近似主領域を考慮した修正情報量に基づく確率分布間の定量的一様近似理論

山田 智哉
Tomoya Yamada

学籍番号: 9 5 2 1 0 4

総合研究大学院大学
数物科学研究科
統計科学専攻

1995.10.30

目 次

第 1 章

緒言

現代統計学において互いに絶対連続な二つの確率分布間の近似を Kullback-Leibler (K-L) 情報量を用いて評価することは極く普通なこととなってきた. その理由として, 我々が扱う基本的な確率モデルが指数型分布族に含まれると想定される場合が多く, そのための理論的な計算が容易になることや, Fisher 情報量の最小化が尤度原理と密接に関わっていることなどが挙げられる.

しかしながら, このように K-L 情報量が統計学の中で多用されてきたにも拘らず, 理論的に見て現代統計学の中でのその用い方に検討を要する点があるように思える.

第 1 の問題点は, 原則として, 比較の対象となる二つの確率分布が正値をとる領域が全く同一となる場合のみを考えていることである. このような場合に該当しない時は, 解析的な都合からある種のコンベンションが置かれることが通常である. しかしこのような設定に納まらない状況は現実の近似問題で極く普通に起こり得る. そのような状況では K-L 情報量は発散し, 近似の尺度としては全く意味をなさないことになる.

第 2 の問題点はこれまでの多くの研究が, 通常の K-L 情報量の定義に現れる尤度比 (or 情報密度比) の分母子の取り方, 従ってどちらの分布で平均を取るのかということの重要さに重分注意を払ってこなかったことにある. 多くの場合, 計算の可能性や容易さを優先させるあまり, Boltzmann (1887) のエントロピーや Kullback (1959) および大偏差確率の理論における情報量の意味と役割に整合しない議論展開をせざるを得なかったように見受けられる. (cf. Matsunawa(1986), 松縄 (1994)). これらの問題点の克服のため, 特に第 1 の点を中心に研究に取り組み, 分布間の一様近似の理論建設に貢献したいということが本研究の

動機である.

このような背景の下, 本論文の目的は考察対象とする分布間の大域的近似問題において, 特に興味を持つ領域を近似の主要な領域として設定し, そこに分布する確率測度とそこでの K-L 情報量に準じた隔たりの量の大きさを評価することにより, 分布間の定量的な一様近似理論を展開することである. 簡単のため以下では, 上述の領域を**近似主領域**, そこでの分布間の隔たりの尺度を**修正情報量**と呼ぶことにする.

さて, X, Y を可測空間 $(R, \mathbf{B})$ 上で定義される確率変数とする. ここで, R は任意の抽象空間, $\mathbf{B}$ は R の部分集合の σ-集合体とする. また, X, Y に対する確率分布をそれぞれ P, Q とする. これらの分布はある σ-有限測度に関して絶対連続とし, それに関する密度をそれぞれ dP, dQ とする. このとき K-L 情報量は

$$I(Q,P) := \int_R \left(\ln \frac{dQ}{dP}\right) dQ$$

と表現される. 良く知られているようにこれは非負の量である. また, 大域的に P が Q で良く近似される時, $I(Q,P) \approx 0$ であることが期待される. しかし, この量は dP または dQ が 0 であったり, dQ/dP の値が局所的に, 例えば一次元分布の場合の裾の部分で, 大きく食い違う場合には K-L 情報量の大きさは無視できないものになる. この不都合を避けるために, 通常は P, Q に関してかなり制限を加えた問題設定の下で, K-L 情報量が発散しないことを暗黙の前提として, 議論を行なっているものが多い. しかしながら分布に対する一般的な制限は K-L 情報量による近似の適用範囲を狭めるものであり, 例えばカイ二乗分布の正規近似のような重要な近似問題も情報量の発散のために理論上扱えないことになる. ところが現代の統計的手法の利用事例の中には, その一般論の建設時に前提とされていた分布に関する条件を忘れ, 形式的に方法の適用を行なっているものも少なくないように見える. このようにして得られた結論が適切であるか否か大層気がかりな点である. このような懸念を少しでも取り除き, K-L 情報量の優れた面を残した近似問題への理論的な接近法の整備が望まれる. このことに対し, 本論文では分布間の近似の尺度として, K-L 情報量そのものではなく, Matsunawa (1982, 1986) にならって, 次の修正情報量と呼ぶ量を主として考

える.

$$I^*(Q,P;A) := \int_A \left(\ln \frac{dQ}{dP}\right) dQ, \quad (A \in \mathbf{B}).$$

これは可測集合 A 上での分布 P と Q の一種の隔たりを表していると見做せるが, 一般に $I(Q,P)$ と違って, 非負性を持たず, 時には負の値を取ることもあり得る. しかし後章で見るように, このことは本論文の議論では特に不都合を生じることはない.

ところで, このような性質を持つ修正情報量を利用して P と Q の間の一様近似を扱うために, 集合 A としてどのようなものを設定すれば良いかが重要な課題となる. このことに関して, 本論文では次の意味での一様性を考える:

$$D(Q,P;\mathbf{B}) = \sup_{E\in\mathbf{B}} \Big|Q(E) - P(E)\Big|.$$

Matsunawa(1986) は $I^*(Q,P;A)$ と $D(Q,P;\mathbf{B})$ の間に後述の定理 2.1.1で引用した不等式関係を与えた. その結果に基づけば, 分布 P が分布 Q で一様近似されるためには集合 A として大略

$$\min\{P(A), Q(A)\} \approx 1$$

となるものを考えればよいことがわかる. これは直感とも良く合致しており, 以下の章で議論する近似主領域と呼ぶ集合は本質的にこの基本性質を持つものである. またこのような状況は中心極限定理や二項分布のポアソン近似など様々な近似問題の本質的部分に内在している.

近似主領域の概念を意識した一様近似理論の必要性は Matsunawa(1982, 1986) で指摘された. そこでは, 理論上の基本的な近似主領域の例として, 任意の $\varepsilon \geq 0$ に対して

$$B_\varepsilon = \left\{x; \left|\ln \frac{dP(x)}{dQ(x)}\right| < \varepsilon, x \in R\right\}$$
$$S_\varepsilon = \left\{x; \max\left\{\left|\frac{dP(x)}{dQ(x)} - 1\right|, \left|\frac{dQ(x)}{dP(x)} - 1\right|\right\} < \varepsilon, x \in R\right\}$$

といった領域を考えた. これらは修正情報量が発散しないために「どちらかの密度が 0 である領域を含めない」という要請を満たすものであり, ε が微小であれば, 密度比が 1 に近

いという極く自然な条件を基盤として考えられている. 本論文では従来の結果を検討し, 改善することを試みる. 特に B_ϵ が理論的な近似主領域となることを, 後述の修正 Affinity の性質を援用して厳密に示し, 関連する条件などについても考察する.

ところで, 上記の領域 B_ϵ や S_ϵ は理論的に自然な近似主領域になり得るが, それらの条件を個別の近似問題の中で翻訳することは複雑になり過ぎて, 必ずしもわかり易く扱い易い条件を持つ領域を構成できないことも多い. このため具体的な問題に対して上記の理論的な条件を緩和化して柔軟に近似主領域を設定する工夫が必要となる. このことを実行する際に, 個々の近似問題における修正情報量の計算およびその定量的評価が近似主領域の構成に要求される条件を見出す時の指針となることが多い.

上述のように本論文で展開しようとする分布間の一様近似理論の骨子は, 問題に適した近似主領域の構成とそこに分布する確率および修正情報量の大きさの評価である. そのために本論文では, これらの量の評価に様々な段階で必要となる有効な各種の不等式を新たに与える. なお本論文の具体的問題の設定の段階において, 時には, 確率分布列間の近似を考えることになる. この場合には, 分布列の隔たりには例えば標本サイズがパラメータとして現れ, 適切な条件下で一様近似理論から近似誤差の定量的評価を含む漸近理論や極限理論がほぼ自動的に誘導できることになる.

本論文の構成は以下の通りである. まず第2章では一様近似理論に有用な良く知られた分布間の隔たりのいくつかの尺度とそれらに対応する修正量を導入し, 後者の諸量の間に成立する不等式関係を与える. なお第2章における議論は近似主領域のみならず一般の可測集合に対して成立するものとなっている. 第3章では近似主領域の構成について議論する. 領域 B_ϵ の近似主領域性について吟味される. また, ガンマ分布やベータ分布の一様正規近似, ordered Dirichlet 分布の多変量正規近似について, 関連する確率分布列と具体的な近似主領域の列について考える. 第4章においては, 近似主領域の概念を用いて多変量指数型分布族に関する一様近似を研究する. 分布の正準パラメータが微小変化した時に修正情報量がどの程度揺動するかを評価することは統計的にも, 物理的にも興味深いことである. Barron and Sheu (1991) はその労作の中で一変量の場合にこの問題を精力的に考察した.

しかし, その接近法は複雑であり, その問題の設定と情報量の定義の仕方は必ずしも広く自然科学一般に共通する認識を得られるとは思えない. このため本論文では第2章で得た一連の不等式を基に, 修正 affinity, 修正 W-divergence を援用して, 当該問題に対して Barron and Sheu よりも見通しよく系統的に修正情報量を評価する.

本論文で提出するいくつかの理論不等式や修正情報量, 近似主領域の確率の定量的評価には, より基本的な不等式を用いて誘導される. その内で特に $\left|t-1\right|$ と $\ln(1+1/t)$ の近似のための不等式とその証明を付録 A に与える. また本論文で考察した個別の近似問題についての各種計算の詳細を付録 B に与える.

第 2 章

分布間の隔たりの大小関係

2.1 はじめに

2 つの確率分布間の隔たりを測ることは, 統計問題において興味があり, いくつかの型が提案されている. 本章では, これらの量の定量的な評価を与える. また, その評価により得られる確率の評価も同時に考察する.

X,Y を可測空間 $(R, \mathbf{B})$ 上で定義される絶対連続な確率変数とする. ここで R は実空間, $\mathbf{B}$ は R のボレル集合族とする. X, Y に対応する確率分布を P, Q で表し, その密度関数を dP, dQ とする. 統計的推測を行なう場合, 分布間の隔たりを測るため, しばしば次の諸量が考えられる:

$$I(Q,P) := \int_R \left(\ln \frac{dQ}{dP}\right) dQ \tag{2.1.1}$$

$$\rho(P,Q) := \int_R \left(dP \cdot dQ\right)^{1/2} \tag{2.1.2}$$

$$V(Q,P) := \int_R \left|dQ - dP\right| \tag{2.1.3}$$

$$W(Q,P) := \int_R \frac{(dQ - dP)^2}{dP}. \tag{2.1.4}$$

これらの分布間の隔たりの尺度は Kullback-Leibler(K-L) 情報量, Matusita の Affinity, L_1 距離, Kagan の W-divergence(または χ^2-divergence) と呼ばれている.(cf. Kullback (1959), Matusita (1955), Kagan (1963)). 特に K-L 情報量は, パラメトリック分布の多くが指数型分布族に従い, 計算的に簡便であることから用いられることが多い. また Matusita の Affinity

は Hellinger 距離と同等の量である．これらの不等式関係は古くから知られていて，例えば

$$1 - \rho(P,Q) \le \frac{1}{2}V(Q,P) \le \sqrt{1 - \left(\rho(P,Q)\right)^2}$$
$$2\left(1 - \left(\rho(P,Q)\right)^2\right) \le I(Q,P) \le \ln\left(1 + W(Q,P)\right)$$

といった不等式関係がある．しかし第1章で触れたように全空間での積分とする分布間の隔たりの量を用いた場合，$I(Q,P)$, $W(Q,P)$ など，大域的に分布が近いにも関わらず分布の近さをうまく表さない量もある．そこで本章では，ある可測集合 $A \in \mathbf{B}$ を考え，この可測集合上で定義される分布間の隔たりの尺度を提案し，これらの非漸近的な定量評価を行なうことを目的とする．なお，本論文では，近似主領域を用いた分布の近似を考えているが，近似主領域の概念は第3章で詳しく述べることにして，本章では dP, dQ どちらも 0 でない任意の可測集合における近似理論を構築する．

本章では可測集合における分布間の隔たりの尺度として，以下の量を考える：可測集合 A に対して，

$$I^*(Q,P;A) := \int_A \left(\ln\frac{dQ}{dP}\right) dQ$$
$$I_a^*(Q,P;A) := \int_A \left|\ln\frac{dQ}{dP}\right| dQ$$
$$\rho^*(Q,P;A) := \int_A \left(dP \cdot dQ\right)^{1/2}$$
$$V^*(Q,P;A) := \int_A \left|dQ - dP\right|$$
$$W^*(Q,P;A) := \int_A \frac{(dQ-dP)^2}{dP}$$
$$\omega^*(Q,P;A) = \int_A \left|\frac{dQ}{dP} - 1\right| dQ$$
$$\Delta^*(Q,P;A) = \int_A \left|\frac{dQ}{dP} - 1\right|^2 dQ.$$

$\omega^*(Q,P;A)$, $\Delta^*(Q,P;A)$ は $V^*(Q,P;A)$, $W^*(Q,P;A)$ の積分をする基の分布関数を入れ換えた量である．これらの分布間の隔たりの尺度は非対称な量であるから，基の分布関数が違った場合，その量が示す意味合いは変わるので，これらがどのような大小関係になっているかが，興味ある．また，多くの統計家は K-L 情報量などの非対称な分布間の隔たりの分母子の取り方や基となる分布の取り方をその意味を考えずに扱っているように見えるが，その差

が本章の結果よりはっきり見通すことができる. 特に本章では, $I^*(Q,P;A)$, $I^*_\alpha(Q,P;A)$ を I^*型の分布間の隔たり, $\omega^*(Q,P;A)$, $V^*(Q,P;A)$ を L_1^*型の分布間の隔たり, $W^*(Q,P;A)$, $\Delta^*(Q,P;A)$ を W^*型の分布間の隔たりと呼ぶ.

具体的問題においては, I^*型の分布間の隔たりと他の分布間の隔たりを表す量との関係を与えることは, 非常に有効である. それは統計モデルの多くは指数型分布族に関係することから, I^*型の分布間の隔たりは, 具体的な積分計算, 特に多変量や多次元の問題において非常に有利になるためである. 本章ではこれらの関係を不等式による定量的な評価を試みる.

またこれらの量に関連して

$$D(Q,P;\mathbf{B}) = \sup_{E\in\mathbf{B}}\Big|Q(E)-P(E)\Big| \tag{2.1.5}$$

が考えられる. これを, 一様誤差と呼ぶ. 一様誤差は, I^*型の分布間の隔たりとの不等式関係として, Matsunawa(1986) の結果を要約し, 次の定理が与えられる.

定理 2.1.1 X,Yを可測空間 $(R,\mathbf{B})$ 上で定義される絶対連続な確率変数とする. ここで R は実空間, $\mathbf{B}$ は Rのボレル集合族とする. X, Yに対応する確率分布を P, Qで表し, その密度関数を dP, dQ とする. A を$\sup_A \dfrac{dQ}{dP} < \infty$ となる可測集合とする. この時, 式 (2.1.5) で定義される $D(Q,P;\mathbf{B})$ に対して次の不等式が得られる:

$$D(Q,P;\mathbf{B}) \le \frac{1}{2}u\left(\sup_A \frac{dQ}{dP}\right)\Big|I^*(Q,P;A)\Big| + \left\{1-\frac{Q(A)+P(A)}{2}\right\}$$
$$D(Q,P;\mathbf{B}) \ge \frac{1}{2}\left[l\left(\inf_A \frac{dQ}{dP}\right)\Big|I^*(Q,P;A)\Big| + \Big|Q(A)-P(A)\Big|\right].$$

ここで $l(t)$, $u(t)$ は補題 A.1.2の式 (A.1.6), (A.1.7) で定義されたものとする.

この不等式を用いれば, $I^*(Q,P;A)\to 0$ と $\min\{P(A),Q(A)\}\to 1$ であることを示せば一様誤差に基づく標題の一様近似が与えられる.

可測集合の例としては,

$$B_\varepsilon = \left\{X; \left|\ln\frac{dP(X)}{dQ(X)}\right| < \varepsilon, X\in R\right\}$$
$$S_\varepsilon = \left\{X; \max\left\{\left|\frac{dP(X)}{dQ(X)}-1\right|, \left|\frac{dQ(X)}{dP(X)}-1\right|\right\} < \varepsilon, X\in R\right\}$$

などが考えられる．本章では，任意の可測集合は暗に確率測度のほとんどを保持している集合であると考えている．この集合から逃げる確率も考慮に入れた大域的な近似理論を構築することが本論文の主目的で，それに準じた分布間の隔たりの大小関係を本章では考えている．第3章で，B_ϵは確率測度のほとんどを保持している領域（近似主領域）であることを明らかにするが，本章では B_ϵを含む任意の可測集合 A を考える．構造的には I^*型の分布間の隔たりに関しては B_ϵが，W^*型や L_1^*型は S_ϵが扱い易い．しかしこの 2 つの領域には 2.2 節で示すように $S_\epsilon \subseteq B_\epsilon$といった包含関係があり，余集合へ逃げる確率を小さくする領域を扱いたいため，B_ϵを用いた評価が好ましい．

本章の構成は以下の通りである．2.2節では，B_ϵと S_ϵの包含関係を与えると共に，これらの領域における量の評価を考える．2.3節では，修正 K-L 情報量とその他の分布間の隔たりとの不等式関係を与え，それらの結果により得られる不等式を与える．2.4節では，L_1^*型，W^*型の分布間の隔たりの評価を与える．

2.2 準備補題

本節では次節以降で必要となる B_ϵ, S_ϵといった領域における諸量の評価を行なう．B_ϵと S_ϵの関係については，$\left|\ln t\right| \leq \max\left\{\left|t-1\right|, \left|1/t-1\right|\right\}$の不等式から，次の結果を得る．

補題 2.2.1 $S_\epsilon \subseteq B_\epsilon$

注 2.1 本定理により本章で考えている可測集合 A は S_ϵも含む集合であると言える．また確率の逃げは $P(A-B_\epsilon) \leq P(A-S_\epsilon)$ を満たす．

注 2.2 本章は，任意の可測集合上の分布間の隔たりを考えている．しかし，その背景には，分布間の隔たりを表す量を近似主領域で用いて分布間の一様近似を行なうという発想がある．近似主領域からの確率測度の逃げを考慮にした評価を行なう時，確率測度の逃げはなるべく小さくすることが望まれる．よって問題によっては B_ϵを用いた方が好ましい．しかし B_ϵや S_ϵの構造上，L_1^*型，W^*型に関しては S_ϵの方が扱い易いという特徴もある．

次の系は Xが B_ϵ, S_ϵの補集合にある時の諸量の評価を表したものである．この結果は，確率測度の逃げを評価する際に有用で，次節以降で用いられる．

系 2.2.1 $X \in S_\varepsilon^c$の時,

$$\left|\frac{dQ(X)}{dP(X)} - 1\right| \geq \frac{\varepsilon}{1+\varepsilon}$$

証明.

$$\left|\frac{dP(X)}{dQ(X)} - 1\right| \geq \varepsilon \quad \text{or} \quad \left|\frac{dQ(X)}{dP(X)} - 1\right| \geq \varepsilon \Leftrightarrow \frac{1}{1+\varepsilon} \geq \frac{dQ(X)}{dP(X)} \quad \text{or} \quad \frac{dQ(X)}{dP(X)} \geq 1+\varepsilon$$
$$\Rightarrow \left|\frac{dQ}{dP} - 1\right| \geq \min\left\{\frac{\varepsilon}{1+\varepsilon}, \varepsilon\right\} = \frac{\varepsilon}{1+\varepsilon}$$

□

系 2.2.2 $X \in B_\varepsilon^c$の時,

$$\left|\frac{dQ(X)}{dP(X)} - 1\right| \geq 1 - e^{-\varepsilon}$$

証明.

$$\begin{aligned}
\left|\ln \frac{dQ(X)}{dP(X)}\right| \geq \varepsilon \quad &\Leftrightarrow \frac{dQ(X)}{dP(X)} \geq e^{\varepsilon} \geq 1 \quad \text{or} \quad \frac{dQ(X)}{dP(X)} \leq e^{-\varepsilon} \leq 1 \\
&\Leftrightarrow \frac{dQ(X)}{dP(X)} - 1 \geq e^{\varepsilon} - 1 \geq 0 \quad \text{or} \quad \frac{dQ(X)}{dP(X)} - 1 \leq e^{-\varepsilon} - 1 \leq 0 \\
&\Rightarrow \left|\frac{dQ(X)}{dP(X)} - 1\right| \geq \min\left\{e^{\varepsilon} - 1, 1 - e^{-\varepsilon}\right\} = 1 - e^{-\varepsilon}
\end{aligned}$$

□

最後に B_ε上での 2 つの分布関数の差の評価を与える.

