)式を再び得る．既ちJ_νの極小値をJ_ν^*とすると

$$J_\nu^*=n\frac{h\nu}{kT}(e^{h\nu/kT}-1)^{-1}-\frac{1}{k}\mathscr{E}^*\left(\frac{u_\nu}{\nu}\right), \tag{4.24}$$

ここに$n=n(\nu,\ d\nu)$．これより，特に$h\nu\ll kT$ならば

$$n-\frac{1}{k}\mathscr{E}^*\left(\frac{u_\nu}{\nu}\right)\sim J_\nu^*\to\text{極小}, \tag{4.25}$$

$h\nu\gg kT$ならば

$$n\frac{h\nu}{kT}e^{-h\nu/kT}-\frac{1}{k}\mathscr{E}^*\left(\frac{u_\nu}{\nu}\right)\sim J_\nu^*\to\text{極小}, \tag{4.26}$$

が従う．(4.25), (4.26)はそれぞれ，共鳴子の振動の自由度にエネルギが等配分される範囲，および振動が励起している共鳴子が殆んどないWienの分布則が成立する範囲に於る熱放射の平衡状態達成の基準と見做される．この事を一般化して考えるならば，(4.23)の第1項u_ν/kTは熱放射の共鳴子モデルに於て振動数νに対する振動の〝広義の自由度″とも呼ぶべき量で，温度Tと関連して測定されるものであると解釈できる．$h\nu/k\ll T$であるなら，例えばTが常温で熱放射を観測するならば(4.25)の基準が使われることになる．これらは統計学に於る最近のいくつかの情報量基準との密接な類推を感じさせる．$h\nu/k\gg T$の場合も調べておこう．Tが常温であっても(4.26)から振動の〝自由度″はエネルギを等配分されておらず殆どゼロ，そしてエントロピ極大原理がこの場合熱平衡条件を決めてしまうことになる．なお，エントロピ極大

が熱平衡条件になるのは，熱力学の理論により，体積 V，エネルギ密度 u_ν が一定(従って放射場の内部エネルギも一定）の時であるから，エントロピ密度 $\mathscr{E}^*$ の独立変数は ν のみとなる．Wien (1896) はエントロピ極大化原理から熱放射分布を導こうとした．彼は完全に光を反射する壁で囲まれた容器内を放射場とし，放射体として理想気体を考えたが，このモデルから放射場のエントロピと ν 又は波長の関係を見出すことができなかった．彼は結局，気体分子運動論と彼自身の変位則が成立するようなある条件を u_ν に仮定した上で，次の Wien の放射分布則

$$u_\nu = A\nu^3 e^{-h\nu/kT} \qquad (A\text{: 定数}) \tag{4.27}$$

に相当する表現を得た．我々の場合，この結果は (4.26) の基準から定数の表現を含めて直ちに得られるものである．歴史的によく知られているように，(4.27) 式は ν/T が小さいとき，即ち (4.25) を熱平衡の基準として使うべき時，に実験結果とかなりずれてしまう．しかし ν/T が大きいときは実験と非常によく合う．Einstein (1905)はこの実験事実を裏付けとして (4.27) 式を彼の光量子理論の初期に用いた事は興味深い．何故彼が Planck の放射分布を用いなかったかを推定すると，彼の興味が黒体放射に於て現れる光の粒子性にあり，放射分布の誘導そのものではなかったからではないかと思われる．即ち，彼は Planck の放射分布の内の Wien 的な部分こそが光の粒子性が現れている部分であると最初から見通していたように思える．なお，注目すべき事は，形式こそ違うが，Einstein も彼の結果を導くに当って Planck と同じく熱力学の基本関係式 (4.8) と（両者に "確率" の解釈の違いはあったが）Blotzmann の原理を用いて，Planck の量子条件に対応する表現を導き，光子の量子性を明確に説明した事である．Einstein はその後 Planck の熱放射論の含む不完全さを種々のアプローチで改善したが，これらの仕事を通じて，Planck の分布は単に黒体の熱放射の問題に留まらず固体の比熱に於ける量子理論等にも密接に関連することが立証された事も周知である．いずれにせよ，熱放射分布を導いた Planck の苦闘こそが今日の量子論に基づく諸科学の巨大な第一歩であったことは明白である．自然現象に対峙する分布を導くことが重大な課題である統計家にとって，彼の仕事は非常に参考になることをもっと一般に認識されてよいのではないかと思われる．

さて，Planck の誘導法を一般化して我々は (4.23) の基準を得た訳であるが，いま

$$\frac{1}{2}\mathrm{FEC} = l_\nu - \mathscr{E}^*(u_\nu/\nu)/k \longrightarrow \text{極小} \tag{4.28}$$

と置き，「**自由エントロピ極小化基準**」と呼ぶことにする．ここに

$l_\nu = u_\nu/kT$：広義の "自由度"

$\mathscr{E}^*(u_\nu/\nu)/k$：熱平衡状態の無次元化エントロピ密度

である．尚「自由エントロピ」は造語であって，Helmholtz の自由エネルギのエントロピと関連する TS 項を無次元化エントロピ密度で書き直したことに由来して名付けた．

空洞放射の分布則を求める場合には外界とエネルギのやり取りのある Planck モデルの様なものこそ意味があり，従って Planck (1901) の最終的な結論は「FEC **極小化基準**」によって，Wien (1896), Rayleigh-Jeans (1900, 1905), Planck (1900b) の三つのモデルから最良なモデルとして自己のものを選択した事に相当しているとも解釈できる．この事はまた (4.24)～(4.26) の比較から，Planck のみが正しい "自由度" を与えた事を意味する．孤立系に於る単純な熱力学的意味のエントロピ最大化原理から空洞放射の分布則が導けない事は以上の考察から明らかである．もしこの原理に頼ろうとするなら，共鳴子の振動エネルギは一定でなく，その平均値が与えられているとする条件が付加されねばならない．しかしこれは結局 Planck のモデルに帰着する．

