論　　文

大域的分類に基づく人的資源の質の向上

Qualitative Improvements of Manpower of Organizations Based on Their Global Classifications

坂　井　一　貴　大　﨑　佑　一　　　　　松　縄　　規
SAKAI Kazutaka, OOSAKI Yuuichi and MATSUNAWA Tadashi

要旨：　何らかの信頼可能な尺度に基づいて評価される，人的資源の簡潔かつ大局的な分類を，K－L情報量による大局的近似独立性を基盤に論じる．また，そのことを踏まえ，多くの分野で有用であると思われる，人的資源の質の向上についての応用に触れる．

キーワード：人的資源，大域的分類，センサードサンプル，近似独立性，質の向上．

１．はじめに

それなりの規模を有する組織において，その人的資源を高い水準に保持したいと考えるのは当然であり，組織のトップはその実現のために出来る限りの措置を講じる努力をしていると推察される．例えば，著者らのような教育に携わってきた者にとっては，学生達の各種成績を基に，①総合評価の大域的分類，引いては，②教育方法やカリキュラム改正による学力引き上げへの応用，等々が現実問題として存在する．以降では，これらの項目を動機として問題を論じる．なお，斯様な問題は，実は，松縄－坂井（2010）が扱った，不確実性下での経営情報学の複合科学的考察と部分的に関連している．事実，そこで触れた基礎モデルや発展モデル等の構築過程に関わる人的資源の質が重要に絡んでくることを考えれば自明である．関連して，様々な分野で人材の質向上が強く望まれてきた．それに対し，集団に一律の仕方で人材強化を図るのではなく，集団によらない強い理論的根拠を持つ，互いに影響をさほど与えない，複数層に分割し，層毎に適切なカリキュラムを提供することが，強化の実を上げるために効果的であろうと推察される．表題で‘大域的分類’としたのも，以後で扱う確率的命題の大域的強さに加え，上記の推察に起因している．このような背景を踏まえつつ，より一般的な状況や問題にも対処できるように，本稿では数理的扱いを中心に課題に取り組むことにする．

さかい　かずたか　　おおさき　ゆういち（経営情報学科）
まつなわ　ただし（情報・システム研究機構）

２．大域的分類のための理論的準備と結果

前章の課題①に応え，大域的分類に取り組むため，確率統計解析の枠組で考察する．まず，評価の低い層ⓐ，中間層ⓑ，および優秀な層ⓒが，ある意味で分割可能であることの理論的セットアップを行う．そのことを基盤に，我々が今後取り組もうとする問題をはじめ，多くの現実の問題にも適用可能な，シンプルな数学モデルを構築する．

$X_1 < X_2 < \cdots < X_N$を累積分布関数$F(x)$，密度関数$f(x)$を持つ一次元連続型分布からの大きさNの無作為標本に基づく順序統計量とする．今，このサイズNの無作為標本の小さいものP個と大きいものQ個を削除するとしよう．その結果として得られる確率変数をTypeⅡセンサードサンプル（censored sample）と呼ぶ（cf. David（1981））．対応して，$X_{P+1} < X_{P+2} < \cdots < X_{N-Q}$が考察対象の順序統計量となる．この状況で，確率ベクトルからなる集合

$$\{\mathbf{X}_{(n)}, \mathbf{Z}_{k(h)}, \mathbf{Y}_{(m)}\} \tag{2.1}$$

を考えよう．ここに，

$$\mathbf{X}_{(n)} = (X_{P+1}, \cdots, X_{P+n}),\ \mathbf{Z}_{k(h)} = (X_k, \cdots, X_{k+h-1}),\ \mathbf{Y}_{(m)} = (X_{N-Q-m+1}, \cdots, X_{N-Q}) \tag{2.2}$$

を表す．但し，

$$X_{P+1} < \cdots < X_{P+n} < X_k < \cdots < X_{k+h-1} < X_{N-Q-m+1} < \cdots < X_{N-Q}. \tag{2.3}$$

従って，以降，常に$P+n<k$，$k+h-1<N-Q-m+1$を仮定し，また，P, Q, n, m, k, hはNの増減と共に変化してよいものとする．すなわち，三部分からなる集合(2.1)は，上記TypeⅡセンサードサンプルからの$n=n(N)$lower extremes，$h=h(N)$middle terms および$m=m(N)$upper extremesによって構成される可変な結合確率ベクトルの集合を表す．

さて，結合確率ベクトル$(\mathbf{X}_{(n)}, \mathbf{Z}_{k(h)}, \mathbf{Y}_{(m)})$は実空間$R_{(n+h+m)}$上のLebesgue測度$\mu$に関して絶対連続と仮定する．また，$f_{\mathbf{X},\mathbf{Z},\mathbf{Y}}(\mathbf{x},\mathbf{z},\mathbf{y})$を結合確率ベクトル$(\mathbf{X}_{(n)}, \mathbf{Z}_{k(h)}, \mathbf{Y}_{(m)})$の同時密度関数とし，$f_{\mathbf{X}}(\mathbf{x}), f_{\mathbf{Z}}(\mathbf{z})$および$f_{\mathbf{Y}}(\mathbf{y})$を確率ベクトル$\mathbf{X}_{(n)}, \mathbf{Z}_{k(h)},$ および$\mathbf{Y}_{(m)}$それぞれの周辺密度関数とする．この時，次の結果が成立する．