補題 2.2.2 確率空間 $(R^{n\times n}, B^{n\times n})$ で定義される分布関数 P, Q に対して

$$\left|P(B_\varepsilon) - Q(B_\varepsilon)\right| \leq 1 - e^{-\varepsilon}. \tag{2.2.1}$$

注 2.3 補題2.2.2の結果より, $\varepsilon \to 0$ とすると, 分布関数間の差は 0 になる. εが 0 であるということは, B_ε上では全ての密度が等しいということを意味しているので, この結果は当然のことといえる.

証明. B_εの条件を書き直すと密度関数 dPと dQ の関係は以下のように書き換えられる:

$$e^{-\varepsilon}dQ \le dP \le e^{\varepsilon}dQ$$
$$e^{-\varepsilon}dP \le dQ \le e^{\varepsilon}dP.$$

この関係について B_εで積分すると, 分布関数 Pと Q の関係は

$$e^{-\varepsilon}Q(B_\varepsilon) \le P(B_\varepsilon) \le \min\left\{e^{\varepsilon}Q(B_\varepsilon), 1\right\}$$
$$e^{-\varepsilon}P(B_\varepsilon) \le Q(B_\varepsilon) \le \min\left\{e^{\varepsilon}P(B_\varepsilon), 1\right\}$$

と与えられる. これにより絶対誤差は次の不等式で与えられる:

$$\begin{aligned}\left|P(B_\varepsilon) - Q(B_\varepsilon)\right| &= \max\{P(B_\varepsilon), Q(B_\varepsilon)\}\left(1 - \min\left\{\frac{Q(B_\varepsilon)}{P(B_\varepsilon)}, \frac{P(B_\varepsilon)}{Q(B_\varepsilon)}\right\}\right) \\ &\le 1 - e^{-\varepsilon}.\end{aligned}$$

□

2.3 修正 K-L 情報量を用いた評価

全空間で平均をとる K-L 情報量

$$I(Q, P) := \int_R (\ln \frac{dQ}{dP}) dQ$$

は, Kullback が体系付けて以降, 多くの論文が発表されてきた. その理由としては, 統計モデルのほとんどが従う指数型分布族に対して非常に計算が簡便であることが挙げられる. この K-L 情報量は, 非負の値を持ち, $I(Q, P) = 0$ で分布が一致することが知られている. このことから, K-L 情報量の最小化原理が得られ, 理論, 応用両面から多くの役割を担ってきた. そこで本節では, 近似主領域の概念に基づいて修正した K-L 情報量と, それに対応する他の分布間の隔たりを表す量との不等式関係を与える.

しかし本論文で扱う修正された K-L 情報量は必ずしも非負になるとは限らない. このため修正 K-L 情報量の評価を与える時は, 上下からの評価を与える必要がある. 本節では, 初めに修正 Affinity と修正 K-L 情報量との不等式評価を与え, その後, 修正 W-divergence と修正 K-L 情報量との不等式評価を考える.

修正 Affinity と修正 K-L 情報量との間には，次の不等式関係が与えられる．

定理 2.3.1 任意の可測集合 $A \in B^{n\times n}$ に対して，

$$\rho^*(P,Q;A) \geq \max\left\{P(A)\cdot\exp\left[\frac{-1}{2P(A)}I^*(P,Q;A)\right], Q(A)\cdot\exp\left[\frac{-1}{2Q(A)}I^*(Q,P;A)\right]\right\}. \tag{2.3.1}$$

証明． ここでは Jensen の不等式を用いて，

$$\rho^*(P,Q;A) = P(A)\int_A\left(\frac{dQ}{dP}\right)^{\frac{1}{2}}dP/P(A) = P(A)\int_A \exp\left\{\frac{1}{2}\ln\frac{dQ}{dP}\right\}\cdot dP/P(A)$$

$$\geq P(A)\exp\left\{\frac{1}{2P(A)}\int_A \ln\frac{dQ}{dP}\cdot dP\right\} = P(A)\exp\left\{-\frac{1}{2P(A)}I^*(P,Q;A)\right\}$$

と表すことができる．また同様に

$$\rho^*(P,Q;A) \geq Q(A)\cdot\exp\left[\frac{-1}{2Q(A)}I^*(Q,P;A)\right]$$

となることから

$$\rho^*(P,Q;A) \geq \max\left\{P(A)\cdot\exp\left[\frac{-1}{2P(A)}I^*(P,Q;A)\right], Q(A)\cdot\exp\left[\frac{-1}{2Q(A)}I^*(Q,P;A)\right]\right\}.$$

これにより定理 2.3.1が証明される． □

系 2.3.1 任意の可測集合 A に対して，定理 2.3.1 を用いると，以下の一連の不等式を得る：

$$I^*(P,Q;A) \geq -2P(A)\cdot\ln\frac{\rho^*(P,Q;A)}{P(A)} \tag{2.3.2}$$

$$I^*(P,Q;A) \geq -2P(A)\cdot\ln\frac{\sqrt{P(A)\cdot Q(A)}}{P(A)} = P(A)\cdot\ln\frac{P(A)}{Q(A)} \tag{2.3.3}$$

$$I^*(P,Q;A) \geq P(A)\left(1-\frac{Q(A)}{P(A)}\right) = P(A)-Q(A). \tag{2.3.4}$$

注 2.4 系 2.3.1の一連の不等式の中で特に式 (2.3.4) をみると修正 K-L 情報量の下界として分布関数の差により評価される．また $A = R^{n\times n}$ とすると，下界が 0 になる．このことから K-L 情報量の最小化原理が得られ，いろいろな応用が考えられている．しかし修正 K-L 情報量は必ずしも非負とはならない．この点で修正 K-L 情報量は上下からの評価を与えねばならず，多少の困難が伴う．

系 2.3.2

$$\rho^*(P,Q;A) \geq \rho^*(P,Q;B_\varepsilon) \geq e^{-\varepsilon/2}\max\Big\{P(B_\varepsilon),Q(B_\varepsilon)\Big\} \tag{2.3.5}$$

証明. 修正 K-L 情報量 $I^*(Q,P;B_\varepsilon)$ の上界として

$$I^*(Q,P;B_\varepsilon) \leq \int_{B_\varepsilon}\left|\ln\frac{dQ}{dP}\right|dQ \leq \varepsilon Q(B_\varepsilon) \tag{2.3.6}$$

が成り立つ. これより (2.3.1) を用いて, $\rho^*(P,Q;B_\varepsilon)$ の下界は

$$\begin{aligned}\rho^*(P,Q;B_\varepsilon) &\geq \max\Big\{P(B_\varepsilon)e^{-\varepsilon/2},Q(B_\varepsilon)e^{-\varepsilon/2}\Big\}\\ &\geq e^{-\varepsilon/2}\max\Big\{P(B_\varepsilon),Q(B_\varepsilon)\Big\}\end{aligned}$$

を得る. □

定理 2.3.1は Jensen の公式を用いて評価を行なった. 定理 2.3.1をさらに一般化することにより, 次の補題が導かれる.

補題 2.3.1

$$\begin{aligned}&\rho^*(P,Q;B_\varepsilon)\\ &\quad\geq \exp\left\{-\frac{I^*}{2}P(B_\varepsilon)\right\}\left[P(B_\varepsilon)+\frac{1}{8}\left(1+\frac{I^*}{2}\right)\left\{\int_{B_\varepsilon}\left(\ln\frac{dP(X)}{dQ(X)}\right)^2\cdot dP-(I^*)^2\right\}\right.\\ &\quad\left.-(I^*)^2\left\{1-P(B_\varepsilon)\right\}-\frac{1}{6}\left\{\int_{B_\varepsilon}\left(\ln\frac{dP(X)}{dQ(X)}\right)^3\cdot dP-(I^*)^3P(B_\varepsilon)\right\}\right]\end{aligned} \tag{2.3.7}$$

$$\begin{aligned}&I^*(P,Q,B_\varepsilon) \geq -2\ln\rho^*(P,Q;B_\varepsilon)\cdot P(B_\varepsilon)+Q(B_\varepsilon)/\Big\{\rho^*(P,Q;B_\varepsilon)\Big\}^2\\ &-\frac{2}{3}e^{-\varepsilon/2}Q(B_\varepsilon)/\Big\{\rho^*(P,Q;B_\varepsilon)\Big\}^3\end{aligned} \tag{2.3.8}$$

証明. $k(y) := e^{-y/2}$とし, $X = \ln\left\{\dfrac{dP(Y)}{dQ(Y)}\right\}$とする. この時, 修正 K-L 情報量と修正 Affinity はそれぞれ

$$\begin{aligned}I^*(P,Q;B_\varepsilon) &= \int_{B_\varepsilon}\ln\frac{dP(Y)}{dQ(Y)}\cdot dP(Y) := E^*\big[X\big]\\ \rho^*(P,Q;B_\varepsilon) &= \int_{B_\varepsilon}\sqrt{\frac{dP(Y)}{dQ(Y)}}\cdot dP(Y) := E^*\left[\exp\left(-\frac{1}{2}X\right)\right]\end{aligned}$$

と表される．ここで

$$E^*[h(x)] = \int_{B_\epsilon} h(x) dP(x)$$

とする．また2次, 3次のモーメントは, それぞれ

$$\begin{aligned}
E^*\left[(X - I^*)^2\right] &= \int_{B_\epsilon} \left(\ln \frac{dP(Y)}{dQ(Y)} - I^*\right)^2 dP(Y) \\
&= \int_{B_\epsilon} \left\{\ln \frac{dP}{dQ}\right\}^2 dP - 2I^* \int_{B_\epsilon} \ln \frac{dP}{dQ} \cdot dP + I^{*2} \int_{B_\epsilon} dP \\
&= \left\{\int_{B_\epsilon} \left(\ln \frac{dP}{dQ}\right)^2 \cdot dP - (I^*)^2\right\} - (I^*)^2 \left\{1 - P(B_\epsilon)\right\} \\
E^*\left[(X - I^*)^3\right] &= \int_{B_\epsilon} \left\{\ln \frac{dP}{dQ} - I^*\right\}^3 dP \\
&= \int_{B_\epsilon} \left(\ln \frac{dP}{dQ}\right)^3 dP - 3I^* \int_{B_\epsilon} \left(\ln \frac{dP}{dQ}\right)^2 dP + 3(I^*)^2 \int_{B_\epsilon} \left(\ln \frac{dP}{dQ}\right) dP - (I^*)^3 \int_{B_\epsilon} dP \\
&= \left\{\int_{B_\epsilon} \left(\ln \frac{dP}{dQ}\right) - (I^*)^3 P(B_\epsilon)\right\} - 3I^* \left\{\int_{B_\epsilon} \left(\ln \frac{dP}{dQ}\right)^2 dP - (I^*)^2\right\}
\end{aligned}$$

と表される．ここで $k(y)$ の展開式として

$$\begin{aligned}
&k(y) \\
&= k(I^*) + \frac{k^{(1)}(I^*)}{1!}(y - I^*) + \frac{k^{(2)}(I^*)}{2!}(y - I^*)^2 + \frac{k^{(3)}(I^*)}{3!}(y - I^*)^3 + \frac{k^{(4)}(\xi)}{4!}(y - I^*)^4
\end{aligned}$$

を用いる．ここで$\xi \in (y, I^*)$, $k^{(1)}(y) = -\frac{1}{2}e^{-y/2}$, $k^{(2)}(y) = \frac{1}{4}e^{-y/2}$, $k^{(3)}(y) = -\frac{1}{8}e^{-y/2}$, $k^{(4)}(y) = \frac{1}{16}e^{-y/2}$である．これらは交互に符号を変えるので, 修正 K-L 情報量による修正 Affinity の評価としては

$$\begin{aligned}
&\rho^*(P, Q; B_\epsilon) \geq e^{-I^*/2} P(B_\epsilon) + \frac{1}{8} e^{-I^*/2} E^*\left[(X - I^*)^2\right] - \frac{1}{48} e^{-I^*/2} E^*\left[(X - I^*)^3\right] \\
&= \exp\left\{-\frac{I^*}{2} P(B_\epsilon)\right\} \left[P(B_\epsilon) + \frac{1}{8}\left(1 + \frac{I^*}{2}\right) \left\{\int_{B_\epsilon} \left(\ln \frac{dP(X)}{dQ(X)}\right)^2 \cdot dP - (I^*)^2\right\}\right. \\
&\left. -(I^*)^2 \left\{1 - P(B_\epsilon)\right\} - \frac{1}{6} \left\{\int_{B_\epsilon} \left(\ln \frac{dP(X)}{dQ(X)}\right)^3 \cdot dP - (I^*)^3 P(B_\epsilon)\right\}\right]
\end{aligned}$$

を得る．これにより式 (2.3.7) が証明される．また式 (2.3.8) については, $h(y) = \ln y$,

$X := \sqrt{dQ(Y)/dP(Y)}$とした時, 修正 K-L 情報量と修正 Affinity はそれぞれ

$$I^*(P,Q;B_\epsilon) = \int_{B_\epsilon} \ln \frac{dP(Y)}{dQ(Y)} \cdot dP(Y) := -2E^*\Big[h(X)\Big]$$
$$\rho^*(P,Q;B_\epsilon) = \int_{B_\epsilon} \sqrt{\frac{dP(Y)}{dQ(Y)}} \cdot dP(Y) := E^*\Big[X\Big]$$

と表される. また2次, 3次のモーメントは, それぞれ

$$\begin{aligned}
&E^*\Big[(X-\rho^*)^2\Big] = \int_{B_\epsilon} dP\left(\sqrt{dQ/dP} - \rho^*\right)^2 = \int_{B_\epsilon} dQ - 2\rho^* \int_{B_\epsilon} (dPdQ)^{1/2} + \rho^{*2} \int_{B_\epsilon} dP \\
&= Q(B_\epsilon) - 2\rho^{*2} + \rho^{*2}P(B_\epsilon) = Q(B_\epsilon) - \rho^{*2} - \rho^{*2}\Big\{1 - P(B_\epsilon)\Big\} \\
&E^*\Big[(X-\rho^*)^3\Big] = \int_{B_\epsilon} dP\left(\sqrt{\frac{dP}{dQ}} - \rho^*\right)^3 \\
&= \int_{B_\epsilon} \sqrt{\frac{dQ}{dP}} dQ - 3\rho^* \int_{B_\epsilon} dQ + 3(\rho^*)^2 \int_{B_\epsilon} (dP \cdot dQ)^{1/2} - (\rho^*)^3 \int_{B_\epsilon} dP \\
&\leq e^{-\epsilon/2}Q(B_\epsilon) - \rho^{*3}Q(B_\epsilon) + \rho^{*3}\Big\{3 - P(B_\epsilon)\Big\}
\end{aligned}$$

と表される. ここで $h(y)$ の展開式として

$$\begin{aligned}
&h(y) \\
&= h(\rho^*) + \frac{h^{(1)}(\rho^*)}{1!}(y-\rho^*) + \frac{h^{(2)}(\rho^*)}{2!}(y-\rho^*)^2 + \frac{h^{(3)}(\rho^*)}{3!}(y-\rho^*)^3 + \frac{h^{(4)}(\xi)}{4!}(y-\rho^*)^4
\end{aligned}$$

を用いる. ここで$\xi \in (y, \rho^*)$, $h^{(1)}(y) = y^{-1}$, $h^{(2)}(y) = -y^{-2}$, $h^{(3)}(y) = 2y^{-3}$, $h^{(4)}(y) = -6y^{-4}$ である. これらは交互に符号を変えるので, 修正 Affinity による修正 K-L 情報量の評価としては

$$\begin{aligned}
&I^*(P,Q,B_\epsilon) \geq -2\ln\rho^*(P,Q;B_\epsilon)P(B_\epsilon) + Q(B_\epsilon)/\Big\{\rho^*(P,Q;B_\epsilon)\Big\}^2 \\
&-\frac{2}{3}e^{-\epsilon/2}Q(B_\epsilon)/\Big\{\rho^*(P,Q;B_\epsilon)\Big\}^3
\end{aligned}$$

を得る. これにより補題 2.3.1が証明された. □

最後に補題 2.3.1の結果よりϵを用いた評価を考える. 次の系で扱う領域の条件で, 下限を置くことは本質的ではないが, より厳密に評価をするため, 下限を置いている. 実際には $\epsilon_1 = 0$ と置けば良い.