ところで，先述の様に，量子力学が建設された今日では放射分布の誘導は Planck のやった方法は取られない．Planck の共鳴子モデルを仮定する場合でも次の様に簡単に誘導できる：放射場は温度 T の熱浴に接し，共鳴子の各固有振動も互に独立と見做せるから，振動の固有エネルギを U_i ($i=0, 1, 2, \cdots$) としてその夫々が一つの量子状態にあるとすると，各 U_i は量子力学の教えるところにより，カノニカル分布に従う．共鳴子を量子論的な調和振動子と考えると，平均エネルギー U は量子論的分配関数 Z (cf. 池田，1975) を用いて次の様に表現される：

$$U=\sum_{i=0}^{\infty} U_i e^{-\beta U_i}/Z,\ Z=\sum_{i=0}^{\infty} e^{-\beta U_i}, \tag{4.29}$$

ここに $\beta=1/kT$ であり $U_i=ih\nu$ ($i=0, 1, \cdots, I$) は一つの要素の共鳴子が取り得るエネルギ水準で，古典的な零点から測られているものとする(U_i をこのように表現しなければならないことは Einstein (1906) が指摘した)．これより

$$U=-\frac{\partial}{\partial\beta}\ln(\sum_{i=0}^{\infty} e^{-\beta U_i})=-\frac{\partial}{\partial\beta}\ln Z=\frac{h\nu}{e^{\beta h\nu}-1} \tag{4.30}$$

即ち，(4.14)式が導かれた．あとは前と同様にして(4.15)式を得る．

上に述べた(4.29)基づく U の分布の導出の基礎にも以下のように統計学に於ける分布の推定の見地から興味ある事が存在しているように思える：Planck のモデルに於て，N 個の共鳴子の全体のエネルギ U_N は

$$U_N=P\varepsilon=NU$$

によって，ε を最小単位として分配された．その方法の数は，(4.2)式で与えられた．ところで全エネルギ U_N の大きさが一定という条件の下で特定の共鳴子がある時刻に於て $i\varepsilon=U_i$ ($i=0, 1, 2, \cdots$) のエネルギを持つ確率は

$$\Pr(U_i\mid i)=\binom{N+P-i-2}{P-i}\Big/ W=\frac{N-1}{N+P-1}\prod_{j=0}^{i-1}\frac{P-j}{N+P-2-j} \tag{4.31}$$

である．P は N に比べて大きく，N も1に比べて十分大きいものとし，$i/P=o(1)$ と仮定すると，近似的に

$$\Pr(U_i\mid i)=\frac{N}{N+P}\left(\frac{P}{N+P}\right)^i=\frac{1}{1+U/\varepsilon}\left(\frac{U/\varepsilon}{1+U/\varepsilon}\right)^i \tag{4.32}$$
$$=\frac{1}{1+U/\varepsilon}e^{-U_i/\tau}\qquad (i=0,\ 1,\ 2,\ \cdots)$$

ここに

$$\frac{1}{\tau}=\frac{1}{\varepsilon}\ln(1+\frac{\varepsilon}{U}) \tag{4.33}$$

と置いた．実は上の τ は(4.29)の β と次の様な関係にある：Planck の結果 $\varepsilon=h\nu$ 及び(4.13)に注意して

$$\tau=kT=1/\beta \tag{4.34}$$

更に(4.32)式の U に Planck の結果(4.14)を代入すると簡単に

$$\Pr(U_i\mid i)=\frac{e^{-\beta U_i}}{(1-e^{-\beta h\nu})^{-1}}=e^{-\beta U_i}/Z \tag{4.35}$$

が出る．即ち Planck の結果を用いれば $\Pr(U_i\mid i)$ がカノニカル分布である事が逆に証明されたことになる．ところで，我々は Planck (4.14)の結果から(4.33)の表現を得たが，今の目的が未知の U を導出した(4.29)式が何を意味するかを問題にしているので，(4.32)式に(4.14)を使う事ができない．けれども(4.32)式から，少なくとも

$$\Pr(U_i\mid i)\propto e^{-U_i/\tau},\qquad (U_i=ih\nu\,;\ i=0,\ 1,\ 2,\ \cdots,\ I) \tag{4.36}$$

であることは見て取れる．I は振動子の励起状態の総数であり，$\sum_{i=0}^{I} \Pr(U_i \mid i)=1$ だから上式右辺の係数は $Z'=\sum_{i=0}^{I} e^{-U_i/\tau}$ である．我々は上の (4.36) を尤度として，確率 $\Pr(i \mid U_i)$ を知りたい立場にあるものとする．その為には事前確率 $\Pr(i)$ $(i=0, 1, 2, \cdots, I)$ が設定できればよい．この確率に対しては我々は特に情報を持っていない．そこで統計力学の基本仮定である等重率に基づいて，U_N が一定という条件の下で一つの共鳴子が取り得るエネルギの状態 i を取る確率は先験的に離散一様分布 $\Pr(i)=1/(I+1)$ $(i=0, 1, \cdots, I)$ で与えられると仮定しよう．すると事後確率はベイズの定理により

$$\Pr(i \mid U_i)=\frac{\Pr(U_i \mid i)\Pr(i)}{\sum_{i=0}^{I} \Pr(U_i \mid i)\Pr(i)}=\frac{\Pr(U_i \mid i)}{\sum_{i=0}^{I} \Pr(U_i \mid i)}$$
$$=\frac{e^{-U_i/\tau}}{\sum_{i=0}^{I} e^{-U_i/\tau}}=\frac{e^{-ih\nu/\tau}}{\sum_{i=0}^{I} e^{-ih\nu/\tau}} \qquad (i=0, 1, \cdots, I). \tag{4.37}$$

となる．さて，一つの共鳴子の長時間に渡る観測の結果に基づいてそのエネルギ状態 i を求めようとする場合，その共鳴子のエネルギ状態 i に関する長時間平均を考えなければならない．この平均は統計力学に於ては，対応する位相集団の位相平均に等しいと捉えられる．$U_i=ih\nu$ $(i=1, \cdots, I)$ と定義されていることを考慮し，(4.30) で用いたのと同様な手法を使えば，平均エネルギ U は I を十分大として

$$U=h\nu \sum_{i=0}^{I} iP(i \mid U_i)=h\nu \sum_{i=0}^{\infty} iP(i \mid U_i)=\frac{h\nu}{e^{h\nu/\tau}-1} \tag{4.38}$$