定理 2.1.（主要な理論的結果）． 集合(2.1)は，$N\to\infty$の時

$$\frac{P+n}{N}\to 0,\ \frac{k}{N}\to\lambda,\ \frac{h}{N}\to\nu,\ \frac{Q+m}{N}\to 0,\quad (0<\lambda<1,\ 0\le\nu<1-\lambda), \tag{2.4}$$

が満たされるならば，また，その時に限り，次の意味で，大域的近似独立である：

$$\int_{R_{(n+h+m)}} f_{\mathbf{X},\mathbf{Z},\mathbf{Y}}(\mathbf{x},\mathbf{z},\mathbf{y}) \ln\frac{f_{\mathbf{X},\mathbf{Z},\mathbf{Y}}(\mathbf{x},\mathbf{z},\mathbf{y})}{f_{\mathbf{X}}(\mathbf{x}) f_{\mathbf{Z}}(\mathbf{z}) f_{\mathbf{Y}}(\mathbf{y})}\, d\mu \lessapprox \varphi(N), \tag{2.5}$$

ここに，$\varphi(N)$は，$\varphi(N)\to 0\quad (N\to\infty)$となる正の関数である．

注 2.1. 本定理および後述の系3.1は，Ikeda－Matsunawa（1970）で得た結果の拡張であり，以下の議論では，以前より精密な道具立てを準備して展開されている．

定理の証明の準備. 目標とする大域的近似独立性は，上記の問題の設定下に於いて，確率積分変換 $U_i := F(X_i)$, $(i=1,\cdots,N)$を利用して，(0,1)上の一様分布からの順序統計量を対象とすることによって証明できる．そこで，(2.1)に対応して，

(2.6) $$\left\{\mathbf{U}_{(n)}, \mathbf{W}_{k(h)}, \mathbf{V}_{(m)}\right\}$$

を考える．但し，

(2.7) $$\mathbf{U}_{(n)} = \left(U_{P+1},\cdots,U_{P+n}\right),\ \mathbf{W}_{k(h)} = \left(U_k,\cdots,U_{k+h-1}\right),\ \mathbf{V}_{(m)} = \left(U_{N-Q-m+1},\cdots,U_{N-Q}\right),$$

を表す．明らかに，要素間に次の関係が満たされる：

(2.8) $$0 < U_{P+1} < \cdots < U_{P+n} < U_k < \cdots < U_{k+h-1} < U_{N-Q-m+1} < \cdots < U_{N-Q} < 1 .$$

この時，確率ベクトル$\left(\mathbf{U}_{(n)}, \mathbf{W}_{k(h)}, \mathbf{V}_{(m)}\right)$は，$\left(\mathbf{u}_{(n)}, \mathbf{w}_{k(h)}, \mathbf{v}_{(m)}\right) = \left(u_{P+1},\cdots,u_{N-Q}\right)$に対し，同時密度関数は次式で与えられる：

(2.9) $$\begin{aligned} f_{\mathrm{U,W,V}}(\mathbf{u},\mathbf{w},\mathbf{v}) &= \frac{N!}{(k-P-n-1)!\,(N-Q-m-k-h+1)!} \\ &\times \left(u_k - u_{p+n}\right)^{k-P-n-1}\left(u_{N-Q-m+1} - u_{k+h-1}\right)^{N-Q-m-k-h+1}, \quad (0 < u_{P+1} < \cdots < u_{N-Q} < 1). \end{aligned}$$

また，$\mathbf{U}_{(n)}, \mathbf{W}_{k(h)}, \mathbf{V}_{(m)}$の周辺密度関数は，それぞれ，$\mathbf{u} = \left(u_{P+1},\cdots,u_{P+n}\right)$, $\mathbf{w} = \left(u_k,\cdots,u_{k+h-1}\right)$

および$\mathbf{v} = \left(u_{N-Q-m+1},\cdots,u_{N-Q}\right)$に対し，

(2.10) $$f_{\overline{\mathrm{U}}}(\mathbf{u}) = \frac{N!}{(N-P-n)!}\left(1-u_{P+n}\right)^{N-P-n}, \quad \left(0 < u_{P+1} < \cdots < u_{P+n} < 1\right),$$

(2.11) $$f_{\overline{\mathrm{W}}}(\mathbf{w}) = \frac{N!}{(k-1)!(N-k-h+1)!}u_k^{k-1}\left(1-u_{k+h-1}\right)^{N-k-h+1}, \ \left(0 < u_k < \cdots < u_{k+h-1} < 1\right),$$