系 2.3.3

$$B'_\varepsilon := \left\{ X; 0 \le \varepsilon_1 \le \left| \ln \frac{dQ(X)}{dP(X)} \right| \le \varepsilon_2 < 2, X \in R^{n\times n} \right\} \text{の時}$$

$$\begin{aligned}
\rho^*(P,Q;B'_\varepsilon) \ \ge \ & \frac{1}{8}\exp\left\{-\frac{\varepsilon_2}{2}P(B_\varepsilon)\right\}\left[\left(8+\varepsilon_1^2-\frac{1}{6}\varepsilon_2^3\right)P(B_\varepsilon)-\left(\frac{\varepsilon_1^2\varepsilon_2}{2}+2\varepsilon_2^2\right)P^2(B_\varepsilon)\right. \\
& \left. +\left(-\frac{\varepsilon_2^3}{2}+\varepsilon_2^2\right)P^3(B_\varepsilon)-\varepsilon_2^3P^4(B_\varepsilon)\right]
\end{aligned}$$

証明. B_εに対して,

$$\begin{aligned}
& \varepsilon_1^2 P(B_\varepsilon) \le \int_{B_\varepsilon}\left(\ln\frac{dP(X)}{dQ(X)}\right)^2 \cdot dP \le \varepsilon_2^2 P(B_\varepsilon) \\
& 0 \le \varepsilon_1 P(B_\varepsilon) \le \left| I^*(P,Q,B_\varepsilon) \right| \le \varepsilon_2 P(B_\varepsilon) \\
& 1+\frac{\varepsilon_2}{2}P(B_\varepsilon) \ge 1+\frac{I^*}{2} \ge 1-\frac{\varepsilon_2}{2}P(B_\varepsilon)(\ge 0) \\
& -\varepsilon_2^3 P(B_\varepsilon) \le \int_{B_\varepsilon}\left(\ln\frac{dP(X)}{dQ(X)}\right)^3 \cdot dP \le \varepsilon_2^3 P(B_\varepsilon) \\
& P^3(B_\varepsilon) \le P^2(B_\varepsilon) \le P(B_\varepsilon) \le 1
\end{aligned}$$

となることから

$$\begin{aligned}
& \left(1+\frac{I^*}{2}\right)\left\{\int_{B_\varepsilon}\left(\ln\frac{dP(X)}{dQ(X)}\right)^2 \cdot dP-(I^*)^2\right\} \\
& \ge \left\{1-\frac{\varepsilon_2}{2}P(B_\varepsilon)\right\}\varepsilon_1^2P(B_\varepsilon)-\left\{1+\frac{\varepsilon_2}{2}P(B_\varepsilon)\right\}\varepsilon_2^2P^2(B_\varepsilon) \\
& -(I^*)^2\left\{1-P(B_\varepsilon)\right\} \ge -\varepsilon_2^2P^2(B_\varepsilon)\left\{1-P(B_\varepsilon)\right\} \\
& -\frac{1}{6}\left\{\int_{B_\varepsilon}\left(\ln\frac{dP(X)}{dQ(X)}\right)^3 \cdot dP-(I^*)^3P(B_\varepsilon)\right\} \ge -\frac{1}{6}\left\{\varepsilon_2^3P(B_\varepsilon)-\varepsilon_2^3P^4(B_\varepsilon)\right\} \\
& \exp\left\{-\frac{I^*}{2}\right\} \ge \exp\left\{-\frac{\varepsilon_2}{2}P(B_\varepsilon)\right\}
\end{aligned}$$

を得る. この結果を式 (2.3.7) に代入することにより,

$$\begin{aligned}
& \rho^*(P,Q;B_\varepsilon) \\
& \quad \ge \exp\left\{-\frac{I^*}{2}P(B_\varepsilon)\right\}\left[P(B_\varepsilon)+\frac{1}{8}\left(1+\frac{I^*}{2}\right)\left\{\int_{B_\varepsilon}\left(\ln\frac{dP(X)}{dQ(X)}\right)^2 \cdot dP-(I^*)^2\right\}\right.
\end{aligned}$$

$$-(I^*)^2\Big\{1-P(B_\varepsilon)\Big\} - \frac{1}{6}\left\{\int_{B_\varepsilon}\left(\ln\frac{dP(X)}{dQ(X)}\right)^3\cdot dP - (I^*)^3P(B_\varepsilon)\right\}\Bigg]$$
$$\ge \frac{1}{8}\exp\left\{-\frac{\varepsilon_2}{2}P(B_\varepsilon)\right\}\Bigg[\left(8+\varepsilon_1^2-\frac{1}{6}\varepsilon_2^3\right)P(B_\varepsilon) - \left(\frac{\varepsilon_1^2\varepsilon_2}{2}+2\varepsilon_2^2\right)P^2(B_\varepsilon)$$
$$+\left(-\frac{\varepsilon_2^3}{2}+\varepsilon_2^2\right)P^3(B_\varepsilon) - \varepsilon_2^3P^4(B_\varepsilon)\Bigg]$$

また $e^{-\frac{\varepsilon}{2}} \le \rho^*(P,Q;B_\varepsilon) \le e^{\frac{\varepsilon}{2}}$を式 (2.3.8) に代入することにより,

$$\begin{aligned}
I^*(P,Q,B_\varepsilon) &\ge -2\ln\rho^*(P,Q;B_\varepsilon)P(B_\varepsilon) + Q(B_\varepsilon)/\Big\{\rho^*(P,Q;B_\varepsilon)\Big\}^2\\
&\quad -\frac{2}{3}e^{-\varepsilon/2}Q(B_\varepsilon)/\Big\{\rho^*(P,Q;B_\varepsilon)\Big\}^3 - 4 + \frac{5}{3}P(B_\varepsilon)\\
&\ge -2\ln e^{\varepsilon/2}\cdot P(B_\varepsilon) + 3Q(B_\varepsilon)e^{-\varepsilon} - \frac{2}{3}e^{-\varepsilon/2}Q(B_\varepsilon)e^{\frac{3}{2}\varepsilon} - 4 + \frac{5}{3}P(B_\varepsilon)\\
&= \left(\frac{5}{3}-\varepsilon\right)P(B_\varepsilon) + \left(3e^{-\varepsilon}-\frac{2}{3}e^{\varepsilon}\right)Q(B_\varepsilon) - 4
\end{aligned}$$

を得る. □

系 2.3.4

$$I^*(P,Q,B_\varepsilon) \ge \left(\frac{5}{3}-\varepsilon\right)P(B_\varepsilon) + \left(3e^{-\varepsilon}-\frac{2}{3}e^{\varepsilon}\right)Q(B_\varepsilon) - 4$$

証明. $e^{-\frac{\varepsilon}{2}} \le \rho^*(P,Q;B_\varepsilon) \le e^{\frac{\varepsilon}{2}}$を式 (2.3.8) に代入することにより,

$$\begin{aligned}
I^*(P,Q,B_\varepsilon) &\ge -2\ln\rho^*(P,Q;B_\varepsilon)P(B_\varepsilon) + Q(B_\varepsilon)/\Big\{\rho^*(P,Q;B_\varepsilon)\Big\}^2\\
&\quad -\frac{2}{3}e^{-\varepsilon/2}Q(B_\varepsilon)/\Big\{\rho^*(P,Q;B_\varepsilon)\Big\}^3 - 4 + \frac{5}{3}P(B_\varepsilon)\\
&\ge -2\ln e^{\varepsilon/2}\cdot P(B_\varepsilon) + 3Q(B_\varepsilon)e^{-\varepsilon} - \frac{2}{3}e^{-\varepsilon/2}Q(B_\varepsilon)e^{\frac{3}{2}\varepsilon} - 4 + \frac{5}{3}P(B_\varepsilon)\\
&= \left(\frac{5}{3}-\varepsilon\right)P(B_\varepsilon) + \left(3e^{-\varepsilon}-\frac{2}{3}e^{\varepsilon}\right)Q(B_\varepsilon) - 4
\end{aligned}$$

を得る. □

次に修正 W-divergence と修正 K-L 情報量との不等式関係を考える. B_ε上での修正 W-divergence については補題 A.1.2を用いると, εを用いた評価として次の結果が得られる.

系 2.3.5 近似主領域 B_εでの修正 W-divergence の評価として以下のように与えられる．

$$\{\varepsilon \cdot l(e^{-\varepsilon})\}^2 \cdot P(B_\varepsilon) \le W^*(Q,P;B_\varepsilon) \le \{\varepsilon \cdot u(e^{\varepsilon})\}^2 \cdot P(B_\varepsilon). \tag{2.3.9}$$

ここで $l(t)$, $u(t)$ は式 (A.1.7), (A.1.6) で与えられたものとする．

また, 修正 W-divergence と修正 K-L 情報量の関係については, Matsunawa(1986) で次の定理が与えられている．本章ではこの結果を用いて議論を進めていく．

定理 2.3.2 [Matsunawa]

$$I^*(Q,P;A) \le Q(A) - P(A) + \frac{1}{2}\left(1 + \frac{1}{3}\sup_{X\in A}\left|\frac{dQ(X)}{dP(X)} - 1\right|\right) W^*(Q,P;A) \tag{2.3.10}$$

$$I^*(Q,P;A) \ge Q(A) - P(A) + \frac{1}{2}\left(1 - \frac{1}{3}\sup_{X\in A}\left|\frac{dQ(X)}{dP(X)} - 1\right|\right) W^*(Q,P;A) \tag{2.3.11}$$

2.4　L_1^*型, W^*型の分布間の隔たりの評価

$\omega^*(Q,P;A)$, $\Delta^*(Q,P;A)$ といった分布間の隔たりの尺度は修正 L_1距離 $V^*(Q,P;A)$, 修正 W-divergence $W^*(Q,P;A)$ の積分の基となる分布でとった量である．特に $W^*(Q,P;A)$, $\Delta^*(Q,P;A)$ は, それぞれピアソン型, ネイマン型のχ^2統計量の連続版とみなすことができる (cf.Neyman (1937))．これらの分布間の隔たりは非対称な量である．よってこのような分布間の隔たりの尺度は基にする分布が違う場合を考えた時, これらの意味が異なるためこれらの大小関係を調べることは興味がある．本節ではこれらの量の定量的な関係を与えるとともに, 領域 S_ε, B_εから逃げる確率の評価を行なう．S_εから逃げる確率の評価としては Matsunawa(1982) で次のように与えられている．

定理 2.4.1 [Matsunawa] S_εを含む任意の可測集合に対して,

$$\min\Big\{P(A - S_\varepsilon), Q(A - S_\varepsilon)\Big\} \le \frac{1}{c(\varepsilon)} V^*(Q,P;A)$$

が成り立つ．ここでεは任意の正数とし, $c(\varepsilon) = \dfrac{\varepsilon}{\varepsilon + 1}$とする．

本章では, 任意の可測集合 A での分布間の隔たりを考えているが, 暗に分布の確率測度のほとんどが S_ε, B_εに保持していると想定している．よって分布間の隔たりは S_ε, B_ε上での量

を評価すれば十分なのだが，本節の結果は $A-S_\varepsilon$, $A-B_\varepsilon$へ逃げる確率の評価を考慮した結果を与える．S_εと B_εを比べる時，補題 2.2.1より $S_\varepsilon \subseteq B_\varepsilon$なる包含関係があり，確率の逃げを小さくした評価を与えるためには，B_εを用いた評価を考える方がよいと思われる．しかし L_1^*型，W^*型の分布間の隔たりを考えた場合，構造上 S_εの方が扱い易いという特徴がある．よって，初めに構造上簡単な S_εを用いた評価を考え，2.4.3節で本節の主要結果である B_εを用いた評価について考える．また，2.4.3節では，I^*型の分布間の隔たりとの関係にもふれる．

2.4.1 L_1^*型の分布間の隔たりの評価

定理 2.4.1を用いると，$\omega^*(Q,P;A)$ と $V^*(Q,P;A)$ の定量的な関係として，次のことが成り立つ．

定理 2.4.2 正数 m および Mが存在して，次の条件を満たすものとする:

$$0 < m \le \frac{dQ(X)}{dP(X)} \le M < \infty, (X \in R).$$

この時，$\omega^*(Q,P;A)$ の上下の評価として次の不等式が成り立つ:

$$\begin{aligned}\omega^*(Q,P;A) &\le \varepsilon + K^* \frac{1}{c(\varepsilon)} V^*(Q,P;A)\\ \omega^*(Q,P;A) &\ge k(\varepsilon)\left[\min\Big\{P(A), Q(A)\Big\} - \min\Big\{P(S_\varepsilon), Q(S_\varepsilon)\Big\}\right].\end{aligned}$$

ここで $K^* := \max\Big(u(M), u(1/m)\Big) \max\Big(|\ln M|, |\ln m|\Big)$, $k(\varepsilon) := \min\left\{\varepsilon, \frac{1}{1+\varepsilon}\right\}$とする．

注 2.5 定理 2.4.2では，条件として

$$0 < m \le \frac{dQ(X)}{dP(X)} \le M < \infty, (X \in R)$$

を置いている．これはかなり強い条件に思われるが，本論文では自然な条件といえる．それは領域 A は2つの密度関数 dP, dQ のどちらも0でないような領域を考えている．そしてこの領域での分布の近似を考えているため，密度の比の上下限を与えておくのは自然な条件といえる．

注 2.6 定理 2.4.2の条件の下で A が

$$\min\Big\{P(A), Q(A)\Big\} \geq 1 - \varepsilon^{\#}, (0 \leq \varepsilon^{\#} < 1)$$

を満たす時，さらに$\omega^*(Q, P; A)$ の下界を用いて，確率の逃げる評価として

$$\begin{aligned}\min\Big\{P(S_\varepsilon), Q(S_\varepsilon)\Big\} &\geq \min\Big\{P(A), Q(A)\Big\} - \frac{1}{k(\varepsilon)}\omega^*(Q, P; A) \\ &\geq 1 - \varepsilon^{\#} - \frac{1}{k(\varepsilon)}\omega^*(Q, P; A).\end{aligned}$$

と与えられる．

証明． f, g を P, Q に対応する確率密度関数とする．この時 $\omega(Q, P; S_\varepsilon)$ の上界に関しては

$$\begin{aligned}\omega(Q, P; A) \quad &= \int_{S_\varepsilon} \left|\frac{g}{f} - 1\right| g d\mu + \int_{A-S_\varepsilon} \left|\frac{g}{f} - 1\right| g d\mu \leq \varepsilon \int_{S_\varepsilon} g d\mu + \int_{A-S_\varepsilon} u\left(\frac{g}{f}\right)\left|\ln\frac{g}{f}\right| g d\mu \\ &= \varepsilon Q(S_\varepsilon) + u(M)\max(|\ln M|, |\ln m|)Q(A - S_\varepsilon),\end{aligned}$$

と与えられる．ここで $u(t)$ は式 (A.1.7) で与えられたものとする．同様に，$\omega(P, Q; S_\varepsilon)$ に関しても

$$\omega(P, Q; A) = \varepsilon P(S_\varepsilon) + u\left(\frac{1}{m}\right)\max\Big(|\ln M|, |\ln m|\Big)P(A - S_\varepsilon)$$

と与えられるので，

$$\begin{aligned}&\omega^*(Q, P; A) := \min\Big\{\omega(Q, P; A), \omega(P, Q; A)\Big\} \\ &\leq \varepsilon + \max\Big(u(M), u(1/m)\Big)\max\Big(|\ln M|, |\ln m|\Big) \cdot \min\Big\{P(A - S_\varepsilon), Q(A - S_\varepsilon)\Big\}\end{aligned}$$

となる．ここで定理 2.4.1の結果を用いると，

$$\omega^*(Q, P; A) \leq \varepsilon + K^*\frac{1}{c(\varepsilon)}V^*(Q, P; A),$$

と得られる．

$\omega(Q,P;A)$ の下界に関しては

$$
\begin{aligned}
&\omega(Q,P;A)\\
&= \int_{A-S_\varepsilon} \left|\frac{g}{f}-1\right| g d\mu + \int_{S_\varepsilon} \left|\frac{g}{f}-1\right| g d\mu \geq \varepsilon Q(A-S_\varepsilon) + \inf_{x\in S_\varepsilon}\left(\frac{g}{f}\right)\int_{S_\varepsilon} |g-f| d\mu\\
&\geq \varepsilon\Big\{Q(A)-Q(S_\varepsilon)\Big\} + \inf_{x\in S_\varepsilon}\left(\frac{g}{f}\right)\Big|Q(S_\varepsilon)-P(S_\varepsilon)\Big|
\end{aligned}
$$

また $\omega(P,Q;A)$ の下界に関しても

$$
\omega(P,Q;A) \quad \geq \varepsilon\Big\{P(A)-P(S_\varepsilon)\Big\} + \inf_{x\in S_\varepsilon}\left(\frac{f}{g}\right)\Big|P(S_\varepsilon)-Q(S_\varepsilon)\Big|
$$

と与えられる．よって

$$
\begin{aligned}
\omega^*(Q,P;A) \quad &\geq k(\varepsilon)\Big[\min(P(A),Q(A)) - \Big\{\max(P(S_\varepsilon),Q(S_\varepsilon)) - |P(S_\varepsilon)-Q(S_\varepsilon)|\Big\}\Big]\\
&\geq k(\varepsilon)\Big[\min\{P(A),Q(A)\} - \min\{P(S_\varepsilon),Q(S_\varepsilon)\}\Big]
\end{aligned}
$$

これにより定理 2.4.2が証明された． □

さて，定理 2.4.2の定量評価により $\omega^*(Q,P;A)$ と $V^*(Q,P;A)$ の間には次の関係を与えることができる．

定理 2.4.3 定理 2.4.2の条件の下 $\min\{P(A),Q(A)\} \approx 1$ を満たす領域を考える．この時 $\omega^*(P,Q;A)$ と $V^*(Q,P;A)$ との関係として

$$
V^*(Q,P;A) \approx 0 \Leftrightarrow \omega^*(P,Q;A) \approx 0
$$

が得られる．

証明． $\Rightarrow$ に関しては定理 2.4.2により与えられた．$\Leftarrow$に関しては次の不等式により得られる．

$$
\begin{aligned}
V^* &= \int_A |g-f| d\mu = \int_A \left|\frac{g}{f}-1\right| f d\mu \leq M\int_A \sqrt{\left|\frac{g}{f}-1\right|}\sqrt{\left|\frac{g}{f}-1\right|} g d\mu\\
&\leq M\cdot\max\left(\sqrt{|M-1|},\sqrt{|m-1|}\right)\int_A \sqrt{\left|\frac{g}{f}-1\right|}\sqrt{g}\sqrt{g} d\mu
\end{aligned}
$$

$$\leq M \cdot \max\left(\sqrt{|M-1|}, \sqrt{|m-1|}\right) \sqrt{\int_A \left|\frac{g}{f}-1\right| g d\mu \cdot \int_A g d\mu}$$
$$= M \cdot \max\left(\sqrt{|M-1|}, \sqrt{|m-1|}\right) \sqrt{\omega(Q,P;A) \cdot Q(A)}.$$

また同様に

$$V^* \leq \frac{1}{m} \cdot \max\left(\sqrt{|M-1|}, \sqrt{|m-1|}\right) \sqrt{\omega(P,Q;A) \cdot P(A)}.$$