となる．(4.34) の $\tau=kT$ を使えば再び Planck の結果 (4.14) と一致する．即ち，(4.38) を参照すれば，量子論的分配関数に基づく (4.29) 式の右辺は実は統計力学に於る等重率の原理を見直した，エネルギ状態 i の事後確率分布による平均値に，エネルギ単位 $\varepsilon=h\nu$ を掛けた量を表現していることになる．上の考え方は，理論的には，エネルギ状態が退化するとき，事前分布 $P(i)$ はもはや離散一様分布でなくなるが，その場合にも有効である．

注意すべき事は，今問題にしている放射の熱平衡状態は，(4.37) に於て最大の確率を与える状態ではなくて，事後確率分布の平均として与えられる状態を指していることである．もし事後確率最大となる状態を取るとそれは明らかに $i=0$ の状態であり，これは共鳴子が励起していない状態である．しかし，振動が実際に起っていてかつ定常状態であることを前提として各共鳴子の平均エネルギを問題としている Planck のモデルに於ては，この確率は特に意味を持つ訳ではない．ただし Einstein (1906) が別の観点から指摘した様に，共鳴子のほんの少しだけが零でないエネルギを持つことはよく表現している．

先に見たように Planck が各 Komplexion に等確率を先験的に仮定した場合の (4.3) 式のエントロピは，放射の熱平衡状態が問題となっていることを考慮して，評価し直されていた．即ち熱力学第 2 法則から導かれる (4.8) 式はあるいはそれからの帰結である Wien の変位則の Planck 表現 (4.10) によって，より正しい解答を与える (4.12) 式のエントロピに書き直された．この操作は，量子的状態関数を使った場合に事前分布から事後分布への見直しをした事に相当している．Planck の Komplexion に基づくエントロピを使った誘導法では，確率や尤度は陽な形での役割を演じないが，上の二つの方法の比較からそれらは表裏一体の関係にあると言える．ベイズの定理は，従って，熱力学第二法則と並行し，考察対象の系が変化する条件を与えるという役目をしていることが改めて認識できる．

再び Einstein のゆらぎの考察に舞い戻ろう．一つの共鳴子の平均エネルギ U は分配関数 Z

を使うと(4.29)又は(4.30)のように表現された．これらの関係式を利用すると共鳴子一個のエネルギの二乗平均，即ち分散は，

$$\overline{(\Delta U)^2}=\frac{\partial^2}{\partial\beta^2}\ln Z \tag{4.39}$$

となる．これに(4.30)，(4.36)，(4.8)を考慮すると

$$\overline{(\Delta U)^2}=-\left(\frac{\partial^2 U}{\partial\beta^2}\right)_V=kT^2\left(\frac{\partial U}{\partial T}\right)_V=-k\left(\frac{\partial U}{\partial(1/T)}\right)_V=-k\Big/\left(\frac{\partial^2 S}{\partial U^2}\right)_V. \tag{4.40}$$

である．最後の項のエントロピ S を熱平衡状態の S_{eq} で置き替ると，$S\leqq S_{eq}$(最大エントロピ)だから

$$\overline{(\Delta U)^2}\geqq -k\Big/\left(\frac{\partial^2 S_{eq}}{\partial U^2}\right)_V \tag{4.41}$$

となる．$-(\partial^2 S_{eq}\big/\partial U^2)_V$ を統計学に於ける Fisher 情報量の類推量と考えるならば，(4.41)は Cramér-Rao の不等式の熱統計学に於ける表現と見做せよう．ところで，我々は熱平衡状態からの U のずれ，ゆらぎ，を考察しているから，(4.41)の不等式で等号が成立する場合を対象とすればよい．即ち U のゆらぎの二乗平均は

$$\overline{(\Delta U)^2}=-k\Big/\left(\frac{\partial^2 S_{eq}}{\partial U^2}\right)_V \tag{4.42}$$

で与えられる．これは Einstein (1909) が少し違った観点で導いた結果と本質的に一致する．(4.19)で述べた彼の結果は上式と並行する関係式を(4.22)式のエントロピ密度 $\mathscr{E}^*$ に適用して導かれたものである．なお，(4.40)を(4.12)の S_{eq} に適用すると，我々は直接 Planck の内挿

$$\overline{(\Delta U)^2}/k=(\nu hU+U^2)/k \tag{4.43}$$

を得る (cf. (4.21))．

この段階迄来ると，Planck (1900b, 1901) は，統計学との類推から言って，間接的に U の分布の最小分散不偏推定に相当する事を行ったと見做せる．この事を可能にしたのはやはり(4.3)式と(4.8)式である．ここでは系の絶対温度の逆数 $1/T$ を推定問題に於る未知パラメータと考えると，(4.8)式の右辺はその $1/T$ に関する熱力学的な不偏推定値と言ったものと考え得る．

(4.3)あるいは(4.12)のエントロピが変数として U のみを含み，パラメータ $1/T$ に無関係なことに注意すると，(4.13)式の右辺が $1/T$ の不偏推定値を与えていると考えてよかろう．現実には理論値 $1/T$ を観測値で代用することになるから，(4.13)を U について解いた(4.14)式が直接的な意味を持つことになる．(4.14)による U の最小分散性は，(4.41)，(4.42)から従うことになる．

上の考察の過程で，Planck の公式の誘導は，統計学の推定論に於る Rao-Blackwell の定理 (cf. Rao, 1973, Blackwell, 1947) の一般化との類似性を感じさせる．その事は次の統計的推定問題へこの定理を応用する時に現れる事柄の物理的意味を検討することによって多少明らかになるであろう：

(問題) 正規分布 $N(\mu,\ \sigma^2)$ の未知パラメータ μ の推定可能関数 $g(\mu)$ の最小分散不偏推定値を，σ^2 を既知として求めよ．