(2.12) $$f_{\overline{\mathrm{V}}}(\mathbf{v}) = \frac{N!}{(N-Q-m)!}u_{N-Q-m+1}^{N-Q-m}, \quad \left(0 < u_{N-Q-m+1} < \cdots < u_{N-Q} < 1\right),$$

と与えられる．ここに，従って，我々は(2.5)式に代えて，条件(2.4)の下で，

(2.13) $$\begin{aligned} &I_N\left(\left(\mathbf{U}_{(n)}, \mathbf{W}_{k(h)}, \mathbf{V}_{(m)}\right) : \left(\mathbf{U}_{(n)}\right)\left(\mathbf{W}_{k(h)}\right)\left(\mathbf{V}_{(m)}\right)\right) \\ &\quad := \int_{R_{(n+h+m)}} f_{\mathrm{U,W,V}}(\mathbf{u},\mathbf{w},\mathbf{v}) \ln \frac{f_{\mathrm{U,W,V}}(\mathbf{u},\mathbf{w},\mathbf{v})}{f_{\overline{\mathrm{U}}}(\mathbf{u})\, f_{\overline{\mathrm{W}}}(\mathbf{w})\, f_{\overline{\mathrm{V}}}(\mathbf{v})}\, d\mu \lessapprox \varphi(N), \end{aligned}$$

$\lim_{N\to\infty}\varphi(N) = 0$となる正の関数$\varphi(N)$の存在を示せばよい．なお，(2.5)と(2.13)で同じ$\varphi(N)$を用いたのは，K－L情報量は変数の可測変換によって不変であることによってい

る．また，$\varphi(N)$が正関数であるとしたのは，K－L情報量の非負性による．

ここで，後々の計算および結果表現の簡素化のため，予め次の諸記号を導入しておく：

記号 2.1.

$$\begin{cases} \delta_0 := N,\ \delta_1 := k-P-n-1,\ \delta_2 := N-Q-m-k-h+1, \\ \delta_3 := N-P-n,\ \delta_4 := k-1,\ \delta_5 := N-k-h+1,\ \delta_6 := N-Q-m, \end{cases} \tag{2.14}$$

また，後で表現の短縮に有用となる，次の符号関数（signum function）を定義する．

$i=0,1,\cdots,6$に対して，

$$\mathrm{S}_i^{\pm} := \begin{cases} +2 & \mathit{if} \quad i=0, \\ +1 & \mathit{if} \quad i \le 2, \\ -1 & \mathit{if} \quad i \ge 3. \end{cases} \tag{2.15}$$

次の補題は，以後の情報量の精密な計算に不可欠となる：

補題 2.1. 有限な正の整数α, βに対し，次の公式が成立する：

$$\int_0^1 z^{\alpha-1}(1-z)^{\beta-1} \ln z \, dz = -\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \sum_{r=1}^{\beta} \frac{1}{\alpha-1+r}, \tag{2.16}$$

$$\int_0^1 z^{\alpha-1}(1-z)^{\beta-1} \ln(1-z) \, dz = -\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \sum_{r=1}^{\alpha} \frac{1}{\beta-1+r}. \tag{2.17}$$

証明概略. 前者を証明する．まずα, βを共に1以上の有限な実数とし，ベータ関数

$$\int_0^1 z^{\alpha-1}(1-z)^{\beta-1} \, dz = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

を考える．左辺の被積分関数および，そのαによる偏導関数は，補題の条件の下，連続であるから，左辺のαでの偏微分は積分記号化で行え，次のように求められる：

$$\frac{\partial}{\partial\alpha} \int_0^1 z^{\alpha-1}(1-z)^{\beta-1} \, dz = \int_0^1 \left(\frac{\partial}{\partial\alpha} z^{\alpha-1} \right) (1-z)^{\beta-1} \, dz = \int_0^1 z^{\alpha-1}(1-z)^{\beta-1} \ln z \, dz. \tag{2.18}$$

一方，上のベータ関数の右辺をαに関し偏微分すると，次のようになる：

$$\begin{aligned} \frac{\partial}{\partial\alpha} \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} &= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \left\{ \frac{\partial}{\partial\alpha} \ln\Gamma(\alpha) - \frac{\partial}{\partial\alpha} \ln\Gamma(\alpha+\beta) \right\} \\ &= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \left\{ \psi(\alpha) - \psi(\alpha+\beta) \right\}, \end{aligned} \tag{2.19}$$

ここに，$\psi(\cdot)$は，ディガンマ関数を表す．そこで，α, βが正の整数の場合を考える．上式最後の$\{\cdots\}$項に，Matsunawa（1976,（4.17）式）が与えた関係式を変形し，逆向きに利用すると次のように表現できる：

$$\psi(\alpha) - \psi(\alpha+\beta) = -\int_0^\infty e^{-\alpha t} \frac{1-e^{-\beta t}}{1-e^{-t}} \, dt = -\sum_{r=1}^{\beta} \frac{1}{\alpha-1+r}. \tag{2.20}$$