となることから

$$0 \leq V^* \leq K_1 \sqrt{\omega^*(Q,P;A)}$$

を得る．ここで $K_1 := \max(M, 1/m) \max\left(\sqrt{|M-1|}, \sqrt{|m-1|}\right)$ とする． □

2.4.2 W^*型の分布間の隔たりの評価

$\Delta^*(P,Q;A)$ と $\omega^*(Q,P;A)$ の間には次の関係が与えられる．

定理 2.4.4 正数 m および M が存在して, 次の条件を満たすものとする:

$$0 < m \leq \frac{dQ(X)}{dP(X)} \leq M < \infty, (X \in R).$$

この時, $\Delta^*(P,Q;A)$ の不等式評価として

$$0 \leq \Delta^*(P,Q;A) \leq \varepsilon^2 + K^2 \frac{1}{k(\varepsilon)} \omega^*(Q,P;A)$$

を得る．ここで $K^* := \max\Big(u(M), u(1/m)\Big) \max\Big(|\ln M|, |\ln m|\Big)$, $k(\varepsilon) := \min\left\{\varepsilon, \frac{1}{1+\varepsilon}\right\}$ とする．

証明. 条件の下では $\sup_{x \in R-A} \left|\frac{g(x)}{f(x)} - 1\right| \leq \max\Big(|m-1|, |M-1|\Big) =: K$ と与えられるので,

$$\Delta(Q,P;A) := \int_A \left(\frac{g}{f}-1\right)^2 g d\mu \leq \int_{S_\varepsilon} \left(\frac{g}{f}-1\right)^2 g d\mu + \int_{A-S_\varepsilon} \left(\frac{g}{f}-1\right)^2 g d\mu$$
$$\leq \varepsilon^2 \int_{S_\varepsilon} g d\mu + c^2 \int_{A-S_\varepsilon} g d\mu = \varepsilon^2 Q(S_\varepsilon) + K^2 Q(A - S_\varepsilon).$$

となる．同様に

$$\Delta(P,Q;A)=\int_A\left(\frac{f}{g}-1\right)^2 fd\mu\le\varepsilon^2P(S_\varepsilon)+K^2P(A-S_\varepsilon)$$

と得られることから

$$\begin{aligned}0\le\Delta^*(P,Q;A)&:=\min\left\{\Delta(Q,P;A),\Delta(P,Q;A)\right\}\\&\le\varepsilon^2+K^2\min\left\{P(A-S_\varepsilon),Q(A-S_\varepsilon)\right\}\end{aligned}$$

ここで定理 2.4.2で

$$\min\left\{P(S_\varepsilon),Q(S_\varepsilon)\right\}\ge\min\left\{P(A),Q(A)\right\}-\frac{1}{k(\varepsilon)}\omega^*(Q,P;A)$$

と与えられることから，

$$0\le\Delta^*(P,Q;A)\le\varepsilon^2+K^2\frac{1}{k(\varepsilon)}\omega^*(Q,P;A).$$

を得る． □

定理 2.4.4の定量評価により$\omega^*(Q,P;A)$ と $V^*(Q,P;A)$ の間には次の関係を与えることができる．

定理 2.4.5 定理 2.4.4の条件の下で，$\min\{P(A),Q(A)\}\approx 1$ を満たす領域を考える．この時，$\Delta^*(P,Q;A)$ と$\omega^*(Q,P;A)$ との関係として，

$$\omega^*(Q,P;A)\approx 0\Leftrightarrow\Delta^*(P,Q;A)\approx 0$$

が得られる．

証明． $\Rightarrow$ に関しては定理 2.4.4により与えられた．$\Leftarrow$に関しては次の不等式により得られる．

$$\begin{aligned}\omega^*&\le\int_A\left|\frac{g}{f}-1\right|gd\mu=\int_A\left|\frac{g}{f}-1\right|\sqrt{g}\sqrt{g}d\mu\\&\le\sqrt{\int_A\left(\frac{g}{f}-1\right)^2gd\mu\cdot\int_A gd\mu}=\sqrt{\Delta^*Q(A)}\le\sqrt{\Delta^*}\end{aligned}$$

□

2.4.3 B_εを用いた評価

2.4.1節, 2.4.2節の結果を援用して, B_εを用いた近似の評価を考える. S_εは構造上 L_1^*型, W^*型に適した領域なので, 2.4.1節, 2.4.2節では, S_εに課した条件により証明を簡単に行なえた. しかし補題 A.1.2を援用することにより, B_εを用いた近似の評価も同様な形で表すことができる. また評価の途中で I^*型の分布間の隔たりの関係も現れる. 以下の評価は多少複雑にはなるが, 2.4.1節, 2.4.2節の結果を改善している. 初めに B_εから逃げる確率の評価を与える.

定理 2.4.6 $\mathbf{B}_\varepsilon$を含む任意の可測集合 A に対して,

$$\min\Bigl\{P(A-B_\varepsilon), Q(A-B_\varepsilon)\Bigr\} \le \frac{1}{1-e^{-\varepsilon}} V^*(X,Y;A)$$

が成り立つ. ここでεは正数とする.

注 2.7 この結果は定理 2.4.1で与えられている結果よりも改善されている.

証明.

$$\begin{aligned} V^*(X,Y;A) \quad &\ge \int_{A-B_\varepsilon} dQ \left|\frac{dP}{dQ} - 1\right| d\mu + \int_{B_\varepsilon} |dQ - dP| \\ &\ge (1-e^{-\varepsilon})\{Q(A) - Q(B_\varepsilon)\} + |Q(B_\varepsilon) - P(B_\varepsilon)| \end{aligned}$$

同様な不等式関係から

$$\begin{aligned} V^*(X,Y;A) \;&\ge\; (1-e^{-\varepsilon})\Bigl\{\min\{Q(A),P(A)\} - \max\{Q(B_\varepsilon),P(B_\varepsilon)\}\Bigr\} + |Q(B_\varepsilon) - P(B_\varepsilon)| \\ &\ge\; (1-e^{-\varepsilon})\min\{P(A-B_\varepsilon), Q(A-B_\varepsilon)\} \end{aligned}$$

を得る. □

次に$\omega^*(Q,P;A)$ の評価と$\Delta^*(Q,P;A)$ の評価を B_εを用いて行なう.

定理 2.4.7 正数 m および Mが存在して, 次の条件を満たすものとする:

$$0 < m \le \frac{dQ(X)}{dP(X)} \le M < \infty, (X \in R).$$

この時, $\omega^*(Q,P;A)$ の不等式評価として

$$
\begin{aligned}
\omega^*(Q,P;A) &\le u(e^\varepsilon)I_a^*(Q,P;B_\varepsilon) + K^*\min\Big\{P(A-B_\varepsilon),Q(A-B_\varepsilon)\Big\} \\
&\le u(e^\varepsilon)\cdot\varepsilon + K^*\frac{1}{1-e^{-\varepsilon}}V^*(Q,P;A) \\
\omega^*(Q,P;A) &\ge (1-e^{-\varepsilon})\min\Big\{P(A-B_\varepsilon),Q(A-B_\varepsilon)\Big\}
\end{aligned}
$$

を得る. ここで

$$
\begin{aligned}
&I_a^*(Q,P;B_\varepsilon) = \int_{B_\varepsilon}\left|\ln\frac{dQ}{dP}\right|dQ \\
&K^* = \max\Big\{u(M),u(1/m)\Big\}\max\Big\{|\ln M|,|\ln m|\Big\}
\end{aligned}
$$

とし, $u(t)$, $l(t)$ は式 (A.1.7), (A.1.6) で与えられるものとする.

証明. 定理 2.4.2の証明と補題 A.1.2を用いると, $\omega^*(Q,P;B_\varepsilon)$ の上下界は

$$
\begin{aligned}
\omega^*(Q,P;B_\varepsilon) &= \int_{B_\varepsilon}\left|\frac{dQ}{dP}-1\right|dQ \\
&\le \int_{B_\varepsilon} u\left(\frac{dQ}{dP}\right)\left|\ln\frac{dQ}{dP}\right|dQ \\
&\le u(e^\varepsilon)I_a^*(Q,P;B_\varepsilon) \\
&\le u(e^\varepsilon)\cdot\varepsilon Q(B_\varepsilon)
\end{aligned}
$$

$$
\begin{aligned}
\omega^*(Q,P;B_\varepsilon) &= \int_{B_\varepsilon}\left|\frac{dQ}{dP}-1\right|dQ \\
&\ge \int_{B_\varepsilon} l\left(\frac{dQ}{dP}\right)\left|\ln\frac{dQ}{dP}\right|dQ \\
&\ge l(e^{-\varepsilon})I_a^*(Q,P;B_\varepsilon) \ge 0
\end{aligned}
$$

と与えられる. これと定理 2.4.6を用いると,

$$
\begin{aligned}
\omega^*(Q,P;A) &= \int_{B_\varepsilon}\left|\frac{dQ}{dP}-1\right|dQ + \int_{A-B_\varepsilon}\left|\frac{dQ}{dP}-1\right|dQ \\
&\le u(e^\varepsilon)\cdot\varepsilon + K^*\min\Big\{P(A-B_\varepsilon),Q(A-B_\varepsilon)\Big\} \\
&\le u(e^\varepsilon)\cdot\varepsilon + K^*\frac{1}{1-e^{-\varepsilon}}V^*(Q,P;A)
\end{aligned}
$$

$$\begin{aligned}\omega^*(Q,P;A) &= \int_{B_\varepsilon}\left|\frac{dQ}{dP}-1\right|dQ+\int_{A-B_\varepsilon}\left|\frac{dQ}{dP}-1\right|dQ\\ &\geq l(e^{-\varepsilon})I_a^*(Q,P;B_\varepsilon)+(1-e^{-\varepsilon})\min\left\{P(A-B_\varepsilon),Q(A-B_\varepsilon)\right\}\\ &\geq (1-e^{-\varepsilon})\min\left\{P(A-B_\varepsilon),Q(A-B_\varepsilon)\right\}\end{aligned}$$

を得る. □

補題 2.4.1 正数 m および Mが存在して, 次の条件を満たすものとする.

$$0<m\leq\frac{dQ(X)}{dP(X)}\leq M<\infty,(X\in R).$$

この時, $\Delta^*(Q,P;A)$ の不等式評価として

$$\Delta^*(Q,P;A)\leq\left\{u(e^\varepsilon)\cdot\varepsilon\right\}^2+K^{*2}\frac{1}{1-e^{-\varepsilon}}\omega^*(Q,P;A)$$

を得る. ここで $K^*=\max\left\{u(M),u(1/m)\right\}\max\left\{|\ln M|,|\ln m|\right\}$とし, $u(t)$ は式 (A.1.7) で与えられるものとする.

証明. 次の一連の不等式により, 補題 2.4.1は証明される.

$$\begin{aligned}\Delta^*(Q,P;A) &= \int_A\left|\frac{dQ}{dP}-1\right|^2dQ=\int_{B_\varepsilon}\left|\frac{dQ}{dP}-1\right|^2dQ+\int_{A-B_\varepsilon}\left|\frac{dQ}{dP}-1\right|^2dQ\\ &\leq\int_{B_\varepsilon}u\left(\frac{dQ}{dP}\right)\cdot\left|\ln\frac{dQ}{dP}\right|dQ+K^{*2}\min\left\{P(A-B_\varepsilon),Q(A-B_\varepsilon)\right\}\\ &\leq\left\{u(e^\varepsilon)\cdot\varepsilon\right\}^2P(B_\varepsilon)+K^{*2}\frac{1}{1-e^{-\varepsilon}}\omega^*(Q,P;A)\end{aligned}$$

□

第 3 章

近似主領域の導入と設定

3.1 はじめに

本章では, 可測集合に近似主領域の概念を導入し, 理論, 応用両面から近似主領域の設定を考える. 特に具体的な問題に対し, 柔軟に近似主領域を設定し, 分布の一様近似を行なう.

局所的に相対的な食い違いが大きい 2 つの分布間の隔たりを測る場合, K-L 情報量などは 2 つの分布間同士が近いにも関わらず良い近似を示さない. このため, K-L 情報量を用いた分布の近似を考えた場合, 全比較の対象となる 2 つの確率分布が正値を取る領域が全く同一となる場合のみを考えている. しかし確率分布がこのような設定に納まらない状況は現実の近似問題で起こり得るため, このような確率分布同士についても測ることのできる分布間の隔たりを提案する必要がある. そこで第 2章では, 分布間の隔たりの不等式評価を得る際に, このような欠点を補うため, ある可測集合 A 上での積分とした $I^{*}(Q, P; A)$ という修正を考えた. この時, 修正 K-L 情報量 $I^{*}(Q, P; A) = 0$ ならば, Q と P は可測集合 A 上で一致することが得られる. しかし可測集合が小さな領域だと, その領域での近似しか言えないため, 局所的な近似理論になってしまう. 本論文では, 大域的な近似理論を扱いたいため, 可測集合はある程度大きな領域を考えねばならない. また大域的な近似を考えることにより, Q と P は可測集合 A のみならず, R 上において, 互いに十分近い分布であることが言える.

領域の大小という抽象的な言い方を確率を用いて数学的に捉えると, 次のようになる.

定義 3.1.1 X, Yを可測空間 $(R, \mathbf{B})$ 上で定義される確率変数とする. ここで Rは任意の抽

象空間, $\mathbf{B}$ を Rの部分集合のσ−集合族とする．X, Yに対応する確率分布を P, Q として，これらはある可測集合 $A \in \mathbf{B}$ において絶対連続とする．可測集合 A に対して

$$\min\Big\{P(A), Q(A)\Big\} = 1 - \epsilon \tag{3.1.1}$$

が成り立つものとする．ここでϵは任意の微小な正数とする．この状況の下で，A における分布の近似問題を考える時，A は近似主領域と呼ぶ．

本章の目的は，第2章で可測集合の例として考えた B_ϵが近似主領域に成り得ること，また具体的問題に対して近似主領域を柔軟に設定することである．

第2章では，理論的な興味から可測集合の例として，密度関数の比の近傍としての領域 B_ϵ, S_ϵを扱ってきた．しかし B_ϵや S_ϵにつけた条件を，具体的問題において扱いやすい形に翻訳し構成することは複雑に過ぎることも多い．そこで分布の一様近似の応用を考える場合には，近似主領域に事前に強い制限を置くことなしに，個々の問題毎に柔軟に構成する立場を取ることにする．また指数型分布族の多くの統計モデルは適切な条件の下で極限分布が正規分布であるということは良く知られている．今までは漸近理論から正規近似を考えてきたが，本章では非漸近的なアプローチから正規近似について考察する．本章の非漸近的な結果を用いることにより，極限理論のみならず，パラメータを固定した時に，どれくらいの近似が成されているかを修正 K-L 情報量を用いて評価することができる．

本章では具体的な分布の近似問題として，一変量の例としてはガンマ分布，ベータ分布の正規近似，多変量の例としては Dirichlet 分布の多変量正規への近似を修正 K-L 情報量を用いて扱う．ガンマ分布，ベータ分布は，それぞれ半開区間，閉区間で定義される分布であるが，近似分布として全空間で定義される正規分布が用いられることもある．これらの近似問題を実用的に解析することは興味深い．また順序付けをしたデータが従う分布として Ordered Dirichlet 分布が知られているが，正規分布に近似されることが多い．よってこの2つの分布の差を測ることも意味のあることである．本章では一部の統計モデルの例を考えているが，多くの統計モデルは同一の困難さがあり，ここで得られる結果を用いれば多くの統計モデルの正規近似はほとんど同じやり方で与えることができるであろう．

本章の構成は以下の通りである．まず 3.2節では，理論的な近似主領域の設定として B_ϵの

妥当性について検討する. 3.3節で緩和化した近似主領域における一変量の分布の一様近似を考え, 3.3.1節では一変量の例として, 3.3.1節ではガンマ分布の正規近似とベータ分布の正規近似を, また3.3.2節では多変量の例として Dirichlet 分布の正規近似を考える. またガンマ分布の特別な場合としてχ^2分布の正規近似についても議論する.

3.2 B_εが近似主領域であることの妥当性

本章では理論的な興味から確率空間 $(R^{n\times n}, \mathbf{B}^{n\times n})$ で定義される絶対連続な2つの分布 P, Q に対して B_εが近似主領域として妥当であることを示す. 定義で表したように, $P(B_\varepsilon)$, $Q(B_\varepsilon)$ がほとんど1に近い時に B_εは近似主領域に成り得る. この結果を得るために, 2.3節で与えた定理2.3.1を用いて, 2つの確率の下界を与える.

定理 3.2.1 B_εに対して

$$\min\Big\{P(B_\varepsilon), Q(B_\varepsilon)\Big\} \geq e^{-\varepsilon}\max\Big\{P(B_\varepsilon), Q(B_\varepsilon)\Big\} \tag{3.2.1}$$

が成り立つ.

注 3.1 B_εに対して $\max\{P(B_\varepsilon), Q(B_\varepsilon)\} \geq e^{-\varepsilon^{\#}}$が満たされる時,

$$\min\{P(B_\varepsilon), Q(B_\varepsilon)\} \geq e^{-\varepsilon^{*}} \tag{3.2.2}$$

が成り立つ. ここで$\varepsilon^* = \varepsilon + \varepsilon^{\#}$とする. この結果により$\varepsilon^*$が0に収束する（つまり$\varepsilon$, $\varepsilon^{\#}$がともに0に収束する）と式(3.2.2)の下界は1に単調に収束していく. これにより微小なεに対して, $\max\{P(B_\varepsilon), Q(B_\varepsilon)\} \geq e^{-\varepsilon^{\#}}$が満たされれば, B_ε上での分布の近似を考える時, B_εは近似主領域とみなせることが数学的に表される.