この問題は例えば次の様にして解ける．$X_1, X_2, \cdots, X_n$ をこの分布からの大きさ n の無作為標本とする．周知の様に，$\bar{X}=\sum_{i=1}^{n} X_i/n$ は μ の十分統計量である．$\bar{X}$ を与えた時の X_1 の条件付分布は μ に無関係な $N(\bar{x},\ (1-1/n)\sigma^2)$ で与えられる．

よって Rao-Blackwell の定理により $g(\mu)$ の最小分散不偏推定値は

(4.44) $$\varphi(\bar{x})=\frac{1}{\sqrt{2\pi}\sigma_0}e^{-(\bar{x}-x)^2/2\sigma_0^2},\ \sigma_0^2=(1-1/n)\sigma^2,\ (n>1)$$

となる．もし g として正規分布の分布関数に関連する量

(4.45) $$\Phi_\sigma(c\,;\,\mu)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{c}\exp\{-(x-\mu)^2/2\sigma^2\}dx,\ (c,\ \sigma^2\text{：既知})$$

を考えるなら，その最小分散不偏推定値は，直ちに

(4.46) $$\Psi(c\,;\,\bar{x})=\frac{1}{\sqrt{2\pi}\sigma_0}\int_{-\infty}^{c}\exp\{-(x-\bar{x}^2)/2\sigma_0^2\}dx$$

と与えられる．上の解法は純粋に統計学的手法に沿ったものである．ところで Kolmogorov (1950) は上の $g(\mu)$ を推定する問題は次の熱伝導方程式

(4.47) $$\frac{\partial\varphi}{\partial t}=\frac{\partial^2\varphi}{\partial z^2}$$

の $t>\sigma^2/2n$ での有界な解で $[\varphi]_{t=0}=g(z)$ となるものを求めることに帰着し得ることを指摘し，その結果として (4.44) を導いた．周知の如く熱伝導方程式は熱力学第 1，第 2 法則を用いて熱流や熱拡散等の不可逆過程の時間的変化を表現する基礎的な偏微分方程式である．従って $g(\mu)$ の最小分散不偏推定を行うという統計的問題に対する熱力学的な類推は，非平衡熱力学の問題として理解されるであろう．問題は (4.47) の方程式の背後にある物理的意味である．この式を扱う代りに以下では少し一般化された，物体表面が熱浴に接している一様的な固体内の三次元熱伝導方程をまず考える．(cf. De Groot, 1963).

$u=u(x,\,y,\,z\,;\,t)$ を固体内の位置 $(x,\,y,\,z)$ で代表される無限小直方体の近傍部分に，時刻 t で含まれている比内部エネルギとする．また ρ をこの物体の密度，$\boldsymbol{J}_u$ をこの微小部分の全表面からその内部へ流れ込む熱流，s を単位質量当りのエントロピそして T を絶対温度とする．非平衡熱力学の理論 (cf. De Groot, 1963) により次の事が要請される：

熱力学の第 1 法則に関連して

(4.48) $$\rho\frac{\partial u}{\partial t}=-\mathrm{div}\boldsymbol{J}_u,$$

熱力学第 2 法則に関連して

(4.49) $$T\frac{\partial s}{\partial t}=\frac{\partial u}{\partial t}.$$

である．$\boldsymbol{x}$ をベクトル量，α をスカラー，$\psi(\alpha)$ を α に関し微分可能な関数とすると $\mathrm{div}(\alpha\boldsymbol{x})=\alpha\mathrm{div}\boldsymbol{x}+\boldsymbol{x}\mathrm{grad}\alpha$ および $\mathrm{grad}\psi(\alpha)=\psi'(\alpha)\mathrm{grad}\alpha$ だから，上の二式より

(4.50) $$\rho\frac{\partial s}{\partial t}=-\mathrm{div}\left(\frac{\boldsymbol{J}_u}{T}\right)-\frac{\boldsymbol{J}_u}{T^2}\mathrm{grad}\,T\equiv-\mathrm{div}\boldsymbol{J}_s+\xi(s)$$

と表現できる．この式は比エントロピ s の時間的変化は $\boldsymbol{J}_s=\boldsymbol{J}_u/T$（エントロピ流）の負の発散と $\xi(s)=-\boldsymbol{J}_uT^{-2}\mathrm{grad}\,T$（エントロピ源又は局所エントロピ生成速度）に基づくことを示している．

いま，注目している熱伝導をしている物体自体については，(4.50) の面積分を行い，ガウスの発散定理を用いると

(4.51) $$\frac{\partial}{\partial t}\int s\rho dV+\int\boldsymbol{J}_s\cdot\boldsymbol{n}dA=\int\xi(s)dV$$

となる．ここに dV は体積要素，dA は面積要素で，$\boldsymbol{n}$ は物体部分の表面の内部向きの法線ベクトルを表す．上式は，右辺のエントロピ源 $\xi(s)$ から生成されたエントロピは，左辺第 1 項の物体の体積内のエントロピ増加（エントロピ輸送）部分と，第 2 項に示される外界へのエントロピ流として散逸することを意味している．

ところで，今の我々の目的は，見なれた形での熱伝導方程式の誘導であるから，(4.48)式を絶対温度 T を用いて表現し直す事にする．先に考察の対象とした無限小直方体にエネルギ保存法則を適用すると

(4.52)
$$\rho\frac{\partial u}{\partial t}dt=\left(\int C_V\rho\frac{\partial T}{\partial t}dV\right)dt=\left(\int \boldsymbol{J}_u\cdot\boldsymbol{n}dA\right)dt$$
$$=\left(-\int \mathrm{div}\boldsymbol{J}_u dV\right)dt$$

となる．ここに C_V は物体の定容比熱を表す．これより熱力学第1法則の次の表現を得る：任意の時刻 t と位置 (x, y, z) について

(4.53)
$$C_V\rho\frac{\partial T}{\partial t}=-\mathrm{div}\boldsymbol{J}_u$$

これに対応する熱力学第2法則の微分表現が必要になる．閉じた系に於ける任意の不可逆過程に対して，第2法則は

(4.54)
$$dS+dS_a\geqq 0$$

である事を要求する．ここに dS は熱伝導している物体のエントロピ変化，dS_a は外部物体のエントロピ変化である．(4.52)に現れた無限小直方体内部のエネルギおよびそこへの供給熱量を考慮すれば上の不等式は