よって，(2.18)～(2.20)から，所要の(2.16)式を得る．(2.17)式も同様にして証明可能．■

次の補題も以下のK－L情報量の組織的計算に有用となる：

補題 2.2. $d_i = n_i - n_{i-1} - 1, \left(i = 1, 2, \cdots, \kappa + 1\right);\quad n_0 = 0, n_{\kappa+1} = N + 1$ に対し，

(2.21)
$$\begin{aligned}\underset{f_{\boldsymbol{\Omega}}\left(\boldsymbol{\omega}_{(\kappa)}\right)}{E}\left[\ln\left\{U_{n_i} - U_{n_i - 1}\right\}\right] &= -\sum_{r=1}^{N-d_i}\frac{1}{d_i + r} = \ln d_i - \ln N + \frac{1}{2}\left(\frac{1}{d_i} - \frac{1}{N}\right)\\ &-\frac{1}{12}\left\{\frac{1}{d_i\left(d_i+1\right)} - \frac{1}{N\left(N+1\right)} + \frac{1}{d_i\left(d_i+1\right)\left(d_i+2\right)} - \frac{1}{N\left(N+1\right)\left(N+2\right)}\right\}\\ &-\left\{\mathcal{T}\left(d_i\right)/d_i - \mathcal{T}\left(N\right)/N\right\},\end{aligned}$$

ここに，$\mathcal{T}\left(s\right)$は，

(2.22)
$$\mathcal{T}\left(s\right) = \sum_{r=3}^{\infty} c_{r+1}\left\{\left(s+1\right)\left(s+2\right)\cdots\left(s+r\right)\right\}^{-1},\ \left(s > 1\right),$$

で与えられる．但し，

(2.23)
$$c_r = \frac{1}{r}\int_0^1 t\left(1-t\right)\left(2-t\right)\cdots\left(r-1-t\right)dt,\ \left(r\text{：3以上の正整数}\right).$$

また，平均を取る密度関数は，$\boldsymbol{\omega}_{(\kappa)} = \left(u_1, \cdots, u_\kappa\right)$に対し，

(2.24)
$$f_{\boldsymbol{\Omega}}\left(\boldsymbol{\omega}_{(\kappa)}\right) = \frac{N!}{\prod_{i=1}^{\kappa+1}\left(d_i!\right)}\prod_{i=1}^{\kappa+1}\left(u_i - u_{i-1}\right)^{d_i},\ \left(0 \equiv u_0 < u_1 < \cdots < u_\kappa < u_{\kappa+1} \equiv 1\right)$$

で与えられるOrdered Dirichlet分布の密度関数を表す（cf. Wilks (1962), David (1981)）．

証明の方針． 式(2.20)の左辺に在る関数$\psi\left(\cdot\right)$に対して逆階乗級数表現を適用すれば所要の結果が得られる（cf. Matsunawa（1976））．■

また，(2.22)の$\mathcal{T}\left(s\right)$についての次の評価はK－L情報量の大きさの評価に役立つ：

補題 2.3. $s > 0$に対して，

(2.25)
$$\underline{\mathcal{T}}\left(s\right) := \frac{19}{120\left(s+1\right)^2\left(s+2\right)} \le \mathcal{T}\left(s\right) \le \frac{19}{90s\left(s+1\right)\left(s+2\right)} =: \overline{\mathcal{T}}\left(s\right).$$

証明概略． $\mathcal{T}\left(s\right)$の表現中の係数$c_r$について，次の不等式が成立する：

(2.26)
$$\frac{19}{60\left(r+1\right)}\Gamma\left(r\right) \le c_{r+1} \le \frac{19}{180\left(r+1\right)}\Gamma\left(r+1\right),\ \left(r \ge 3\right).$$

ここで，$\Gamma\left(\cdot\right)$は実ガンマ関数

(2.27)
$$\Gamma\left(x\right) = \int_0^\infty e^{-t}t^{x-1}dt, \left(x > 0\right)$$

を意味する．この不等式(2.26)と，Matsunawa(1976)の関連する証明方法と結果を用い，所要の評価不等式(2.25)を導出できる．■

補題 2.4.　$0<x<1$に対し

(2.28)
$$-\ln(1-x)=x+\frac{x^2}{2}+\mathfrak{S}(x),$$

ここに,

(2.29)
$$\frac{x^3/3}{1-\sqrt{2}\,x/2}<\mathfrak{S}(x)<\frac{x^3/3}{1-\sqrt{3}\,x/2}.$$

証明概略.　$0<x<1$に対し，Maclaurin級数

$$-\ln(1-x)=x+\frac{x^2}{2}+\frac{x^3}{3}+\frac{x^4}{4}+\cdots=x+\frac{x^2}{2}+\frac{x^3}{3}\left[1+\frac{3}{4}x+\frac{3}{5}x^2+\frac{3}{6}x^3+\cdots\right]$$