証明. 次の一連の不等式

$$P(B_\varepsilon)\cdot Q(B_\varepsilon) \geq \left(\int_{B_\varepsilon}(dPdQ)^{\frac{1}{2}}\right)^2 \geq \Big\{\rho^*(P,Q;B_\varepsilon)\Big\}^2 \geq e^{-\varepsilon}\left[\max\Big\{P(B_\varepsilon), Q(B_\varepsilon)\Big\}\right]^2$$
$$P(B_\varepsilon)\cdot Q(B_\varepsilon) = \min\{P(B_\varepsilon), Q(B_\varepsilon)\}\cdot\max\{P(B_\varepsilon), Q(B_\varepsilon)\}$$

が与えられ，これにより $\min\{P(B_\varepsilon), Q(B_\varepsilon)\}$ の下界は

$$\min\{P(B_\varepsilon), Q(B_\varepsilon)\} \geq e^{-\varepsilon} \max\{P(B_\varepsilon), Q(B_\varepsilon)\}$$

を得る．これにより定理 3.2.1が証明された． □

定理 3.2.1の結果より，εの取り方と $\max\{P(B_\varepsilon), Q(B_\varepsilon)\}$ の値によって B_εにほとんどの確率測度を保持しているということが理論的に示された．第2章では，理論的な裏付けなしに暗黙の了解として B_εを用いたが，本定理により B_εが近似主領域に成り得ることが明らかになった．また，第2章の不等式評価を与える命題の証明で，任意の可測集合 A に対し，$A - B_\varepsilon$へ逃げる確率も考慮にした結果を与えたが，実際には，B_εから逃げる確率はあまり大きくないため，B_εでの議論に集約することもできる．

3.3 緩和化した近似主領域における分布の一様近似

3.3.1 一変量の場合

本節では，修正 K-L 情報量を用いた，具体的な分布の近似問題について，一変量の例をいくつかあげる．最初にガンマ分布の正規近似を例にとって，修正 K-L 情報量に基づいた分布の近似，近似主領域の設定の仕方について考えていく．問題の背景についてはそれぞれの例の中で述べることにして，最初に緩和化した近似主領域における分布間の隔たりを測る際の計算上の困難さや問題点などについてまとめる．

近似主領域は定義 3.1.1で定義したが，本節ではパラメトリック分布の場合のパラメータの変動をよりわかり易くしたいということから次のような定義の下で分布の近似を考える．

定義 3.3.1 X_n, Y_nを可測空間列 $(R_n, \mathbf{B})$ 上で定義される確率変数とする．ここで各 n に対して，R_nは任意の抽象空間，$\mathbf{B}$ を R_nの部分集合のσ−集合族とする．X_n, Y_nに対応する確率分布列を P_n, Q_nとして，これらはある可測集合列 $A_n \in \mathbf{B}_n$において絶対連続とする．$n \geq n_0$に対して可測集合列 A_nに対して

$$\min\Big\{P_n(A_n), Q_n(A_n)\Big\} = 1 - \eta(\varepsilon_n) \tag{3.3.1}$$

となる正の整数 $n_0 = n_0(\epsilon; L, M)$ が存在するとする．ここで $0 < \epsilon_n < 1$, $0 < \eta_n < 1$ とし，ϵ_n, η_n は $n \to \infty$ とした時に $\epsilon_n \to 0$, $\eta_n \to 0$ となるとする．この状況の下で，A_n における分布の近似問題を考える時，A_n は近似主領域と呼ぶ．

定義 3.1.1と違うのは，パラメータの変動により近似主領域も変化していくことにある．これはパラメータの変動により平均や分散も変化することを考慮に入れれば，近似主領域が変化することは自然なことである．

本節ではパラメータの変動により分布の近さがどのように変化するかを考察するため，X_n, Y_n をそれぞれパラメータ n のガンマ分布，パラメータ (n, n) の正規分布とする．またそれぞれの密度関数を f_g, f_N とし，

$$f_g(x) = \frac{1}{\Gamma(n)} x^{n-1} e^{-x} \tag{3.3.2}$$

$$f_N(x) = \frac{1}{\sqrt{2\pi n}} e^{-\frac{(x-n)^2}{2n}} \tag{3.3.3}$$

と表す．ここで $n > 1$ とする．この条件により本節では凸関数となるガンマ分布のみを近似させる対象としている．図 3.1は，パラメータを $n = 5$ とした時の密度関数である．図 3.1のようにガンマ分布，正規分布はそれぞれ非負数，全空間で定義される確率分布なので，通常の K-L 情報量ではどんなに近い分布でも発散してしまう．これにより I^*型の分布間の隔たりを考える場合には，近似主領域を用いた修正 K-L 情報量を考えなければならない．近似主領域を A_n とすると，修正 K-L 情報量は

$$I^*(f_N, f_g; A_n) = \int_{A_n} \log \frac{f_N(x)}{f_g(x)} f_N(x) dx. \tag{3.3.4}$$

$$= \int_{A_n} \left\{ \log \frac{\Gamma(n)}{\sqrt{2\pi n}} + x - \frac{(x-n)^2}{2n} - (n-1) \log x \right\} \tag{3.3.5}$$

$$= \log \frac{\Gamma(\alpha)}{\sqrt{2\pi n}} Pr(x \in A) + E^*[x] - E^*\left[\frac{(x-n)^2}{2n}\right] - (n-1) E^*[\ln x] \tag{3.3.6}$$

と表される．ここで $E^*[h(x)]$ は

$$E^*[h(x)] = \int_{A_n} h(x) f_N(x) dx \tag{3.3.7}$$

と定義する．これにより修正 K-L 情報量 $I^*(f_N, f_g; A_n)$ を得るためには，近似主領域 A_n での積分として $E^*[x]$, $E^*[(x-n)^2]$, $E^*[\ln x]$ といった量が必要になる．ここで問題となるの

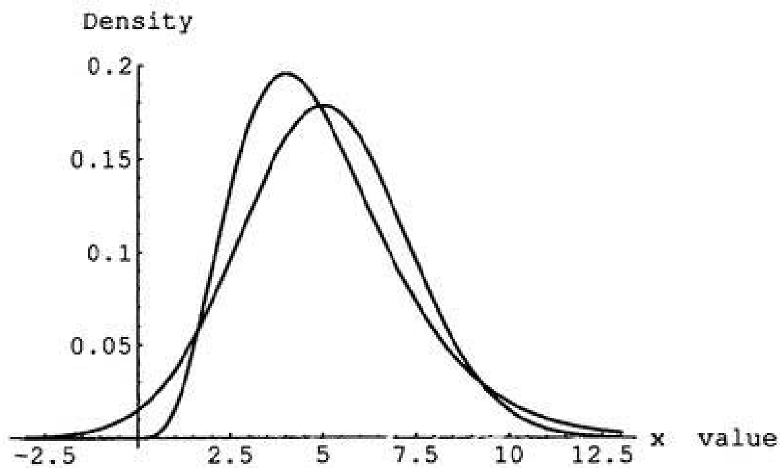

Figure 3.1: ガンマ分布と正規分布の密度関数

は, $E^*[\ln x]$ は今の所, 正確には得ることができないということである. しかし修正 K-L 情報量 $I^*(f_N, f_g; A_n)$ が

$$\xi(n) \leq I^*(f_N, f_g; A_n) \leq \eta(n) \tag{3.3.8}$$

を満たし, $n \to \infty$ とした時に$\xi(n) \to 0$, $\eta(n) \to 0$ が成立する時, ガンマ分布列 f_gは正規分布列 f_Nに収束すると言える. これにより修正 K-L 情報量の不等式評価を考えていく. なお, ここで必要となる諸量の計算は付録 Bにある. また領域 A_nを近似主領域と見なすためには A_nに入る確率が 1 に近い, つまり任意のε_nに対して

$$P(A_n) > 1 - \varepsilon_n \ (0 \leq \varepsilon_n \leq 1, \varepsilon_n \to 0 \text{ as } n \to \infty)$$

となる条件が必要である. このために修正 K-L 情報量と確率の評価を考慮に入れて, 近似主領域の設定を考えねばならない. 本節ではパラメータを固定した時の修正 K-L 情報量と確率の評価を与え, パラメータを大きくした時に正規近似が言えることを示していく.

また本節での結果と第 2 章の分布間の隔たりの大小関係の結果を用いて, 方向性のない一様近似に関する諸量 $D(Q_n, P_n; \mathbf{B}_n)$ での近似なども同時に言え, 分布の一様近似が証明される.

注 3.2 本節では, 平均, 分散が基礎分布と同じ近似分布を考えている. 平均, 分散がそろっているのは自然な近似分布の取り方だと言えるが, いろいろな応用や別の考え方をする時はより一般的な形で分布間の隔たりを与えた方がよい. しかし本節では近似の度合をよりわかりやすくしていきたいため, 一般形は付録でまとめていく. 本節では付録で与えた一般的な結果を用いて証明を行なっている.

例 1：ガンマ分布の正規近似

式 (3.3.2) で定義したガンマ分布の式 (3.3.3) で定義した正規分布への近似を考える.

定理 3.3.1 近似主領域として $A_n = \left\{x; n - n^{\xi} \leq x \leq \infty\right\}$を考える. ここで$\frac{1}{2} < \xi < 1$ とする. この時, $n \to \infty$ とした時に, ガンマ分布は漸近正規性を有する. また, 直ちに一様誤差 Dでの一様近似が示される.

証明. 定理3.3.1を示すためには,

$$I^*(f_N, f_g; A_n) \to 0 \tag{3.3.9}$$

$$Pr(X \in A_n) \to 1 \tag{3.3.10}$$

となることを示せば良い.

$$\gamma(a,b) := \int_0^b x^{a-1}e^{-x}$$

を定義する. 系B.2.1の条件の下で$\alpha = n$, $\beta = 1$とする. また定理3.3.1の条件の下では, a^*, b^*はそれぞれ$\frac{1}{2}n^{2\xi-1}$, ∞となる. また任意の正値iに対して

$$\gamma(i,\infty) = 1$$

となる. この時, 修正K-L情報量$I^*(f_N, f_g; A_n)$の上下界はそれぞれ

$$\begin{aligned}
I^*(f_N,f_g;A) \le& \frac{1}{2}\left[\frac{1}{12n} - \underline{R}(n)\right]\left(\gamma(\frac{1}{2},\frac{1}{2}n^{2\xi-1})+1\right) \\
&-\frac{1}{\sqrt{2\pi n}}\left(\gamma(1,\frac{1}{2}n^{2\xi-1})-1\right)-\frac{1}{4n}\left(\gamma(\frac{3}{2},\frac{1}{2}n^{2\xi-1})+1\right) \\
&+\frac{\sqrt{2}(n-1)}{3n^{3/2}\sqrt{\pi}}\left(\gamma(2,\frac{1}{2}n^{2\xi-1})-1\right)+\frac{n-1}{2n^2}\left(\frac{1}{1-1/n^{1-\xi}}\gamma(\frac{5}{2},\frac{1}{2}n^{2\xi-1})+1\right) \qquad (3.3.11) \\
I^*(f_N,f_g;A) \ge& \frac{1}{2}\left[\frac{1}{12n} - \overline{R}(n)\right]\left(\gamma(\frac{1}{2},\frac{1}{2}n^{2\xi-1})+1\right) \\
&-\frac{1}{\sqrt{2\pi n}}\left(\gamma(1,\frac{1}{2}n^{2\xi-1})-1\right)-\frac{1}{4n}\left(\gamma(\frac{3}{2},\frac{1}{2}n^{2\xi-1})+1\right) \\
&+\frac{\sqrt{2}(n-1)}{3n^{3/2}\sqrt{\pi}}\left(\gamma(2,\frac{1}{2}n^{2\xi-1})-1\right) \\
&-\frac{n-1}{4n^2(1-1/n^{1-\xi})^2}\gamma(\frac{5}{2},\frac{1}{2}n^{2\xi-1}) \qquad (3.3.12)
\end{aligned}$$

と表される. よって$n \to \infty$とした時, ここで得た修正K-L情報量$I^*(f_N, f_g; A)$の上下界はそれぞれ0になることから$I^*(f_N, f_g; A) \to 0$とわかる. また正規変数, ガンマ変数がそれぞれA_nに入る確率は

$$\begin{aligned}
Pr_n(X \in A) &= \frac{1}{2\sqrt{\pi}}\left\{\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},\frac{1}{2}n^{\xi-1})+1\right)\right\} \\
Pr_g(X \in A) &= 1-\gamma(n, n-n^{\xi})
\end{aligned}$$

と表される．これより $n \to \infty$ とした時，$Pr_N(A) \to 1$, $Pr_g(A) \to 1$, $I^*(f_N, f_g; A) \to 0$ となり，これよりガンマ分布のパラメータを大きくした時の漸近分布は正規分布になる．これにより定理 3.3.1が証明された． □

次にガンマ分布の特別な場合として，χ^2分布の正規近似に関する結果を与える．χ^2分布の正規近似はよく知られていて，いろいろな応用分野で用いられている．次の結果は，この近似を修正 K-L 情報量を用いた評価を与えているが，今までの近似理論とは別の角度からこの近似を捉えているので，大変興味のある結果である．

定理 3.3.2 系 B.2.2の条件の下で，近似主領域として $A_n = \{x; n - n^\xi \le x \le \infty\}$ を考える．ここで$\frac{1}{2} < \xi < 1$ とする．この時，$n \to \infty$ とした時に，χ^2 分布は漸近正規性を有する．また，直ちに一様誤差 Dでの一様近似が示される．

証明． 定理 3.3.2を示すためには

$$I^*(f_N, f_{\chi^2}; A_n) \to 0$$
$$Pr(X \in A_n) \to 1$$

を示せば良い．修正 K-L 情報量 $I^*(f_N, f_{\chi^2}; A_n)$ の上下界はそれぞれ

$$\begin{aligned}
&I^*(f_N, f_{\chi^2}; A_n) \le \frac{1}{2}\left[\frac{1}{6n} - \underline{R}(\frac{n}{2})\right]\left(\gamma(\frac{1}{2}, \frac{1}{2}n^{2\xi-1}) + 1\right) \\
&-\frac{1}{\sqrt{\pi n}}\left(\gamma(1, \frac{1}{2}n^{2\xi-1}) - 1)\right) - \frac{1}{2n}\left(\gamma(\frac{3}{2}, \frac{1}{2}n^{2\xi-1}) + 1\right) \\
&+\frac{2(n-2)}{3n^{3/2}\sqrt{\pi}}\left(\gamma(2, \frac{1}{2}n^{2\xi-1}) - 1\right) + \frac{n-2}{2n^2}\left(\frac{1}{1-n^{\xi-1}}\gamma(\frac{5}{2}, \frac{1}{2}n^{2\xi-1}) + 1\right) \\
&I^*(f_N, f_{\chi^2}; A_n) \ge \frac{1}{2}\left[\frac{1}{6n} - \overline{R}(\frac{n}{2})\right]\left(\gamma(\frac{1}{2}, \frac{1}{2}n^{2\xi-1}) + 1\right) \\
&-\frac{1}{\sqrt{\pi n}}\left(\gamma(1, \frac{1}{2}n^{2\xi-1}) - 1\right) - \frac{1}{2n}\left(\gamma(\frac{3}{2}, \frac{1}{2}n^{2\xi-1}) + 1\right) \\
&+\frac{2(n-2)}{3n^{3/2}\sqrt{\pi}}\left(\gamma(2, \frac{1}{2}n^{2\xi-1}) - 1\right) - \frac{n-2}{2n^2(1-n^{\xi-1})^2}\gamma(\frac{5}{2}, \frac{1}{2}n^{2\xi-1})
\end{aligned}$$

と表される．また正規変数，ガンマ変数がそれぞれ A_nに入る確率は

$$Pr_N(X \in A_n) = \frac{1}{2\sqrt{\pi}}\left\{\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2}, a^*) + \gamma(\frac{1}{2}, b^*)\right)\right\}$$
$$Pr_{\chi^2}(X \in A_n) = \gamma(n, n + n^\xi) - \gamma(n, n - n^\xi)$$

と表される．これより $n \to \infty$ とした時, $Pr_N(A_n) \to 1$, $Pr_{\chi^2}(A_n) \to 1$, $I^*(f_N, f_{\chi^2}; A_n) \to 0$ となり，これよりχ^2分布はパラメータ n を大きくした時の漸近分布は正規分布になる．これにより定理3.3.2が証明された． □

例2：ベータ分布の正規近似

$f_\beta(x)$ を

$$f_\beta(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}$$

で定義されるパラメータ (α, β) のベータ分布の密度関数とする．ここで$\alpha > 0$, $\beta > 0$ とし, $B(\alpha,\beta)$ は

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

によって定義されるベータ関数とする．ベータ分布は, $[0,1]$ 上で定義される分布で, 有限区間上で定義される連続分布に対してベータ近似が行なわれることがある．しかし, 分布の近似を考える場合, データの特性を考えずに正規近似が行なわれることが多い．よって，この2つの分布の隔たりを測ることは興味がある．本節では修正 K-L 情報量に基づいたベータ分布の正規近似を考える．なお, 本節では$\alpha + \beta = n$, $\alpha = \alpha(n) = k_1 n$, $\beta = \beta(n) = k_2 n$ (k_1, k_2:定数) である状況を考え, n が変動する時の分布の隔たりを考察する．この時次の定理が成り立つ．

定理 3.3.3 系 B.3.1の結果と, 確率を考慮に入れ, 領域 A_nとして

$$A_n = \{x; k_1 - \delta_n \le x \le k_1 + \delta'_n\}$$

なる領域を考える．ここで

$$\delta_n = \sqrt{\sigma^2(k_1 n)^\tau} = \sqrt{\frac{n^\tau}{n+1}} \cdot k_1^{(1+\tau)/2} k_2^{1/2}$$
$$\delta'_n = \sqrt{\sigma^2(k_1 n)^\kappa} = \sqrt{\frac{n^\kappa}{n+1}} \cdot k_1^{(1/2} k_2^{(1+\kappa)/2}$$

とし, τ, κは, $0 < \tau < 1$, $0 < \kappa < 1$ を満たす実数とする．この時, $n \to \infty$ とした時, ベータ分布は漸近正規性を有する．また, 直ちに一様誤差 Dでの一様近似が示される．