(4.55)
$$\int C_V\rho\frac{1}{T}\frac{\partial T}{\partial t}dV+\int\frac{1}{T}\mathrm{div}\boldsymbol{J}_u dV>0$$

に等しい．よって熱伝導をしている無限小の物体に対して熱力学第2法則は次の形で表現される：時刻 t と位置 (x, y, z) について

(4.56)
$$C_V\rho\frac{1}{T}\frac{\partial T}{\partial t}+\mathrm{div}\boldsymbol{J}_s\geqq 0$$

が成立する．上の左辺は(4.48), (4.49), (4.50), (4.53)から $\xi(s)$ に等しい．故に

(4.57)
$$\xi(s)=-\boldsymbol{J}_u(\mathrm{grad}\,T)/T^2\geqq 0$$

でなければならない．(4.50)式の等式に対応して，(4.56)の不等式を等号で置き替る事を考えよう．そこで $\boldsymbol{J}_u$ と $\mathrm{grad}\,T$ の間に次の線型関係

(4.58)
$$\boldsymbol{J}_u=-\lambda\cdot\mathrm{grad}\,T$$

を導入する (cf. Planck, 1930, De Groot, 1963)．これを(4.57)に代入すれば，$\lambda\geqq 0$ でなければならないことが分る．また(4.53), (4.58)から熱伝導方程式

(4.59)
$$\frac{\partial T}{\partial t}=(\lambda/\rho C_V)\mathrm{div}\cdot\mathrm{grad}\,T\equiv\varkappa\Delta T$$

を得る．ここに $\Delta=\mathrm{div}\cdot\mathrm{grad}$（ラプラシアン）を表す．実際の測定上の問題から，上の絶対温度 T の代りに，セッ氏温度 Θ を用いて表現すれば，通常の見なれた熱伝導方程式

(4.60)
$$\frac{\partial\Theta}{\partial t}=\varkappa\left(\frac{\partial^2\Theta}{\partial x^2}+\frac{\partial^2\Theta}{\partial y^2}+\frac{\partial^2\Theta}{\partial z^2}\right)$$

を得る．ついでながら，(4.58)式は Fourier の熱伝導の経験式で，λ は物体の熱伝導率である．

一次元熱伝導方程式は(4.59)の特別な場合として誘導できるから，(4.47)式は $T=\varphi$, $\varkappa=1$ の場合の一次元熱伝導方程式に相当することが分る．以上が Kolmogorov の熱伝導方程式を用いた推定法の一つの物理的背景である．注意すべき点は，我々が不偏推定すべき関数は温度 T (or Θ)であるということが以上の熱伝導の議論では明確に表れていることである．この事は先に注意したPlanckの熱放射分布の誘導で効いているのは $1/T$ の不偏推定とも考え得た事に通じている．ただし熱伝導問題に於る T は内部エネルギだけでなく位置及び時間を含む関数であるだけ複雑であることは上に見た通りである．

なお，Kolmogorov の様に熱伝導方程式を介して得た推定値が最小分散不偏推定になっているという事と，熱伝導方程式が熱力学第 2 法則に於る不等式を，(4.51) に見るように，系内の不可逆過程にエントロピ生成を考える事によって，等号で置き替えたことを考え併せると，対応する Rao-Blackwell の推定法は不偏推定法に於る Cramér-Rao の不等式を等号で置き替る事に成功したものだとも解釈できよう．後者の場合，θ を未知パラメータとして推定可能関数 $g(\theta)$ の任意の**不偏**推定量を $\hat{g}$，その改良を $\tilde{g}(t^*)=E[\hat{g} \mid T^*=t^*]$ とする時，平均二乗誤差 $E[\hat{g}-\tilde{g}]^2 \geqq 0$ がエントロピ生成に対応している．通常統計量 T^* としては応用上便利な θ の十分統計量が用いられるが，T^* は必ずしも十分統計量である必要はない．要は $\tilde{g}$ が θ に無関係，即ち統計量になるという事である (cf. Blackwell, 1947)．この事が本質的に $\tilde{g}$ が $g(\theta)$ の不偏推定量でかつ分散最小である事を導く．なお，Rao-Blackwell の結果は θ および T^* がベクトル値である場合にも成立するし，損失関数も分散だけでなく一般に凸関数であれば成立することが知られている (Arnold and Kattii, 1972)．従っていわゆる負のエントロピ関数もその資格がある．これに対応する熱伝導理論に基づく Kolmogorov の方法も，一般化された Stokes の定理を使えば多次元のモデルに対する熱伝導方程式を得る事が可能だから，上述の関連する部分は平行的に拡張できよう．

以上の本節の考察を通じて，Akaike (1973) の情報量基準 (AIC) の，これ迄それ程指摘されなかったと思われるが，エントロピを物指しとする熱統計学を主体とする物理的，統計的な背景が，ある側面からではあるが，浮び上って来たように思われる．この事は AIC が単に統計学のみを基盤とするものではなく，より広い科学諸分野の問題の解析，設計に際しても共通する指導的原理となり得ることを十分に示唆している．

この事を直観的に理解する為には，エントロピと情報量の関係を明らかにする必要がある．この為次節では AIC に於て重要な役割りを演じる K-L 情報量とエントロピの関係やこの情報量の若干の性質を考察することにする．

5. K-L 情報量とエントロピ

可測空間 $(\Omega, \mathscr{B})$ 上で定義される二つの確率測度 P, Q を考える．$\mathscr{B}$ 上で $P \ll Q$ の時，

$$I(P, Q)=\int \ln R(x) \cdot dP(x)=\int R(x)\ln R(x)dQ(x), \ if\ P \ll Q \tag{5.1}$$