を上下から評価する．上式最後の大括弧部分は

$$[\cdots]<1+\frac{x}{2/\sqrt{3}}+\left(\frac{x}{2/\sqrt{3}}\right)^2+\left(\frac{x}{2/\sqrt{3}}\right)^3+\left(\frac{x}{2/\sqrt{3}}\right)^4+\cdots=\frac{1}{1-\sqrt{3}/2\cdot x}.$$

$$[\cdots]>1+\frac{x}{\sqrt{2}}+\left(\frac{x}{\sqrt{2}}\right)^2+\left(\frac{x}{\sqrt{2}}\right)^3+\left(\frac{x}{\sqrt{2}}\right)^4+\cdots=\frac{1}{1-x/\sqrt{2}}.$$

以上の等式および不等式から所要の(2.29)の関係式を得る．■

注 2.2.　上記補題の(2.28)における最終項は，対応するMaclaurinの定理のLagrange剰余項表現を用いると，$x^3/\{3(1-\xi)^3\}, (0<\xi<x)$ と与えられる．しかし，ξ が x に従属するため後々の理論誤差評価処理が複雑になる．補題 2.4の使用は，必要な近似精度を保ちながら，上記の複雑さから我々を解放する．

次の公式も本稿の複数個の階乗計算の過程で重要な道具となる．

補題 2.5.　(ガンマ関数の表現)．　$x>0$に対して,

(2.30)
$$\ln\Gamma(x+1)=\frac{1}{2}\ln 2\pi+\left(x+\frac{1}{2}\right)\ln x-x+\frac{1}{12x}-\mathcal{R}(x).$$

なお，剰余項とその大きさは次のように与えられる：

(2.31)
$$\mathcal{R}(x)=\sum_{i=2}^{\infty}a_{i+1}\{x(x+1)(x+2)\cdots(x+i)\}^{-1},$$

(2.32)
$$a_r=\frac{1}{r}\int_0^1 t(1-t)(2-t)\cdots(r-1-t)\left(\frac{1}{2}-t\right)dt,\ (r\geq 2),$$

(2.33)
$$\underline{\mathcal{R}}(x)<\mathcal{R}(x)<\overline{\mathcal{R}}(x).$$

ここに,

(2.34)
$$\underline{\mathcal{R}}(x)=\frac{1}{360x(x+1)(x+2)}\ \text{および}\ \overline{\mathcal{R}}(x)=\frac{1}{360x^3}.$$

証明略．上記不等式は形の簡潔さから採用した．複雑ではあるが，より精密な表現が与えられる(cf.松縄－武井(1999))．■

3．人的資源三層の大域的近似独立性

本章では，ここまでの準備の下，定理 2.1の証明を行う．

【条件(2.4)が十分であることの証明】　前章の式(2.9)～(2.12)および記号2.1の定義(2.14), (2.18)を用いて計算すると，(2.13)のK－L情報量は

(3.1)
$$I_N := I_N\left(\left(\mathbf{U}_{(n)}, \mathbf{W}_{k(h)}, \mathbf{V}_{(m)}\right) : \left(\mathbf{U}_{(n)}\right)\left(\mathbf{W}_{k(h)}\right)\left(\mathbf{V}_{(m)}\right)\right) =: A_N + B_N$$

と表せる．但し

(3.2)
$$A_N = -\sum_{i=0}^{6} \mathrm{S}_i^{\pm} \ln \delta_i!,$$

(3.3)
$$\begin{aligned} B_N = {} & \delta_1 \mathbf{E}\left[\ln\left(U_k - U_{P+n}\right)\right] + \delta_2 \mathbf{E}\left[\ln\left(U_{N-Q-m+1} - U_{k+h-1}\right)\right] - \delta_3 \mathbf{E}\left[\ln\left(1 - U_{P+n}\right)\right] \\ & - \delta_4 \mathbf{E}\left[\ln U_k\right] - \delta_5 \mathbf{E}\left[\ln\left(1 - U_{k+h-1}\right)\right] - \delta_6 \mathbf{E}\left[\ln U_{N-Q-m+1}\right], \end{aligned}$$

と表せる．なお，上式で$\mathbf{E}$は，次の意味の平均作用素を意味している：

(3.4)
$$\mathbf{E}\left[\ln(\cdot)\right] =: \int_{R_{(n+h+m)}} f_{\mathbf{U},\mathbf{W},\mathbf{V}}(\mathbf{u}, \mathbf{w}, \mathbf{v}) \cdot \ln(\cdot)\, d\mu.$$

式(3.1)の右辺の各項を逐次評価する：

まず，A_Nについて，階乗部分に補題2.5の近似公式を適用して，整理すると

(3.5)
$$A_N = -\sum_{i=0}^{6} \mathrm{S}_i^{\pm} \delta_i \ln \delta_i - \sum_{i=0}^{6} \mathrm{S}_i^{\pm} \left(\frac{1}{2} \ln \delta_i + \frac{1}{12\delta_i} - \mathcal{R}\left(\delta_i\right) \right).$$