証明. 定理3.3.3を証明するためには

$$I^*(f_N, f_\beta; A_n) \to 0$$

$$Pr(X \in A_n) \to 1$$

を示せば良い. 系B.3.1の結果に $a = k_1 - \delta_n$, $b = k_1 + \delta_n'$ とすると, $a^* = (k_1 n)^\tau$, $b^* = (k_2 n)^\kappa$ と表されるので, 修正K-L情報量の上下界は

$$
\begin{aligned}
&\overline{I}^*(f_N, f_\beta; A_n) = \frac{1}{2}\left[\frac{1}{2}\log\frac{n+1}{n} + \frac{1}{12n}\left(\frac{1}{k_1} + \frac{1}{k_2} - 1\right) + \overline{R}(n) - \underline{R}(k_1 n) - \underline{R}(k_2 n)\right] \\
&\left(\gamma(\frac{1}{2}, (k_1 n)^\tau) + \gamma(\frac{1}{2}, (k_2 n)^\kappa)\right) + \frac{1}{\sqrt{n+1}} \cdot \frac{k_1 - k_2}{\sqrt{2\pi k_1 k_2}}\left(\gamma(1, (k_1 n)^\tau) - \gamma(1, (k_2 n)^\kappa)\right) \\
&-\frac{1}{n+1} \cdot \frac{1 - k_1 k_2}{4 k_1 k_2}\left(\gamma(\frac{3}{2}, (k_1 n)^\tau) + \gamma(\frac{3}{2}, (k_2 n)^\kappa)\right) \\
&+\left[\frac{n}{(n+1)^{3/2}} \cdot \frac{\sqrt{2}(k_2 - k_1)}{\sqrt{\pi k_1 k_2}} + \frac{1}{(n+1)^{3/2}} \cdot \frac{\sqrt{2}(k_1^3 - k_2^3)}{3\sqrt{\pi}(k_1 k_2)^{3/2}}\right]\left(\gamma(2, (k_1 n)^\tau) - \gamma(2, (k_2 n)^\kappa)\right) \\
&+\frac{n}{2(n+1)^2}\left[\left(\frac{k_1^2}{k_2} + \frac{k_2^2}{k_1\left(1 - \frac{n^{\tau/2}}{(n+1)^{1/2}} k_1^{(\tau-1)/2} k_2^{1/2}\right)}\right)\gamma(\frac{5}{2}, (k_1 n)^\tau)\right. \\
&\left.+\left(\frac{k_1^2}{k_2\left(1 - \frac{n^{\kappa/2}}{(n+1)^{1/2}} k_1^{1/2} k_2^{(\kappa-1)/2}\right)} + \frac{k_2^2}{k_1}\right)\gamma(\frac{5}{2}, (k_2 n)^\kappa)\right] \\
&-\frac{1}{2(n+1)^2}\left[\left(\frac{k_1^2}{k_2^2} + \frac{k_2^2}{k_1^2\left(1 - \frac{n^{\tau/2}}{(n+1)^{1/2}} k_1^{(\tau-1)/2} k_2^{1/2}\right)}\right)\gamma(\frac{5}{2}, (k_1 n)^\tau)\right. \\
&\left.+\left(\frac{k_1^2}{k_2^2\left(1 - \frac{n^{\kappa/2}}{(n+1)^{1/2}} k_1^{1/2} k_2^{(\kappa-1)/2}\right)} + \frac{k_2^2}{k_1^2}\right)\gamma(\frac{5}{2}, (k_2 n)^\kappa))\right] \\
&\underline{I}^*(f_N, f_\beta; A) = \frac{1}{2}\left[\frac{1}{2}\log\frac{n+1}{n} + \frac{1}{12n}\left(\frac{1}{k_1} + \frac{1}{k_2} - 1\right) + \underline{R}(n) - \overline{R}(k_1 n) - \overline{R}(k_2 n)\right] \\
&\left(\gamma(\frac{1}{2}, (k_1 n)^\tau) + \gamma(\frac{1}{2}, (k_2 n)^\kappa))\right) + \frac{1}{\sqrt{n+1}} \cdot \frac{k_1 - k_2}{\sqrt{2\pi k_1 k_2}}\left(\gamma(1, (k_1 n)^\tau) - \gamma(1, (k_2 n)^\kappa)\right) \\
&-\frac{1}{n+1} \cdot \frac{1 - k_1 k_2}{4 k_1 k_2}\left(\gamma(\frac{3}{2}, (k_1 n)^\tau) + \gamma(\frac{3}{2}, (k_2 n)^\kappa)\right) \\
&+\left[\frac{n}{(n+1)^{3/2}} \cdot \frac{\sqrt{2}(k_2 - k_1)}{\sqrt{\pi k_1 k_2}} + \frac{1}{(n+1)^{3/2}} \cdot \frac{\sqrt{2}(k_1^3 - k_2^3)}{3\sqrt{\pi}(k_1 k_2)^{3/2}}\right](\gamma(2, (k_1 n)^\tau) - \gamma(2, (k_2 n)^\kappa)) \\
&+\frac{n}{2(n+1)^2}\left[\left(-\frac{k_1^2}{k_2(1 + \frac{n^{\tau/2}}{(n+1)^{1/2}} k_1^{(\tau+1)/2} k_2^{-1/2})} + \frac{k_2^2}{k_1\left(1 - \frac{n^{\tau/2}}{(n+1)^{1/2}} k_1^{(\tau-1)/2} k_2^{1/2}\right)^2}\right)\gamma(\frac{5}{2}, (k_1 n)^\tau)\right.
\end{aligned}
$$

$$+\left(\frac{k_1^2}{k_2\left(1-\frac{n^{\kappa/2}}{(n+1)^{1/2}}k_1^{1/2}k_2^{(\kappa-1)/2}\right)^2}-\frac{k_2^2}{k_1(1+\frac{n^{\kappa/2}}{(n+1)^{1/2}}k_1^{-1/2}k_2^{(\kappa+1)/2})}\right)\gamma(\frac{5}{2},(k_2n)^{\kappa})\Bigg]$$

$$-\frac{1}{2(n+1)^2}\Bigg[\left(-\frac{k_1^2}{k_2^2(1+\frac{n^{\tau/2}}{(n+1)^{1/2}}k_1^{(\tau+1)/2}k_2^{-1/2})}+\frac{k_2^2}{k_1^2\left(1-\frac{n^{\tau/2}}{(n+1)^{1/2}}k_1^{(\tau-1)/2}k_2^{1/2}\right)^2}\right)\gamma(\frac{5}{2},(k_1n)^{\tau})$$

$$+\left(\frac{k_1^2}{k_2^2\left(1-\frac{n^{\kappa/2}}{(n+1)^{1/2}}k_1^{1/2}k_2^{(\kappa-1)/2}\right)^2}-\frac{k_2^2}{k_1^2(1+\frac{n^{\kappa/2}}{(n+1)^{1/2}}k_1^{-1/2}k_2^{(\kappa+1)/2})}\right)\gamma(\frac{5}{2},(k_2n)^{\kappa})\Bigg]$$

となる．よって $n\to\infty$ とした時に，$\overline{I}^*(f_N,f_\beta;A_n)\to 0$, $\underline{I}^*(f_N,f_\beta;A)\to 0$ と与えられたことにより $I^*(f_N,f_\beta;A_n)\to 0$ が導かれる．また補題 B.1.1により，確率 $Pr(X\in A)$ は

$$Pr(X\in A)=\frac{1}{2}\left(\gamma(\frac{1}{2},(k_1n)^{\tau})+\gamma(\frac{1}{2},(k_2n)^{\kappa})\right)$$

と与えられるので，$Pr(X\in A_n)\to 1$ が得られる．これにより定理が証明された． □

3.3.2 多変量の場合

多変量の例として，一様分布から無作為抽出された順序統計量に基づく ordered Dirichlet 分布の近似問題について考察する．$U_{n1}<U_{n2}<\cdots<U_{nn}$を一様分布 $U(0,1)$ から無作為抽出された大きさ n に基づく順序統計量とする．この順序統計量から $k=k(n)(<n)$ 個選び，それらを新たに $U_{nn_1}<U_{nn_2}<\cdots<U_{nn_k}(n_1<n_2<\cdots<n_k)$ と置く．ここで $n_0=0$, $n_{k+1}=n+1$, $U_0=0$, $U_{k+1}=n+1$ とする．この時，ランダムベクトル $U_{n(k)}=(U_{nn_1},\cdots,U_{nn_k})$ の同時確率密度関数は ordered Dirichlet 密度

$$h_n\left(z_{n(k)}\right)=\left\{n!/\prod_{i=1}^{k+1}d_i!\right\}\cdot\prod_{i=1}^{k+1}(z_i-z_{i-1})^{d_i},\ (0=z_0<z_1<\cdots<z_k<z_{k+1}=1)\quad(3.3.13)$$

によって与えられる (cf. Wilks(1962))．ここで $z_{n(k)}=(z_1,\cdots,z_k)\in R_{(k)}$, $d_i=n_i-n_{i-1}-1$ とする．$U_{n(k)}$ に対応して同時確率密度

$$g_n\left(z_{n(k)}\right)=(2\pi)^{-k/2}\cdot\left|L_{n(k)}\right|^{-k/2}\cdot\exp\left[-\frac{1}{2}\left(z_{(k)}-l_{n(k)}\right)^tL_{n(k)}^{-1}\left(z_{(k)}-l_{n(k)}\right)\right]\quad(3.3.14)$$

を持つ正規ランダムベクトル $Z_{n(k)}=(Z_{n1},\cdots,Z_{nk})$ を考える．ここで $l_{n(k)}=(l_{n1},\cdots,l_{nk})^t$, $l_{ni}=n_i/(n+1)$, $i=1,\cdots,k$, $l_{n0}=0$, $l_{nk+1}=1$ とし，

$$L_{n(k)}:=\frac{1}{n+2}\left(l_{ni}(1-l_{nj})\right)_{k\times k}\ (1\le i\le j\le k)\ ,\ \left|L_{n(k)}\right|=(n+2)^{-k}\prod_{i=1}^{k+1}\left(l_{ni}-l_{nj}\right).$$

と置く.

この近似問題に対して Mosteller (1946), Walker (1968) らは k を固定した時の弱収束の意味での同時正規性の極限理論を与えた. また Weiss (1969), Ikeda-Matsunawa (1972), Reiss (1975) らは kが n につれて変動する時の 2 つのランダムベクトルの確率分布の間の一様誤差の意味での漸近正規性を与えている. 本論文では後者の立場を取り, 近似問題を考えていく. しかし 彼らの内容とは違う点は $h_n(z_{n(k)})$ に関する $g_n(z_{n(k)})$ の近似の度合を近似主領域における修正 K-L 情報量と近似主領域から逃げる確率を測ることにより, 一様誤差による強い一様近似を展開している所である. 我々の修正 K-L 情報量を用いることにより k個のサンプルの同時正規性について, 基の確率空間が変動する時にどうなるかを近似の誤差評価を用いて調べる. この結果を得るためにある近似主領域での修正 K-L 情報量を用いる. なお, 本節の結果は系 A.2.1と B.4節で与える ordered Dirichlet 分布の密度関数の近似展開を用いて得られる.

例 3 : Ordered Dirichlet 分 布 の 正 規 近 似

本節では, 一様分布 $U(0,1)$ から無作為抽出されたランダムサンプルに基づく n 個の順序統計量から $k(n)$ 個選んだ時の同時正規性についての評価を考える. 今

$$Q^0_{n(k)} = \left\{ z_{(k)} = (z_1, \cdots, z_k); 0 \equiv z_0 < z_1 < \cdots < z_k < z_{k+1} \equiv 1 \right\} \tag{3.3.15}$$

と定義する. またある与えられた正数 L, Mに対して

$$Q^{L,M}_{n(k)} = \left\{ z_{(k)} = (z_1, \cdots, z_k);\ \begin{array}{l} 1 < \frac{z_i - z_{i-1}}{l_{ni} - l_{ni-1}} < 1 + \frac{1}{L} \text{ or} \\ \frac{M}{1+M} < \frac{z_i - z_{i-1}}{l_{ni} - l_{ni-1}} < 1, i = 0, 1, \cdots, k+1 \end{array} \right\}, \tag{3.3.16}$$

を定義する. 本節では, 近似主領域として $A^{L,M}_{n(k)} := Q^0_{n(k)} \cap Q^{L,M}_{n(k)}$をとる時の修正 K-L 情報量による評価を考える. つまり

$$I^*(g_n, h_n; A^{L,M}_{n(k)}) \to 0 \text{ as } n \to \infty \tag{3.3.17}$$

を評価することになる. ここで $g_n(z_{(k)})$ は (3.3.14) 式により与えられた正規変数 $Z_{n(k)}$の確率密度関数, $h_n(z_{(k)})$ は (3.3.13) 式により与えられた一様変数から $k(n)$ 個選んだ時の同時確率密度関数とする. 以後, $U_{n(k)}$, $Z_{n(k)}$の確率分布をそれぞれ $P^{U_{n(k)}}$, $P^{Z_{n(k)}}$ と表す.

Ikeda-Matsunawa (1972) では関連した K-L 情報量 $I(h_n, g_n; Q^0_{n(k)})$ を評価している．この情報量は本節で考えている $I^*(g_n, h_n; A^{L,M}_{n(k)})$ より計算上便利であるが，近似の方向性や関連した物理的な意味合いから直接近似問題を扱っているのではない (cf. Matsunawa(1995))．逆に，$I^*(g_n, h_n; A^{L,M}_{n(k)})$ を用いれば，近似主領域として扱われている $A^{L,M}_{n(k)}$ に十分配慮を置かねばならない．

以上の設定の下で，ordered Dirichlet 変数 $U_{n(k)}$ の漸近正規性について次の定理を与える．

定理 3.3.4 次の条件

$$\frac{k}{\min\limits_{1\le i\le k+1}(n_i - n_{i-1})} \to 0 \text{ as } n \to \infty \tag{3.3.18}$$

を満たすと仮定する．この時に近似主領域として

$$A^{L,M}_{n(k)} = \left\{ z = (z_1, \cdots, z_k) \,\middle|\, \begin{array}{l} 0 \equiv z_0 < z_1 < \cdots < z_k < z_{k+1} \equiv 1, \\ \dfrac{z_i - z_{i-1}}{l_{ni} - l_{ni-1}} > 1\text{or} \\ \dfrac{K}{K+1} < \dfrac{z_i - z_{i-1}}{l_{ni} - l_{ni-1}} < 1,\ i = 0, 1, \cdots, k+1 \end{array} \right\} \tag{3.3.19}$$

を考える．ここで L, M は与えられた正数とする．この時に ordered Dirichlet 分布は多変量正規分布に漸近的に近似される．また ordered Dirichlet 分布は正規分布に一様誤差 D の意味で一様に近似される．

証明． 定理 3.3.4を示すためには，

$$I^*(g_n, h_n; A^{L,M\,L,M}_{n(k)\,n(k)}) \to 0 \text{ as } n \to \infty \tag{3.3.20}$$

$$P^{Z_{n(k)}}\left(A^{L,M\,L,M}_{n(k)\,n(k)}\right) \to 1 \text{ as } n \to \infty \tag{3.3.21}$$

を示せば良い．$z_{n(k)} \in A^{L,M}_{n(k)}$ に対して，定理 B.4.1を用いれば，

$$\begin{aligned}\left|I^*(g_n, h_n; A^{L,M}_{n(k)})\right| &< \int_{A^{L,M}_{n(k)}} g_n(z_{(k)}) \left|\ln\left(g_n(z_{(k)})/h_n(z_{(k)})\right)\right| dZ_{(k)} \\ &=: E^*_g\left[\left|\ln\left(g_n(z_{(k)})/h_n(z_{(k)})\right)\right|\right] \\ &< \frac{k}{2}\left\{\frac{1}{n+1} - \frac{1}{6k(n+1)} - \frac{1}{2(n+1)(n+2)} - \frac{1}{6(n+1)(n+2)^2}\right\}\end{aligned}$$

$$+\frac{1}{12(n+1)}\sum_{i=1}^{k+1}\frac{1}{l_{ni}-l_{ni-1}}+R(n+1)$$
$$+\frac{1}{2(n+2)}E_g^*\left[(z_{(k)}-l_{(k)})^t L_{n(k)}^{-1}(z_{(k)}-l_{(k)})\right]$$
$$+\sum_{i=1}^{k+1}\frac{1}{l_{ni}-l_{ni-1}}E_g^*\left[|(z_i-l_{ni})-(z_{i-1}-l_{ni-1})|\right]$$
$$+\frac{1}{2}\sum_{i=1}^{k+1}\frac{1}{(l_{ni}-l_{ni-1})^2}E_g^*\left[|(z_i-l_{ni})-(z_{i-1}-l_{ni-1})|^2\right]$$
$$+\frac{1}{3}\sum_{i=1}^{k+1}\left(\frac{n+1}{(l_{ni}-l_{ni-1})^2}+\frac{1}{(l_{ni}-l_{ni-1})^3}\right)E_g^*\left[|(z_i-l_{ni})-(z_{i-1}-l_{ni-1})|^3\right]$$
$$+|\theta|\sum_{i=1}^{k+1}\left(\frac{n+1}{(l_{ni}-l_{ni-1})^3}+\frac{1}{(l_{ni}-l_{ni-1})^4}\right)E_g^*\left[|(z_i-l_{ni})-(z_{i-1}-l_{ni-1})|^4\right]$$

となる．ただし θ は

$$\frac{1}{6}\left\{1-\frac{1}{(L+1)^2}\right\}<\theta=\theta(L,M)<\max\left\{\frac{1}{3},\frac{1}{4}\left(1+\frac{1}{M}\right)\right\}$$

を満たす実数である．この不等式に次の関係式

$$\begin{aligned}
&E_g^*[(z_{(k)}-l_{(k)})^t L_{n(k)}^{-1}(z_{(k)}-l_{(k)})]\\
&\quad< E_g[(z_{(k)}-l_{(k)})^t L_{n(k)}^{-1}(z_{(k)}-l_{(k)})]=tr\left(L_{n(k)}^{-1/2}L_{n(k)}L_{n(k)}^{-1/2}\right)=tr\left(I_{(k)}\right)=k\\
&E_g^*\left[|(z_i-l_{ni})-(z_{i-1}-l_{ni-1})|\right]\\
&\quad< E_g\left[|(z_i-l_{ni})-(z_{i-1}-l_{ni-1})|\right]=\frac{\sqrt{2}}{\sqrt{\pi}(n+2)^{1/2}}[(l_{ni}-l_{ni-1})\{1-(l_{ni}-l_{ni-1})\}]^{1/2}\\
&E_g^*\left[|(z_i-l_{ni})-(z_{i-1}-l_{ni-1})|^2\right]\\
&\quad< E_g\left[|(z_i-l_{ni})-(z_{i-1}-l_{ni-1})|^2\right]=\frac{1}{n+2}[(l_{ni}-l_{ni-1})\{1-(l_{ni}-l_{ni-1})\}]\\
&E_g^*\left[|(z_i-l_{ni})-(z_{i-1}-l_{ni-1})|^3\right]\\
&\quad< E_g\left[|(z_i-l_{ni})-(z_{i-1}-l_{ni-1})|^3\right]=\frac{\sqrt{2}}{\sqrt{\pi}(n+2)^{3/2}}[(l_{ni}-l_{ni-1})\{1-(l_{ni}-l_{ni-1})\}]^{3/2}\\
&E_g^*\left[|(z_i-l_{ni})-(z_{i-1}-l_{ni-1})|^4\right]\\
&\quad< E_g\left[|(z_i-l_{ni})-(z_{i-1}-l_{ni-1})|^4\right]=\frac{3}{(n+2)^2}[(l_{ni}-l_{ni-1})\{1-(l_{ni}-l_{ni-1})\}]^2
\end{aligned}$$

を用いると，

$$\begin{aligned}
&\left|I^*(g_n,h_n;A_{n(k)}^{L,M})\right|\\
&\quad<\frac{k}{n+1}\left\{1-\frac{1}{12k}-\frac{3}{4(n+2)}-\frac{1}{12(n+2)^2}\right\}
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{360(n+1)(n+2)(n+3)}+\frac{1}{32(n+1)^2(n+2)(n+3)}\\
&+\frac{1}{12}\sum_{i=1}^{k+1}\frac{1}{n_i-n_{i-1}}+3|\theta|\left(\frac{n+1}{n+2}\right)^2\sum_{i=1}^{k+1}\frac{\{1-(l_{ni}-l_{ni-1})\}^2}{n_i-n_{i-1}}+\frac{n+1}{2(n+2)}\sum_{i=1}^{k+1}\frac{1-(l_{ni}-l_{ni-1})}{n_i-n_{i-1}}\\
&+\sqrt{\frac{2}{\pi}}\left(\frac{n+1}{n+2}\right)^{1/2}\sum_{i=1}^{k+1}\left[\frac{1-(l_{ni}-l_{ni-1})}{n_i-n_{i-1}}\right]^{1/2}+\frac{2}{3}\sqrt{\frac{2}{\pi}}\left(\frac{n+1}{n+2}\right)^{3/2}\sum_{i=1}^{k+1}\frac{\{1-(l_{ni}-l_{ni-1})\}^{3/2}}{(n_i-n_{i-1})^{1/2}}\\
&+\frac{2}{3}\sqrt{\frac{2}{\pi}}\left(\frac{n+1}{n+2}\right)^{3/2}\sum_{i=1}^{k+1}\left[\frac{1-(l_{ni}-l_{ni-1})}{n_i-n_{i-1}}\right]^{3/2}+3|\theta|\left(\frac{n+1}{n+2}\right)^2\sum_{i=1}^{k+1}\left[\frac{1-(l_{ni}-l_{ni-1})}{n_i-n_{i-1}}\right]^2
\end{aligned}
$$