で定義される量は P の Q に関する Kullback-Leibler の情報量(略して K-L 情報量)と呼ばれる．ただし，$R(x)$ は $\mathscr{B}$ 上での密度 (Radon-Nikodym 導関数)

$$R(x)=\frac{dP(x)}{dQ}(\geqq 0), \tag{5.2}$$

を表す．なお，$P \ll Q$ でなければ $I(P, Q)=\infty$ と定義する．

通常の応用に於て，P および Q が $(\Omega, \mathscr{B})$ 上のある σ-有限測度 μ に関する密度 f, g をそれぞれ持つ場合を考える事が多い．この時 K-L 情報量は見なれた形

$$I(f, g)=\int f(x)\ln\frac{f(x)}{g(x)}d\mu \tag{5.3}$$

で表現される．以下の議論ではこの形で K-L 情報量を扱う．Kullback (1959)は仮説弁別の立場から，この量を $g(x)$ に対し $f(x)$ が有利となる弁別の為の，分布 $f(x)$ からの一観測値当りの平均情報と呼んだ．しかしこの量はエントロピと関連づけて考えるとき，もっと重要な意味を持つことが分る．

この量は既に Boltzmann (1878) や Gibbs (1902) によって本質的に気づかれていたもので

ある．しかし彼等がこの量が分布の隔りを測る測度として重要であることを指摘していたにもかかわらずその後殆ど人々の注意を引くことがなかったようで，前者に対しては Akaike (1982) でその先駆性が明確に指摘されるまで全く見落されていたと言える．

K-L 情報量の重要な性質の一つはそれが非負（or 凸）である性質を持っていることである．この性質は Gibbs によっても，彼の分布のエントロピの極値性を証明する為に，既に明らかにされていたものである．

例えば熱源と接触して熱平衡状態にある閉じた系を記述する分布のモデルとしてカノニカル分布を採用するのが自然である事が，前節の関連箇所と同じ設定で，次の様にして示される：g_N を (3.25) に与えたカノニカル分布の密度関数，f_N を g_N と同じ平均エネルギを持つ任意の密度関数とする．即ち

$$\int f_N H d\tau = \int g_N H d\tau = \langle H \rangle, \tag{5.4}$$

$$\int f_N d\tau = \int g_N d\tau = 1, \tag{5.5}$$

ここに $d\tau$ は (3.22) と同じ意味を持つ位相空間に於る体積要素を，H は (3.25) の中のハミルトニアンを表すものとする．性質 $I(f_N\,;\,g_N) \geqq 0$ を利用すると

$$\begin{aligned} \mathrm{S}_G(f_N)/k &= -\int f_N \ln f_N d\tau \leqq -\int f_N \ln g_N d\tau = C + \beta \int f_N H d\tau \\ &= C + \beta \int g_N H d\tau = -\int g_N \ln g_N d\tau \equiv S_G(g_N)/k \end{aligned} \tag{5.6}$$

となる．ここに

$$C = -\frac{3N}{2}\ln(\beta N/2\pi M) + \ln \mathrm{Q}$$

である．(5.6) はカノニカル分布が条件 (5.4), (5.5) の下で，Gibbs のエントロピを最大にするものである事を示している．更に，上の様な設定で，熱平衡状態に於ては

$$\frac{S_G(g_N) - S_G(f_N)}{k} = -\int g_N \ln g_N d\tau - \left(-\int f_N \ln f_N d\tau\right) = \int f_N \ln \frac{f_N}{g_N} d\tau \geqq 0 \tag{5.7}$$

が成立する．即ち，上記の熱平衡状態では，カノニカル分布 g_N に基づく Gibbs エントロピ $S_G(g_N)$ と分布 f_N に基づく $S_G(f_N)$ の差を Boltzmann 定数で除した無次元量が分布 f_N の g_N に関する K-L 情報量に等しいことが分かる．この性質の類推的な結果は，統計学に於る指数分布族の特徴付け，分布の近似理論，極限理論などにも大いに役立つものである．

Boltzmann のエントロピ S_B に対しても，十分多くの粒子からなる孤立系の場合，分子間の相互作用を持たない一原子分子の理想気体については，(5.7) と同様な結果が成立つ：

$$\frac{S_B(g_N) - S_B(f_N)}{k} = N \int f_1 \ln \frac{f_1}{g_1} d\tau_1 \geqq 0, \tag{5.8}$$

ここに N は系内の分子数，f_1, g_1 は (3.23) と同様に定義される任意の一粒子確率密度関数で，$f_N(\boldsymbol{x}, \boldsymbol{P}) = \prod_{i=1}^{N} f_1(x_i, p_i)$, $g_N(\boldsymbol{x}, \boldsymbol{p}) = \prod_{i=1}^{N} g_1(x_i, p_i)$ であり，N 個の粒子同時密度関数 g_N はミクロカノニカル分布

$$g_N = \begin{cases} \text{一定}, & \text{if } E < H < E + \varDelta E, \ (\varDelta E \ll E) \\ 0, & \text{otherwise} \end{cases} \tag{5.9}$$

である．このとき g_1 は一粒子に付随する位相空間 (μ-空間) ではカノニカル分布である．(5.8) は (5.7) と同様に誘導できるが，ここでは (3.13) 式を誘導した時の例で与えてみよう：g_N として r 個のセルに等確率を与える分布を考える．即ち，(3.13)′ で f_1 の代りに $1/r$ とする．$n_i/N_0 = f_1 dV$ に注意して，(3.13)′, (3.15) を使うと

$$(5.10)\qquad \frac{S_B(g_N)-S_B(f_N)}{k}=(N_0\ln r+C_4)-\left\{-N_0\sum_{i=1}^{r} f_1\ln(f_1 dV)dV+C_4\right\}$$
$$=N_0\sum_{i=1}^{r} f_1\ln\left(f_1 dV\Big/\frac{1}{r}\right)dV$$

r を十分大きく取り，dV を十分小さくすると各セルで $1/r \fallingdotseq g_1 dV$ とできるから，近似的に (5.8) の右辺が導かれる．

上の事の統計学に於る類推は Sanov (1957) によって与えられた：$F_N(x)$ を大きさ N の独立なサンプルに基づく経験分布関数，$F(x)$ を X の分布関数，$\Phi(x)$ を他の任意の分布関数で，$\int_{-\infty}^{\infty}\ln(dF/d\Phi)d\Phi$ が存在するものと仮定する．このとき任意の $\varepsilon>0$ に対し

$$(5.11)\qquad P(\sup_x|F_N(x)-\Phi(x)|<\varepsilon)=\exp\left[N\int_{-\infty}^{\infty}\ln\frac{dF}{d\Phi}\cdot d\Phi+\gamma_\varepsilon+\mu_{N,\varepsilon}\right]$$