次に，B_Nについて，補題 2.2を援用して，例えば(3.1)式右辺の第 2 項の平均値は

$$\mathbf{E}\left[\ln\left(U_k - U_{P+n}\right)\right] = \mathbf{E}\left[\ln\left(U_k - U_{P+n}\right)\right] \quad = -\sum_{r=1}^{N-\delta_1} \frac{1}{\delta_1 + r}$$

となる．同様にして他の平均値も計算できる．その結果,

(3.6)
$$B_N = \sum_{i=1}^{6} \left\{ \mathrm{S}_i^{\pm} \delta_i \left(-\sum_{r=1}^{N-\delta_i} \frac{1}{\delta_i + r} \right) \right\}.$$

さらに，この右辺有限和部分に補題 2.2を適用して整理すると,

(3.7)
$$B_N = \sum_{i=0}^{6} \mathrm{S}_i^{\pm} \delta_i \ln \delta_i - \frac{1}{12} \sum_{i=1}^{6} \mathrm{S}_i^{\pm} \left(\frac{2}{\delta_i + 1} - \frac{1}{\delta_i + 2} \right) - \sum_{i=0}^{6} \mathrm{S}_i^{\pm} \mathcal{T}\left(\delta_i\right).$$

と変形できる．(3.1)，(3.5)，(3.7)からK－L情報量は次のようになる：

(3.8)
$$I_N = -\frac{1}{2} \sum_{i=0}^{6} \mathrm{S}_i^{\pm} \ln \delta_i + \frac{1}{12} \sum_{i=0}^{6} \mathrm{S}_i^{\pm} \left(-\frac{1}{\delta_i} - \frac{2}{\delta_i + 1} + \frac{1}{\delta_i + 2} \right) + \sum_{i=0}^{6} \mathrm{S}_i^{\pm} \left\{ \mathcal{R}\left(\delta_i\right) - \mathcal{T}\left(\delta_i\right) \right\}.$$

上式右辺第 1 項は，$\ln N$，$\ln k$ の項が相殺でき，次のように表せる：

(3.9)
$$-\frac{1}{2}\sum_{i=0}^{6}\mathrm{S}_i^{\pm}\ln\delta_i=-\frac{1}{2}\left\{+\ln\left(1-\frac{\delta_0-\delta_2}{\delta_0}\right)-\ln\left(1-\frac{\delta_0-\delta_3}{\delta_0}\right)-\ln\left(1-\frac{\delta_0-\delta_5}{\delta_0}\right)\right.$$
$$\left.-\ln\left(1-\frac{\delta_0-\delta_6}{\delta_0}\right)+\ln\left(1-\frac{\delta_4+1-\delta_1}{\delta_4+1}\right)-\ln\left(1-\frac{1}{\delta_4+1}\right)\right\}.$$

右辺の各対数項は，条件(2.4)の下，$\delta_0(\equiv N)\to\infty$でゼロに収束するが，情報量の大きさを知るため，補題 2.4を用い剰余項$\mathfrak{S}(\cdot)$等を評価・整理すれば，(3.8)式は

(3.10)
$$-\frac{1}{2}\sum_{i=0}^{6}\mathrm{S}_i^{\pm}\ln\delta_i=\frac{1}{2}\left(\frac{P+n}{k}-\frac{P+n}{N}\right)+\frac{1}{k}+\frac{(Q+m)(k+h-1)}{2N^2}$$
$$+\frac{1}{4}\left\{\left(\frac{P+n}{k}\right)^2-\left(\frac{P+n}{N}\right)^2\right\}+\frac{P+n+1}{2k^2}+\frac{1}{2}\sum_{i=1}^{6}\mathrm{S}_i^{\pm}\mathfrak{S}(\rho_i)$$

と表せる．また，(3.8)右辺の中間項は，

(3.11)
$$+\frac{1}{12}\sum_{i=0}^{6}\mathrm{S}_i^{\pm}\left(-\frac{1}{\delta_i}-\frac{2}{\delta_i+1}+\frac{1}{\delta_i+2}\right)=-\frac{1}{6}\sum_{i=0}^{6}\mathrm{S}_i^{\pm}\frac{1}{\delta_i}+\frac{1}{6}\sum_{i=0}^{6}\mathrm{S}_i^{\pm}\frac{1}{\delta_i(\delta_i+1)(\delta_i+2)}$$

と表現できる．上式右辺の第 1 項は，さらに変形できて

(3.12)
$$-\frac{1}{6}\sum_{i=0}^{6}\mathrm{S}_i^{\pm}\frac{1}{\delta_i}=+\frac{1}{6}\left(-\frac{P+n}{k^2}+\frac{P+n}{N^2}\right)-\frac{1}{6}\sum_{i=1}^{6}\mathrm{S}_i^{\pm}\left(\frac{\rho_i^2}{\delta_i}\right)$$