を得る．よって

$$
\begin{aligned}
&\left|I^*(g_n,h_n;A_{n(k)}^{L,M})\right|<\\
&\frac{k}{n+1}\left\{1-\frac{1}{12k}-\frac{3}{4(n+2)}-\frac{1}{12(n+2)^2}\right\}\\
&+\frac{1}{360(n+1)(n+2)(n+3)}+\frac{1}{32(n+1)^2(n+2)(n+3)}\\
&+\left[\frac{1}{12}+\frac{n+1}{2(n+2)}+3|\theta|\left(\frac{n+1}{n+2}\right)^2\right]\frac{k+1}{\min\limits_{1\le i\le k+1}(n_i-n_{i-1})}\\
&+\sqrt{\frac{2}{\pi}}\left(\frac{n+1}{n+2}\right)^{1/2}\left[1+\frac{2}{3}\left(\frac{n+1}{n+2}\right)\right]\frac{k+1}{\min\limits_{1\le i\le k+1}(n_i-n_{i-1})^{1/2}}\\
&+\frac{2}{3}\sqrt{\frac{2}{\pi}}\left(\frac{n+1}{n+2}\right)^{3/2}\frac{k+1}{\min\limits_{1\le i\le k+1}(n_i-n_{i-1})^{3/2}}+3|\theta|\left(\frac{n+1}{n+2}\right)^2\frac{k+1}{\min\limits_{1\le i\le k+1}(n_i-n_{i-1})^2}\\
&\to 0\ (n\to\infty),\ \text{if } \varepsilon_n:=\frac{k}{\min(n_i-n_{i-1})}\to 0\ (n\to\infty),
\end{aligned}
\tag{3.3.22}
$$

から式 (3.3.20) が示される．

次に式 (3.3.21) を示す．式 (3.3.1) が得られるならば，$n\to\infty$ とした時に式 (3.3.21) が示される．A を任意の正の整数 L, M に対して $A_{n(k)}^{L,M}=A_{n(k)}^{+L}\cup A_{n(k)}^{-M}$ と分解する．ここで部分集合 $A_{n(k)}^{+L}$, $A_{n(k)}^{-M}$ は

$$
A_{n(k)}^{+L}=Q_{n(k)}^0\cap\left\{Z_{(k)}=(z_1,\cdots,z_k)\ \middle|\ \begin{array}{l}1<\frac{z_i-z_{i-1}}{l_{ni}-l_{ni-1}}<1+\frac{1}{L}\\ (i=0,1,\cdots,k+1)\end{array}\right\}
\tag{3.3.23}
$$

$$
A_{n(k)}^{-M}=Q_{n(k)}^0\cap\left\{Z_{(k)}=(z_1,\cdots,z_k)\ \middle|\ \begin{array}{l}\frac{M}{M+1}<\frac{z_i-z_{i-1}}{l_{ni}-l_{ni-1}}<1\\ (i=0,1,\cdots,k+1)\end{array}\right\}
\tag{3.3.24}
$$

と与えられる．ここで $Q_{n(k)}^0$ は式 (3.3.15) で定義した集合とする．確率を評価するために，次の変換を行なう．各々の n に対して $Z_{n(k)}=(Z_{n1},\cdots,Z_{nk})$ を $N(l_{n(k)},L_{n(k)})$ に従う正規変

数とする．この時 $Z_{n(k)}$の変換された変数として次の $V_{n(k)} = (V_{n1}, \cdots, V_{nk})$ を考える:

$$V_{ni} := \frac{\sqrt{n+2}\,\{(Z_{ni} - l_{ni}) - (Z_{ni-1} - l_{ni-1})\}}{\sqrt{l_{ni} - l_{ni-1}}}\ (i = 1, \cdots, k). \tag{3.3.25}$$

この時

$$Z_{ni} - l_{ni} = \sum_{j=1}^{i} (l_{nj} - l_{nj-1})^{1/2}\, V_{ni}/\sqrt{n+2}\ (i = 1, \cdots, k), \tag{3.3.26}$$

となるので，変換のヤコービアンは $J(z_{(k)} \to \nu_{(k)}) = \prod_{i=1}^{k}(l_{ni} - l_{ni-1})^{1/2}/\sqrt{n+2}$であることから $V_{n(k)}$の確率分布は$\nu_{(k)} = (\nu_1, \cdots, \nu_k)$ に対して

$$\phi\left(\nu_{(k)}\right) = (2\pi)^{-k/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{k} \nu_i^2\right),\ (-\infty < \nu_i < \infty, i = 1, \cdots, k), \tag{3.3.27}$$

と与えられる．$V_{n(k)}$の各々の成分は独立に同一の標準正規分布に従っている．変換 (3.3.26) の下で，式 (3.3.23), (3.3.24) によって与えられる $A_{n(k)}^{+L}$, $A_{n(k)}^{-M}$に対応する次の集合を考える:

$$B_{n(k)}^{+L} = \left\{ \nu_{(k)} = (\nu_1, \cdots, \nu_k)^t \;\middle|\; \begin{array}{l} 0 \le \nu_i < \sqrt{\frac{n+2}{n+1}}\frac{\sqrt{n_i - n_{i-1}}}{L} \\ (i = 1, \cdots, k) \end{array} \right\}, \tag{3.3.28}$$

$$B_{n(k)}^{-M} = \left\{ \nu_{(k)} = (\nu_1, \cdots, \nu_k)^t \;\middle|\; \begin{array}{l} \sqrt{\frac{n+2}{n+1}}\frac{\sqrt{n_i - n_{i-1}}}{1+M} < \nu_i \le 0, \\ (i = 1, \cdots, k) \end{array} \right\}, \tag{3.3.29}$$

よって

$$\begin{aligned}
&P^{Z_{n(k)}}\left(A_{n(k)}^{L,M}\right) = P^{V_{n(k)}}\left(B_{n(k)}^{+L}\right) + P^{V_{n(k)}}\left(B_{n(k)}^{-M}\right) \\
&= (2\pi)^{-k/2} \int_{B_{n(k)}^{+L} \cup B_{n(k)}^{-M}} \exp\left[-\frac{1}{2}\sum_{i=1}^{k} \nu_i^2\right] d\nu_{(k)} \\
&= \prod_{j=1}^{k} \left\{ \frac{1}{\sqrt{2\pi}} \int_0^{\sqrt{\frac{n+2}{n+1}}\frac{\sqrt{n_i - n_{i-1}}}{L}} \exp\left(-\frac{\nu_i^2}{2}\right) \cdot d\nu_j + \frac{1}{\sqrt{2\pi}} \int_{\sqrt{\frac{n+2}{n+1}}\frac{\sqrt{n_i - n_{i-1}}}{1+M}}^{0} \exp\left(-\frac{\nu_i^2}{2}\right) \cdot d\nu_j \right\} \\
&=: \prod_{j=1}^{k} \left\{ N\left(\sqrt{\frac{n+2}{n+1}}\frac{\sqrt{n_i - n_{i-1}}}{L}\right) + N\left(\sqrt{\frac{n+2}{n+1}}\frac{\sqrt{n_i - n_{i-1}}}{1+M}\right) \right\}
\end{aligned}$$

を得る．ここで

$$N(x) := \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\nu^2/2} d\nu,\ (x > 0), \tag{3.3.30}$$

$$\frac{1}{2} - \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}x}\left\{1 + \frac{3}{2(x^2+2)}\right\} < N(x) < \frac{1}{2} - \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}x}\left\{1 - \frac{5}{2(x^2+2)}\right\} \tag{3.3.31}$$

となる．結果として，補題 A.2.1と条件 (3.3.18) を用いれば，

$$
\begin{aligned}
&P^{Z_{n(k)}}\left(A_{n(k)}^{L,M}\right)\\
&> \prod_{j=1}^{k}\left[\begin{array}{l}
\dfrac{1}{2}-\dfrac{L}{\sqrt{2\pi(n_j-n_{j-1})}}\sqrt{\dfrac{n+2}{n+1}}\left\{1+3/\left\{2\left(\dfrac{n+2}{n+1}\cdot\dfrac{n_j-n_{j-1}}{L^2}+2\right)\right\}\right\}\\
\qquad\qquad\qquad\cdot\exp\left\{-\dfrac{1}{2}\left(\dfrac{n+2}{n+1}\cdot\dfrac{n_j-n_{j-1}}{L^2}+2\right)^2\right\}+\\
\dfrac{1}{2}-\dfrac{1+M}{\sqrt{2\pi(n_j-n_{j-1})}}\sqrt{\dfrac{n+2}{n+1}}\left\{1+3/\left\{2\left(\dfrac{n+2}{n+1}\cdot\dfrac{n_j-n_{j-1}}{(1+M)^2}+2\right)\right\}\right\}\\
\qquad\qquad\qquad\cdot\exp\left\{-\dfrac{1}{2}\left(\dfrac{n+2}{n+1}\cdot\dfrac{n_j-n_{j-1}}{(1+M)^2}+2\right)^2\right\}
\end{array}\right]\\
&\to 1 \text{ as } n\to\infty,
\end{aligned}
\tag{3.3.32}
$$

を得る．よって定理が証明された． □

3.4 数値例

本節では本章で与えたいくつかの具体的な近似問題に関する数値例を示す．ここでは 3.3.1 節で与えたガンマ分布の正規近似に関する問題を考える．

定理 3.3.1の結果は漸近的な形で与えているが，これは式 (3.3.11), (3.3.12) の定量評価により得られたものである．定量的な評価によりパラメータ n を固定した時の近似具合を議論することもできる．図 3.2, 3.3は定理 3.3.1における$\xi = 0.8$ とした時の修正 K-L 情報量の上下界と確率の変化を与えたものである．この結果が示すように n が大きくなるにつれて定理 3.3.1の結果が視覚的に捉えることができる．

また，ξの取り方により近似主領域は変化するため，それを表にまとめた．ここで，$Pr_g(A_n)$ はガンマ変数が A_nに入る確率，$Pr_N(A_n)$ は正規変数が A_nに入る確率，$\overline{I}^*(f_N, f_g; A_n)$ は修正 K-L 情報量の上界，$\underline{I}^*(f_N, f_g; A_n)$ は修正 K-L 情報量の下界を表す．これらの結果を比較すると，確率はξが大きいほど 1 に近い値を示しているのに対して修正 K-L 情報量はξが小さいほど 0 に近い値を示している．ξは近似主領域の大きさを表していて，ξが大きいほど近似主領域は大きくなる．しかし近似主領域が大きくなることにより 2 つの分布間の密度の相対的な食い違いが大きくなり，修正 K-L 情報量が大きくなってしまう．すなわち，近似主

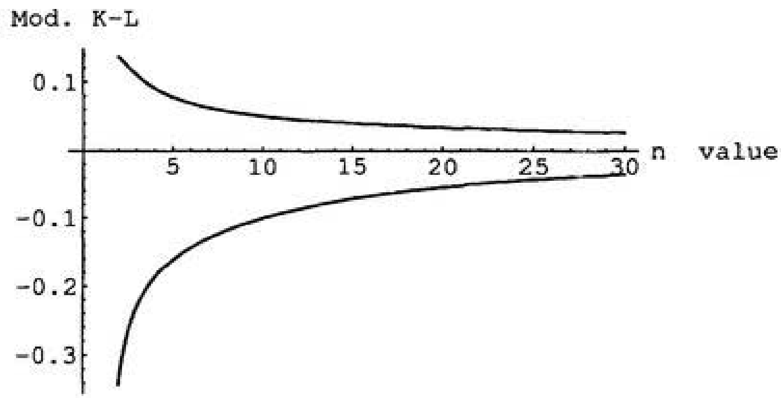

Figure 3.2: 修正 K-L 情報量の上下界

Table 3.1: $\xi = 0.6$

n	$Pr_g(A_n)$	$Pr_N(A_n)$	$\bar{I}^*(f_N, f_g; A_n)$	$\underline{I}^*(f_N, f_g; A_n)$
2	0.914524	0.858089	0.107669	-0.0976542
3	0.907007	0.867815	0.0832418	-0.0727894
4	0.906421	0.87466	0.0644419	-0.0614797
5	0.907429	0.879926	0.0513115	-0.0541318
6	0.908877	0.884197	0.0418986	-0.0487823
7	0.910421	0.887781	0.0349085	-0.0446527
8	0.91194	0.890866	0.029552	-0.0413419
9	0.913393	0.893568	0.0253378	-0.038614
10	0.914767	0.895971	0.021949	-0.0363189
20	0.924883	0.911377	0.0069458	-0.0241948
100	0.949863	0.943505	-0.0023111	-0.00937378

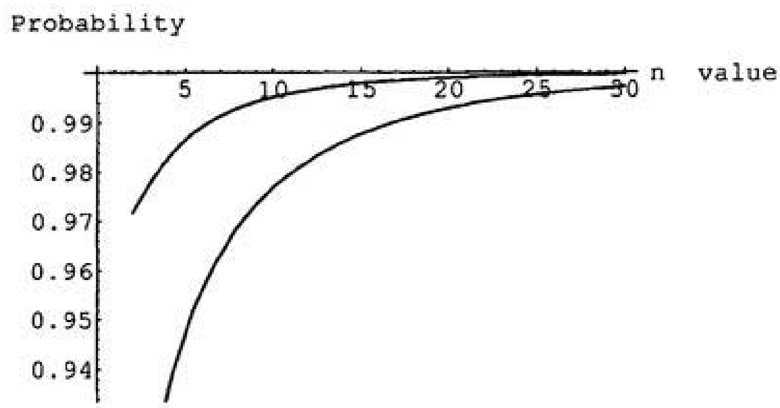

Figure 3.3: ガンマ変数と正規変数の確率分布

Table 3.2: $\xi = 0.7$

n	$Pr_g(A_n)$	$Pr_N(A_n)$	$\overline{I}^*(f_N, f_g; A_n)$	$\underline{I}^*(f_N, f_g; A_n)$
2	0.944895	0.87466	0.111624	-0.177756
3	0.946299	0.893568	0.0840582	-0.129511
4	0.950577	0.9065	0.0641946	-0.108743
5	0.954704	0.916165	0.0507492	-0.0955827
6	0.958322	0.92378	0.0413419	-0.0860406
7	0.961447	0.929998	0.0345057	-0.078636
8	0.964155	0.935204	0.0293726	-0.0726449
9	0.966519	0.93965	0.0254121	-0.067656
10	0.9686	0.943505	0.0222867	-0.0634125
20	0.980916	0.965663	0.00942655	-0.040046
100	0.996791	0.993996	0.00283539	-0.0100825

Table 3.3: $\xi = 0.8$

n	$Pr_g(A_n)$	$Pr_N(A_n)$	$\overline{I}^*(f_N, f_g; A_n)$	$\underline{I}^*(f_N, f_g; A_n)$
2	0.971745	0.890866	0.137185	-0.405307
3	0.977689	0.917795	0.109438	-0.283914
4	0.982879	0.935204	0.0903141	-0.232914
5	0.986669	0.947454	0.0776601	-0.201043
6	0.989438	0.95653	0.0688883	-0.177983
7	0.9915	0.963497	0.0624948	-0.160045
8	0.993067	0.968984	0.0576237	-0.145484
9	0.994281	0.973393	0.0537699	-0.133328
10	0.995236	0.976993	0.0506224	-0.122978
20	0.998924	0.992984	0.0342454	-0.0669853
100	0.999998	0.999966	0.0088796	-0.0110165

Table 3.4: $\xi = 0.9$

n	$Pr_g(A_n)$	$Pr_N(A_n)$	$\overline{I}^*(f_N, f_g; A_n)$	$\underline{I}^*(f_N, f_g; A_n)$
2	0.991793	0.9065	0.252629	-1.71137
3	0.995982	0.93965	0.23675	-1.17501
4	0.998013	0.959167	0.222252	-0.945116
5	0.998976	0.971522	0.209495	-0.798663
6	0.999453	0.979704	0.19765	-0.690967
7	0.999698	0.985293	0.186394	-0.606364
8	0.999829	0.989202	0.175655	-0.537487
9	0.999901	0.991985	0.165446	-0.480175
10	0.999942	0.993996	0.155795	-0.431773
20	0.999999	0.999541	0.0889189	-0.18989
100	1.	1.	0.0141961	-0.0223398

領域の補集合の確率と修正 $K-L$ 情報量のトレードオフの関係が数値的に現れる.

第 4 章

近似主領域を考慮した指数型分布族の一様近似

4.1 はじめに

第3章で近似主領域を考慮に入れた具体的な分布の近似問題を考察した．この近似主領域を用いた分布の近似は他方面で扱われる統計問題へ応用されることが期待される．本章では，近似主領域を考慮した理論的な応用として，指数型分布族の一様近似について議論する．

統計モデルの多くは指数型分布族に従う．よって，指数型分布族は統計理論において重要な立場を占めてきた．また関連分野でも早くから登場していた (cf. Boltzmann(1887), Gibbs)．この分布族のパラメータが微小に変化された分布がどのように近似されるかを考察することは理論および応用両面で興味深い．これについて Barron and Sheu(1991) は有限区間で定義される 2 つの一次元密度関数について，経験的制約条件の下で K-L 情報量最小化法に基づいて考え，有界で線形独立な関数系を用いて、指数型分布族による近似を行なった．このため彼らは多くの事柄と共に，近似に必要となる K-L 情報量に関するいくつかの両側不等式を与えた．特に，彼らは指数型分布族のパラメータの差を用いた評価を与えている．これは分布の変化が背後の考察対象の系の物理量の変化として良く表現されていると解釈される．しかし彼らの問題設定が有限区間に限定されていることと K-L 情報量の適用の仕方に物理的には必ずしも自然ではないものがある．そこで本章ではそのような制限と適用を排除し，ある領域での多変量一般指数型分布内でパラメータ等が微小に変動する時の近似問題へ拡張し，それにより修正された K-L 情報量を評価することが本章の目的

である. また K-L 情報量の解釈も Kullback(1959) の線に沿って行なう.