が成立する．ここに $\lim_{\varepsilon\to 0}\gamma_\varepsilon=0$, $\lim_{\varepsilon\to 0}\lim_{N\to\infty}\mu_{N,\varepsilon}=0$ である．

これより

$$(5.12)\qquad \lim_{\substack{\varepsilon\to 0\\ \varepsilon>0}}\lim_{N\to\infty}\frac{1}{N}\ln P(\sup_x|F_N(x)-\Phi(x)|<\varepsilon)=\int_{-\infty}^{\infty}\ln\frac{dF}{d\Phi}d\Phi.$$

(5.8) と (5.12) を比べると，(5.8) でミクロカノニカル分布 g_N と (5.12) に於る経験分布関数 F_N が対応していることが分る．

ところで，(5.7), (5.8) の中のエントロピは (3.22), (3.23) の H-関数 H_G 又は H_B に基礎を置くもので，f_N が熱平衡状態の時に，f_N によるエントロピと g_N によるエントロピの差が K-L 情報量と結びつくものであった．しかし K-L 情報量そのものは，Sanov の結果からも想像されるように，熱平衡状態だけでなく，非平衡状態に於ても，任意の二つの分布の隔りを評価する有効な測度として使えるものである．この事を再び (3.13) 式を使って，少し統計的見地から検討してみよう：$\sum_{i=1}^{r} n_i=N_0$ に注意すれば (3.13) は次の様に表現できる：

$$(5.13)\qquad \ln P=-N_0\sum_{i=1}^{r}\frac{n_i}{N_0}\ln\left(\frac{n_i}{N_0}\Big/\frac{1}{r}\right)+C_6.$$

従って，S_B/k は r 項確率分布 $\{n_1/N, n_2/N, \cdots, n_r/N\}\equiv f_N$ の r 項離散型一様分布 $\{1/r, 1/r, \cdots, 1/r\}\equiv g_N$ に関する K-L 情報量の $-N_0$ 倍に比例する．そこで，(3.13)′ 式の様に $n_i=N_0 f_1 dE$ とし，g_N も，Sanov の結果からの類推で，密度関数 g_1 を連続型分布からの r 個の独立なサンプルによって与えられた経験測度であると考えると，前と同様な極限操作により，(5.13) の右辺の第 1 項は

$$(5.14)\qquad S_B/k\propto -N_0\int f_1\ln(f_1/g_1)d\tau_1=-N_0\cdot I(f_1, g_1)$$
$$\equiv S_B^*(f_N, g_N)/k$$

と表現される．これとの類推で

$$(5.15)\qquad -\int f_N\ln(f_N/g_N)d\tau=-I(f_N, g_N)\equiv S_G^*(f_N, g_N)/k$$

と置くのは自然であろう．古典熱力学に於るエントロピとの対応を持つ S_B 或いは S_G に於て通常意味を持つのはそれら自身の値ではなく，異なる状態あるいは異なる分布に対する S_B あるいは S_G の差である．その観点からすると S_B^* あるいは S_G^* はより実用的で広範囲な応用を持つ量であると言える．(5.14) から S_B 自身が適当な条件下で S_B^* に比例しているから，K-L 情報量が時には**負のエントロピ**（より正確には f の g に関する負のエントロピ）と呼ばれるのも納得できる．

ところで，統計学に於てK-L情報量は次の様なパラメトリックな構造の下で考える事により，応用範囲を広いものにしている：$\{P_\theta;\ \theta\in\Theta\}$を可測空間$(\Omega, \mathscr{B})$上の確率測度とする．$X$をこの空間で定義される確率変数とし，$P_\theta$は$\sigma$-有限測度$\mu$に関し絶対連続とする．このとき$X$の密度を$f(x, \theta)=dP_\theta/d\mu$と表す．今$\tau(\theta)$を$\Theta$から$\Theta$上への1対1写像とし，任意の$\theta\in\Theta$に関し$P$は$P_{\tau(\theta)}$に関し絶対連続と仮定する．この時$P_\theta$の$P_{\tau(\theta)}$に関するK-L情報量は

(5.16) $$I(\theta,\ \tau(\theta))=\int f(x,\ \theta)\ln\frac{f(x,\ \theta)}{f(x,\ \tau(\theta))}d\mu$$

で定義される．Ωを一次元実空間，θを一変量とすると，よく知られている様に，$\theta-\tau(\theta)$が微少な場合，近似式

(5.17) $$I(\theta,\ \tau(\theta))\fallingdotseq I_F(\theta)\cdot(\theta-\tau(\theta))^2/2$$

が成立する．ここに$I_F(\theta)$はθに関するFisterの情報量で，次式で定義される：

(5.18) $$I_F(\theta)=E_\theta\left[\frac{\partial\ln f(x,\ \theta)}{\partial\theta}\right]^2=E_\theta\left[-\frac{\partial\ln^2 f(x,\ \theta)}{\partial\theta^2}\right].$$

これはθの$\tau(\theta)$の周りのゆらぎの二乗平均と結びつく量であり熱放射の場合に出て来た(4.40)と関係する．この結果，P_θの$P_{\tau(\theta)}$の周りでのゆらぎの確率P^*は近似的に

(5.19) $$P^*\sim\exp[-I(\theta,\ \tau(\theta))]$$

で与えられる．θが多変量の場合にも上の議論は拡張できる．この様に，パラメータθの$\tau(\theta)$からの微小なずれに対し，K-L情報量が分布のゆらぎに関連して登場してくるという事はその物理的背景と共に実に重要な性質である。もし，P_θが時間tに関係するならば，我々は上の性質を拡張して，非平衡であるが定常で安定な分布を議論する問題へ発展させ得る．その事は，頑健性を持つ拡張された推定問題に於ても，K-L情報量，ひいては情報量基準が重要な役割りを演じ得る事を十分示唆している．

参 考 文 献

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle (eds. B.N. Petrov and F. Csaki), *2nd International Symposium on Information Theory*, Budapest, Akademiai Kiado, 267-281.