と表せる．但し，上で，計算簡略化のために次のように置いた：

(3.13)
$$\rho_1=\frac{P+n+1}{k},\ \rho_2=\frac{Q+m+k+h-1}{N},\ \rho_3=\frac{P+n}{N},$$
$$\rho_4=\frac{1}{k},\ \rho_5=\frac{k+h-1}{N},\ \rho_6=\frac{Q+m}{N}.$$

(3.8), (3.10), (3.11), (3.12)から，次の表現が従う：

(3.14)
$$I_N=\frac{1}{2}\left(\frac{P+n+1}{k}-\frac{P+n}{N}\right)+\frac{(Q+m)(k+h-1)}{2N^2}+\frac{1}{4}\left\{\left(\frac{P+n}{k}\right)^2-\left(\frac{P+n}{N}\right)^2\right\}$$
$$+\frac{2(P+n)+3}{6k^2}+\frac{P+n}{6N^2}+\sum_{i=0}^{6}\mathrm{S}_i^{\pm}\left\{\mathcal{R}(\delta_i)-\mathcal{T}(\delta_i)+\frac{1}{6\,\delta_i(\delta_i+1)(\delta_i+2)}\right\}$$
$$+\frac{1}{2}\sum_{i=1}^{6}\mathrm{S}_i^{\pm}\left\{\mathfrak{S}(\rho_i)-\frac{\rho_i^2}{3\delta_i}\right\}=:\varphi(N).$$

条件(2.4)の下で，$I_N\to 0\ (N\to\infty)$が成立するから，(2.13)の$\varphi(N)$として(3.14)の右辺あるいはその上限を考えてよい．以上で，(2.4)が定理の十分条件であることが証明された．必要であることは，(3.14)等を考慮し，定理の対偶を示すことにより証明される．■

注 3.1. (3.14)のI_Nはexactな表現である．ただし，中に無限級数$\mathcal{R}(\delta_i),\mathcal{T}(\delta_i),\mathfrak{S}(\rho_i)$を

含んでいる．これらが，それぞれ$N \to \infty$でゼロに収束することが，それぞれ不等式(2.33), (2.25), (2.29)から分かる．また，これらの不等式から，有限なNに対して，I_Nの精密な上下限が与えられる（詳細略）．

系 3.1. 前章でのTypeⅡセンサードサンプルからの$n = n(N)$ lower extremesおよび$m = m(N)$ upper extremes $\{\mathbf{X}_{(n)}, \mathbf{Y}_{(m)}\}$は，$0 \le \nu < 1$として，$N \to \infty$で

(3.15) $$\frac{P+n}{N} \to \nu,\ \frac{Q+m}{N} \to 0,\ \text{または}\ \frac{P+n}{N} \to 0,\ \frac{Q+m}{N} \to \nu$$

が満たされる時，かつその時に限り，大域的近似独立である．この場合，関連する情報量は次のように評価できる：Nが十分に大きいとき，

(3.16) $$\int_{R_{(n+m)}} f_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y}) \ln \frac{f_{\mathbf{X},\mathbf{Y}}(\mathbf{x},\mathbf{y})}{f_{\mathbf{X}}(\mathbf{x}) f_{\mathbf{Y}}(\mathbf{y})}\ d\mu \lessapprox \frac{(P+n)(Q+m)}{2N^2} =: \psi(N). \quad \blacksquare$$

注 3.2. （**大域的近似独立の理論的根拠**）．上で示したK－L情報量に基づく近似独立性は，数理的にかなり強い概念である．実際，

(3.17) $$D_N := \sup_{E \in \mathbf{B}_{(n+h+m)}} \left| P^{(\mathbf{X}_{(n)}, \mathbf{Z}_{k(h)}, \mathbf{Y}_{(m)})}(E) - P^{(\mathbf{X}_{(n)})(\mathbf{Z}_{k(h)})(\mathbf{Y}_{(m)})}(E) \right| \le \sqrt{I_N/2}$$

となることが，数理統計学の一般論から成立する．ここで，不等式左辺の$P^{(\mathbf{X}_{(n)}, \mathbf{Z}_{k(h)}, \mathbf{Y}_{(m)})}$および$P^{(\mathbf{X}_{(n)})(\mathbf{Z}_{k(h)})(\mathbf{Y}_{(m)})}$は，それぞれ，第2章の結合確率ベクトル$(\mathbf{X}_{(n)}, \mathbf{Z}_{k(h)}, \mathbf{Y}_{(m)})$および$(\mathbf{X}_{(n)})(\mathbf{Z}_{k(h)})(\mathbf{Y}_{(m)})$の確率測度を表す．また，$\mathbf{B}$は実空間$R_{(n+h+m)}$上の部分集合のBorel集合体を表す．従って，K－L情報量が$I_N \to 0\ (N \to \infty)$ならば，$\mathbf{B}$に属する$R_{(n+h+m)}$の任意の部分集合$E$に対し定義される一様距離$D_N$に基づく，集合$\{\mathbf{X}_{(n)}, \mathbf{Z}_{k(h)}, \mathbf{Y}_{(m)}\}$の近似独立性が自動的に成立することになる．すなわち，第3章で得た近似独立性は将に大域的な性質を持つ強い結果と言える．