Akaike, H. (1982). Prediction and Entropy, *MRC Technical Summary Report,* No. **2397**, Math. Res. Center, Univ. of Wisconsin ; *Revised Version,* to appear in *A Celebration of Statistics,* (eds. A. C. Atkinson and S.E. Fienberg), Spring-Verlag, New York (1985).

天野 清 (1948).「量子力学史」, 天野清選集 1, 日本科学社.

Arnold, J.C. and Katti, S.K. (1972). An application of Rao-Blackwell theorem in preliminary test estimators, *J. Multivariate Analysis,* **2**, 236-238.

Blackwell, D. (1947). Conditional expectation and unbiased sequential estimation, *Ann. Math. Statist.,* **18**, 105-110.

Boltzmann, L. (1872). Weitere Studies über das Wärmegleichgewicht unter Gasmolekülen, *Wiener Berichte,* **66**, 275-370.

Boltzmann, L. (1877). Über die Beziehung zweischen dem zweiten Hauptsatze der mechanichen Wärmetheorie und der Wahrscheinlichkeitsrechung respective den Sätzen über das Wärmegleichgewicht, *Wiener Berichte,* **76,** 373-435.

Boltzmann, L. (1878). Weitere Bemerkungen über einige Probleme der mechanishen Wärmetheorie, *Wiener Berichte,* **78**, S. 7-46.

Boltzmann, L. Part I, (1896). Part II, (1898). *Vorlesungen über Gastheorie,* Leipzig ; English translation (1964). *Lecture on Gas Theory,* Univ. of California Press, Berkeley and Los Angels.

Carnot, S. (1824). *Réflections sur la Pussance Motrice du feu et sur les Machines Properes â Développer cette Puissance,* Paris ; 英訳, (ed. Mendoza) (1960) *Reflections on the Motive Power of Fire and on Machines fitted to Develop that Power,* Dover ; 和訳, 広重徹 (1973).「カルノー・熱機

関の研究」みすず書房.

Clausius, R. (1850). Ueber die bewegende Krafte der Wärme und Gesetze die sich daraus für die Wärmelehre selbst ableiten lassen, *Poggendorff's Ann. d. Phys.*, **74**, 368-500; 英訳 (ed. Mendoza) (1960). Reflections on the Motive Power of Fire, 107-152. Dover.

Clausius, R. (1865). Ueber verschiedene für die Anwendung bequeme Formen der Hauptgleichungen der mechanischen Wärmetheorie, *Poggendorff's Ann. d. Phys.*, **125**, 353-400.

De Groot, S.R. (1963). *Thermodynamics of Irreversible Processes,* North-Holland.

Einstein, A. (1905). Über eine Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt, *Ann. d. Phys,* **17**, 132-148.

Einstein, A. (1909). Zum gegenwärtigen Stand die Strahlungsproblem, *Phys. Zeitschrift,* **16**, 185-193.

Gibbs, J.W. (1902). *Elementary Principles in Statistical Mechanics,* Yale Univ. Press, Dover reprint (1960).

池田和義 (1975).「統計熱力学」, 共立出版.

Jaynes, E.T. (1965). Gibbs vs Boltzmann Entropies, *Am. J. Phys.*, **33**, 391-398.

Jeans, J.H. (1905). On the partition of energy between matter and aether, *Phylosophical Magazine,* **10**, 91-98.

Kolmogorov, A.N. (1950). Unbiased estimates, *Izvestiya Akademii Nauk SSSR. Ser. Mat.*, **14**, 303-326; 英訳 Amer. Math. Soc. Trans. Ser. 1, No. **11** (1953), 144-170.

Kullback, S. (1959). *Information and Statistics,* Wiley.

Planck, M. (1900). Ueber irreversible Strahlungsvorgänge, *Ann. d. Phys.*, **1**, 69-122.

Planck, M. (1900a). Ueber eine Verbesserung der Wien'schen Specralgleichung, *Verhandungen der Deutschen Physiklischen Gesellschaft,* **2**, 202-204.

Planck, M. (1900b). Zur Theorie des Gesetzes der Energieverteilung im Normalspectrum, *Verhandungen der Deutschen Physikalischen Gesellschaft,* **2**, 237-245.

Planck, M. (1901). Ueber das Gesetz der Energieveiteilung im Normalspectrum, *Ann. d. Phys.* **4**, 553-563.

Planck M. (1930). *Einführung in die Theorie der Wärme*, Hirzel., Leipzig.

Rao, C.R. (1973). *Linear Statistical Inference and its Applications,* 2nd ed., Wiley.

Rayleigh, L. (1905). Remarks upon the laws of complete radiation, *Philosophical Magazine,* **49**, 539-540.

Sanov, I.N. (1957). On the probability of large deviations of random variables, *Math. Sb. N. S.*, **42** (84), 11-44, 英訳, *Selected Trans. in Math. Statist. and Prob. Inst. of Statist. and Amer. Math. Soc.* Vol. **1**, 213-244.

谷下市松 (1981)「工業熱力学, 基礎編」裳華房.

Wien, W. (1893). Eine neue Beziehung der Strahlung schwarzen Körper zum zweiten Hauptsatz der Wärmetheorie, *Sitzungsberichte der Berlinen Akademic*, 55-62.

Wien, W. (1896). Ueber die Energievertheilung in Emissionsspectrum eines schwarzen Körpers, *Wiedemann Annalen, d. Phys.* **58**, 662-669.

Entropy, Information and Statistics

Tadashi Matsunawa
(The Institute of Statistical Mathematics)

The concepts of entropy are discussed through the thermodynamical background, Boltzman-Plancks' principle and derivations of the Planck radiation formula in the classical statistical mechanics. It is pointed out that Planck's approach contains many useful results even to current statistical theory. For instance, in the principle behind his derivation we can find close relationships to the Rao-Blackwell Theorem and the Akaike Information Criterion in the modern statistics.

To recognize the meaning and role of the K-L information number relations between the information and the entropy are also investigated.