４．人的資源の質向上への応用

およそ，複数個の何らかの意味で比較可能な情報が得られる時，その構成要素間に順序付けを考える状況は少なくない．第１章の課題②もそのような順序付けがなされている状況下にあるものとして，若干の考察を行う．

近年，我が国の多くの組織で，組織メンバー，特に若年層の質の低下が問題視されている．それは，報道されてきた，生徒の学力の国際比較から，大いにありそうなことである．問題を簡単にするため，今，考察対象のメンバー数は相当数いるある組織で，成績評価が中庸な層の教育・訓練はそれなりの成果を上げている場合を想定する．そのような時，中庸層への教育・訓練は従来のそれの改善程度に行われるものとする．一方，その組織では，人的資源の質的向上のため，何らかの手立てを講じなければならないと

しよう．この時，前章の系 3.1が有用である．それは，(3.16)における$\psi(N)$は近似独立性の速度がかなり良いことを示していることによる．従って，サンプルサイズNがそれなりの大きさであれば独立性が成立しているとして議論をしても許されよう．すなわち，評価の低いn人からなる層と，m人からなる評価の高い層は，対応する二つの評価の分布は，大局的に独立とみなせる．この時，層のサイズをどの程度にするかは，系3.1が成立するための必要十分条件(3.15)を参考にするとよい．経験的にも感じていることであるが，nやmはNに比べ十分小さく取るのがよい．

このような認識の下，当該の組織は，この二つの層向けに，独立に質向上用のカリキュラムないしメニューを用意するのが諸観点から勧められる．なお，付言すると，前章のモデルは，現実でもよく起こり得るTypeⅡセンサードサンプルからのデータを扱えることである．例えば，大学教育などで，病気，成績不良あるいは経済的理由などでP人の学生が進級できないとか，大学をやめてしまうようなケースである．また，このようなプログラムを実施するに当たり，計画を立て，実施し，結果を分析の後，さらなる改善に結びつける指導者，スタッフの存在が欠かせない．上記の二層の場合に加え，評価が中庸の層の人的資源の質を改善をも図りたい時は，定理2.1の三層の大局的近似独立性の結果を利用して計画・実施することが考えられる．直感的にも予想できるが，この場合，上記の二層の時よりもサンプルサイズが大きいことが要求される．このことは，式(3.14)の$\varphi(N)$の表現にも表れている．なお，本稿では触れなかったが，サンプルをセンサー（＝打ち切り）する際に，個数ではなく，予め設定した割合（例えば下から５％）でセンサーするモデルを考えることがある．これをType Iセンサリングという．この場合，センサーされる個数が確率変数となり，数学的構造が複雑になる（cf. Ikeda－Matsunawa（1970））．

５．あとがき

本稿では，順序付け可能な人的資源に対し，その質の向上を図るために，それなりの根拠を持った，何らかの工夫が組織の内外から要請されている場合を想定して考察した．

もともとの研究動機は，ここ数年間に渡り学生の学力を観察してきた中で，学科の学生の学力向上に早急に取り組まなければならないという思いにあった．カリキュラムの改変時期とも連動させることも考慮した．しかしこの思いを実現させるにはかなりのハードルがあることも予想される．多くの大学では，主に経営上の理由から，計画実行にとって，望むだけの質と量の教職員を確保することが容易ではない．従って，実行可能と思われる範囲内で，合理的かつ効率的な計画を立てる必要がある．本稿で明らかにした理論と応用の可能性はその一助になるものと確信している．今後は，これまでの蓄積されたデータと，計画導入がされたならば，その後の関連データを基盤にした定量的な分析も加えた効果の検討とその先のさらなる改善を考えることが必要である．現在我が国では大学生の学力の評価は，基本的には，個々の大学の中に留まっている．しかし，そう

遠くない時期に，国内の大学である程度統一された，学力評価が実施される可能性もあり得る．我々は，そのようなことも見通して，地道に学生の学力を向上させる様々な努力をしてゆく必要を感じている．

参考文献

David, H. A. (1981). *Order Statistics 2nd ed.*, John Wiley & Sons, Inc., New York.

Ikeda, S. and Matsunawa, T. (1970). On asymptotic independence of order statistics, *Annals of Inst. Statist. Math.*, **22**, 435－449.

Matsunawa, T. (1976). Some inequalities based on inverse factorial series, *Annals of Inst.* Statist. Math., **28**, 291－305.

松縄規，武井智裕.（1999）. 不完全ガンマ関数比の評価不等式，*統計数理*, 47, 119－142.

松縄規，坂井一貴.（2010）. 不確実性下での経営情報学の複合科学的考察，*富山短期大学紀要*，**45**, 111－125.

Wilks, S. S.（1962）. *Mathematical Statistics 2nd ed.* John Wiley & Sons, Inc. New York.

（平成22年10月29日受付、平成22年11月11日受理）