本章の概要は以下の通りである. 4.2節では多変量一般指数型分布を定義し, K-L 情報量の代表的な性質として有名な, 情報量収支（ピタゴラスの定理とも呼ばれる）の近似主領域を考えた場合の拡張を考える. 4.3節では多変量指数型分布のパラメータが微小変動した時の分布の近似について, 修正 Affinity, 修正 W-divergence を援用して修正 K-L 情報量の評価を与える. 4.4節では本章の主要結果であるパラメータ等が微小変動する時の分布の近似誤差の評価を考える. この結果は Barron and Sheu(1991) とは違った方法によるパラメータの差による評価を与えている.

4.2 多変量一般指数型分布族

Xを可測空間 $(R^{n\times n}, \mathbf{B}^{n\times n})$ 上で定義される $n \times n$ ランダム行列とする. ここで $R^{n\times n}$ は任意の実空間, $\mathbf{B}^{n\times n}$ は $R^{n\times n}$の部分集合の σ-集合族とする. また $u_i(X)$ を可測空間 $(R^{n\times n}, \mathbf{B}^{n\times n})$ 上で定義される実行列値関数とし, $(R^{n\times n}, \mathbf{B}^{n\times n})$ 上で定義される統計基礎モデルを Pとする. 統計基礎モデル Pをどう取るかは重要であるが, 本論文では与えられたものとする.(選び方としては例えば松縄 (1994) を参照されたい.) 統計基礎モデル Pにより誘導される大きさ m の多変量一般指数型分布族として Q^*_mとし, その密度関数を

$$dQ^*_m(X;\beta_1,\cdots,\beta_m) = \exp\left[-tr\left\{\sum_{i=1}^{m}\beta_i^t u_i(X)\right\} - \Psi(\beta_1,\cdots,\beta_m)\right]\cdot dP(X) \quad (4.2.1)$$

とする. ここで$\beta_i (i = 1,\cdots,m) (\in R^{n\times n})$ はパラメータ行列で, $\Psi(\beta_1,\cdots,\beta_m)$ は正規化定数とする. また$\Psi(\beta_1,\cdots,\beta_m)$は

$$\begin{aligned}\Psi(\beta_1,\cdots,\beta_m) &= \ln Z(\beta_1,\cdots,\beta_m) \quad (4.2.2)\\ &:= \ln\int_{R^{n\times n}}\exp\left[-tr\left\{\sum_{i=1}^{m}\beta_i^t u_i(X)\right\}\right]\cdot dP(X) < \infty\end{aligned}$$

となり $u_i(X)$ の基礎の確率測度 Pに関するキュミュラント母関数とみることができる. さて, Barron and Sheu の類推として $u_i(X)$ が微小なランダム変動をする時, この状況の下で Q^*_mの平均が存在するとし, その平均を $U_i(\beta_1,\cdots,\beta_m)$ とする:

$$U_i(\beta_1,\cdots,\beta_m) = \int_{R^{n\times n}} u_i(X)dQ^*_m \quad (4.2.3)$$

$$= -\frac{\partial}{\partial \beta_i}\Psi(\beta_1,\cdots,\beta_m)(i=1,\cdots,m) < \infty.$$

この条件の下で Q_m^* が同じ統計基礎モデルを持つ多変量一般指数型分布族の中で次の分布 Q_T に移行する状況を考える:

$$dQ_T(X;\beta_{1T},\cdots,\beta_{mT}) = \exp\left[-tr\left\{\sum_{i=1}^{m}\beta_{iT}^t u_i(X)\right\} - \Psi(\beta_T)\right]\cdot dP(X). \tag{4.2.4}$$

ここで $\beta_T = (\beta_{1T},\cdots,\beta_{mT})$ は移行した分布のパラメータ行列とする．この時，この変動によってひきおこされるゆらぎの大きさを量るため本章では修正 K-L 情報量の大きさを考える．

注 4.1 Barron and Sheu(1991) は，一変量 m 母数の指数分布族を扱っている．彼らは P を reference distribution と呼び，その確率測度がすべて $[0,1]$ にのみ分布するものに限定している．このため問題の適用範囲を狭くしてしまっている．

K-L 情報量の性質の一つとして，K-L 情報量を擬似距離と見る立場ではピタゴラスの定理が成り立つことが知られている．これは情報の移動と生成という立場から情報量収支ということもできる．近似主領域を考える場合，負になることもあるため，距離とはなり得ないが，次の結果は近似主領域を考えた場合に情報量収支がどのように変化するかを表す．

補題 4.2.1 有界な数 U_m^* が存在して，

$$\left|tr\left\{\sum_{i=1}^{m}\beta_i^t\int_A u_i(X)d(Q_T - Q_m^*)\right\}\right| \le \left|U_m^*\right|\cdot\left|Q_T(A) - Q_m^*(A)\right| \tag{4.2.5}$$

が成立するものとする．この時，次の修正された情報量収支が得られる:

$$\begin{aligned}&\left|I^*(Q_T,P;A) - I^*(Q_T,Q_m^*;A) - I^*(Q_m^*,P;A)\right| \\ &\qquad \le \left(\left|U_m^*\right| + \Psi(\beta_1,\cdots,\beta_m)\right)\left|Q_m^*(A) - Q_T(A)\right|.\end{aligned} \tag{4.2.6}$$

注 4.2 式 (4.2.6) を見てわかるように，修正 K-L 情報量は，ピタゴラスの定理とも呼ばれる情報量収支は一般には成り立たない．しかし A を近似主領域とした時，$\min\left\{Q_T(A), Q_m^*(A)\right\} \approx 1$ となるので，$I^*(Q_T,P;A) \approx I^*(Q_T,Q_m^*;A) + I^*(Q_m^*,P;A)$ が成り立つ．また特別な場合

として $A = R^{n\times n}$ とした時, 式 (4.2.6) はピタゴラスの定理そのものとなる. なお, Barron and Sheu(1991) ではここでの関係と双対な $I(P, Q_T) = I(P, Q_m^*) + I(Q_m^*, Q_T)$ が触れられている. この違いは, K-L 情報量の解釈から起きてくる. 式 (4.2.1) の Q_m^* が正準分布の一般化で, それに対応する系がどう変化するかを考察する際には, Kullback(1959) に沿った K-L 情報量の与え方の方が, 関連する大偏差確率の解釈や物理的観点から適切であると思われる (cf. Matsunawa(1995)).

証明. 対数は次のように分割することができる:

$$\ln \frac{dQ_T}{dP} = \ln \frac{dQ_T}{dQ_m^*} + \ln \frac{dQ_m^*}{dP}.$$

これにより $I^*(Q_T, P; A) - I^*(Q_T, Q_m^*; A)$ は

$$\begin{aligned}
I^*(Q_T, P; A) - I^*(Q_T, Q_m^*; A) &= \int_A dQ_T \ln \frac{dQ_T}{dP} - \int_A dQ_T \ln \frac{dQ_T}{dQ_m^*} \\
= \int_A dQ_T \ln \frac{dQ_m^*}{dP} &= \int_A dQ_T \left[-tr \left\{ \sum_{i=1}^m \beta_i^t u_i(X) \right\} - \Psi(\beta_1, \cdots, \beta_m) \right] \\
= -tr \left\{ \sum_{i=1}^m \beta_i^t \int_A u_i(X) dQ_T \right\} &- \Psi(\beta_1, \cdots, \beta_m) Q_T(A)
\end{aligned}$$

と変形できる. 一方 $I^*(Q_m^*, P; A)$ については

$$\begin{aligned}
I^*(Q_m^*, P; A) &= \int_A dQ_m^* \ln \frac{dQ_m^*}{dP} \\
&= -tr \left\{ \sum_{i=1}^m \beta_i^t \int_A u_i(X) dQ_m^* \right\} - \Psi(\beta_1, \cdots, \beta_m) Q_m^*(A)
\end{aligned}$$

となる. 従って,

$$\begin{aligned}
&I^*(Q_T, P; A) - I^*(Q_T, Q_m^*; A) - I^*(Q_m^*, P; A) \\
&= -tr \left\{ \sum_{i=1}^m \beta_i^t \int_A u_i(X) d(Q_T - Q_m^*) \right\} - \Psi(\beta_1, \cdots, \beta_m)(Q_T(A) - Q_m^*(A))
\end{aligned}$$

を得る. 条件 (4.2.5) を考慮して

$$\begin{aligned}
&\left| I^*(Q_T, P; A) - I^*(Q_T, Q_m^*; A) - I^*(Q_m^*, P; A) \right| \\
&\qquad \le \left(\left| U_m^* \right| + \Psi(\beta_1, \cdots, \beta_m) \right) \left| Q_m^*(A) - Q_T(A) \right|.
\end{aligned}$$

これにより補題 4.2.1を得る. □

4.3 多変量指数型分布の近似

本節では 4.2 節で導入した多変量一般指数型分布が Pから Q_Tへ変化した時の修正 K-L 情報量の評価を, 修正 Affinity, 修正 W-divergence を援用して考える.

本章での近似主領域の取り方については, B_εにならって次の C_εを考える:

$$C_\varepsilon = \left\{ \max\left\{ \left| \log \frac{dQ_m^*}{dP} \right|, \left| \log \frac{dQ_T}{dQ_m^*} \right| \right\} \leq \varepsilon, X \in R^{n\times n} \right\}. \tag{4.3.1}$$

これは、Q_m^*と Pの比と Q_m^*と Q_Tの比がそれぞれε近傍にあるという設定で, 自然な近似主領域の取り方といえる. また定理 3.2.1から,

$$\min\Big\{ P(C_\varepsilon), Q_m^*(C_\varepsilon), Q_T(C_\varepsilon) \Big\} \geq e^{-\varepsilon^*} \tag{4.3.2}$$

と表されることから, C_εがこれら 3 つの分布に対して近似主領域とみなすことができる. ここで$\varepsilon^* = \varepsilon + \varepsilon^\#$とし, $\varepsilon^\#$は $\max\{P(C_\varepsilon), Q_m^*(C_\varepsilon), Q_T(C_\varepsilon)\} \geq e^{\varepsilon^\#}$を満たすものとする. さらに C_εに対して条件 (4.2.5) が成立するものとする. この状況の下で次の定理が得られる.

定理 4.3.1 Q_m^*, Q_Tをそれぞれ (4.2.1), (4.2.4) で定義した統計基礎モデル Pから誘導される多変量一般指数型分布とする. また条件 (4.2.5), (4.3.2) が成り立つとする. この時, 修正 K-L 情報量について次の一連の不等式を得る:

$$I^*(Q_T, Q_m^*; C_\varepsilon) \leq \varepsilon \tag{4.3.3}$$

$$I^*(Q_T, Q_m^*; C_\varepsilon) \geq e^{-\varepsilon} - 1 \tag{4.3.4}$$

$$I^*(Q_m^*, P; C_\varepsilon) \leq 1 - e^{-\varepsilon} + \frac{1}{2}\Big\{\varepsilon \cdot u\left(e^\varepsilon\right)\Big\}^2 + \frac{1}{6}\Big\{\varepsilon \cdot u\left(e^\varepsilon\right)\Big\}^3 \tag{4.3.5}$$

$$I^*(Q_m^*, P; C_\varepsilon) \geq \left(e^{-\varepsilon} - 1\right) + \frac{1}{2}\left[\left\{l(e^{-\varepsilon})\right\}^2 \left(e^{-\varepsilon^*}\right) - \frac{1}{3}\varepsilon \cdot \left\{u(e^\varepsilon)\right\}^3\right]\varepsilon^2 \tag{4.3.6}$$

$$\begin{aligned} I^*(Q_T, P; C_\varepsilon) \leq& \left(1 - e^{-\varepsilon}\right)\left\{1 + \left|U_m^*\right| + \left|\Psi(\beta_1, \cdots, \beta_m)\right|\right\} \\ &+ \frac{1}{2}\Big\{\varepsilon \cdot u\left(e^\varepsilon\right)\Big\}^2 + \frac{1}{6}\Big\{\varepsilon \cdot u\left(e^\varepsilon\right)\Big\}^3 + \varepsilon \end{aligned} \tag{4.3.7}$$

$$\begin{aligned} I^*(Q_T, P; C_\varepsilon) \geq& \left(e^{-\varepsilon} - 1\right)\left\{2 + \left|U_m^*\right| + \left|\Psi(\beta_1, \cdots, \beta_m)\right|\right\} \\ &+ \frac{1}{2}\left[\left\{l(e^{-\varepsilon})\right\}^2 \left(e^{-\varepsilon^*}\right) - \frac{1}{3}\varepsilon \cdot \left\{u\left(e^\varepsilon\right)\right\}^3\right]\varepsilon^2 . \end{aligned} \tag{4.3.8}$$

ここで $l(t)$, $u(t)$, ε^* はそれぞれ (A.1.6), (A.1.7), (3.2.2) で定義されたものとする.

注 4.3 εを0に収束させることにより, 式 (4.3.3)-(4.3.8) の全ての修正 K-L 情報量は0へ収束することがわかる. よって考慮の対象としている分布が集合 C_ε上に確率測度のほとんどを保持していることが分かれば, 考察している分布間の近似は意味のあることが言える.

証明. 本定理の証明は主に Affinity, W-divergence と K-L 情報量との関係を与えた補題 2.3.1, 定理 2.3.2と近似主領域における情報量収支を与えた補題 4.2.1を用いる.

式 (4.3.3) , 式 (4.3.4) の証明.

修正 Affinity と修正 K-L 情報量との関係を与えた補題 2.3.1および系 2.3.1を用いて Q_m^* と Q_Tの修正 K-L 情報量の上下界の評価について考える. $I^*(Q_T, Q_m^*; C_\varepsilon)$ の上界については式 (2.3.6) より

$$I^*(Q_T, Q_m^*; C_\varepsilon) \le \int_{C_\varepsilon} \left| \ln \frac{dQ_T}{dQ_m^*} \right| dQ_T \le \varepsilon Q_T(C_\varepsilon) \le \varepsilon$$

を得る. $I^*(Q_T, Q_m^*; C_\varepsilon)$ の下界については系 2.3.1より

$$\begin{aligned} I^*(Q_T, Q_m^*; C_\varepsilon) &\ge Q_T(C_\varepsilon) \ln \left\{ Q_T(C_\varepsilon)/Q_m^*(C_\varepsilon) \right\} \\ &\ge Q_T(C_\varepsilon) - Q_m^*(C_\varepsilon) \ge -\left| Q_T(C_\varepsilon) - Q_m^*(C_\varepsilon) \right| \end{aligned}$$

と与えられる. ここで下界を与えるためには, 分布関数の差の評価が必要になる. これは補題 2.2.2を用いて

$$I^*(Q_m^*, Q_T; C_\varepsilon) \ge -\left| Q_m^*(C_\varepsilon) - Q_T(C_\varepsilon) \right| \ge e^{-\varepsilon} - 1$$

と表される. これより式 (4.3.3), (4.3.4) が得られる.

式 (4.3.5) , 式 (4.3.6) の証明.

修正 W-divergence と修正 K-L 情報量との関係を与える式 (2.3.10), (2.3.11) を用いる. 系 2.3.1より $W^*(Q_m^*, P; C_\varepsilon)$ の上下界は

$$\left\{ l(e^{-\varepsilon}) \cdot \varepsilon \right\}^2 \cdot P(C_\varepsilon) \le W^*(Q_m^*, P; C_\varepsilon) \le \left\{ u(e^{\varepsilon}) \cdot \varepsilon \right\}^2 \cdot P(C_\varepsilon)$$

と表される．また式 (2.3.10), (2.3.11) の右辺にある上界の評価が必要となるのだが，これは補題 A.1.2の不等式を用いて

$$\sup_{X\in C_\varepsilon}\left|\frac{dQ_m^*(X)}{dP(X)}-1\right|\leq u(e^\varepsilon)\cdot\varepsilon$$

と表される．これより $I^*(Q_m^*,P;C_\varepsilon)$ の上下界はそれぞれεを使った評価が得られる：

$$\begin{aligned}I^*(Q_m^*,P;C_\varepsilon)&\leq Q_m^*(C_\varepsilon)-P(C_\varepsilon)+\frac{1}{2}\left(1+\frac{1}{3}\sup_{X\in C_\varepsilon}\left|\frac{dQ_m^*(X)}{dP(X)}-1\right|\right)W^*(Q_m^*,P;C_\varepsilon)\\&\leq 1-e^{-\varepsilon}+\frac{1}{2}\Big\{\varepsilon\cdot u\left(e^\varepsilon\right)\Big\}^2+\frac{1}{6}\left\{\varepsilon\cdot u\left(e^\varepsilon\right)\right\}^3\end{aligned}$$

$$\begin{aligned}I^*(Q_m^*,P;C_\varepsilon)&\geq Q_m^*(C_\varepsilon)-P(C_\varepsilon)+\frac{1}{2}\left(1-\frac{1}{3}\sup_{X\in C_\varepsilon}\left|\frac{dQ_m^*(X)}{dP(X)}-1\right|\right)W^*(Q_m^*,P;C_\varepsilon)\\&\geq\left(e^{-\varepsilon}-1\right)+\frac{1}{2}\left[\left\{l(e^{-\varepsilon})\right\}^2\left(e^{-\varepsilon^*}\right)-\frac{1}{3}\varepsilon\cdot\left\{u\left(e^\varepsilon\right)\right\}^3\right]\varepsilon^2.\end{aligned}$$

式 (4.3.7) , 式 (4.3.8) の証明.

以上の結果と 4.2節で述べた修正された情報量収支を用いて Q_Tと Pの修正 K-L 情報量の評価を考える．式 (4.2.6) にある分布関数の差は補題 2.2.2を用いて得られるので，$I^*(Q_T,P;C_\varepsilon)$ の上下界は

$$\begin{aligned}&I^*(Q_T,P;C_\varepsilon)\\&\quad\leq I^*(Q_T,Q_m^*;C_\varepsilon)+I^*(Q_m^*,P;C_\varepsilon)+\left(\left|U_m^*\right|+\left|\Psi(\beta_1,\cdots\beta_m^*)\right|\right)\cdot\left|Q_T(C_\varepsilon)-Q_m^*(C_\varepsilon)\right|\\&\quad\leq 1-e^{-\varepsilon}+\frac{1}{2}\left\{\varepsilon\cdot u\left(e^\varepsilon\right)\right\}^2+\frac{1}{6}\left\{\varepsilon\cdot u\left(e^\varepsilon\right)\right\}^3+\varepsilon\\&\quad+\left(\left|U_m^*\right|+\left|\Psi(\beta_1,\cdots\beta_m^*)\right|\right)\left(1-e^{-\varepsilon}\right)\\&\quad=\left(1-e^{-\varepsilon}\right)\left\{1+\left|U_m^*\right|+\left|\Psi(\beta_1,\cdots,\beta_m)\right|\right\}\\&\quad+\frac{1}{2}\Big\{\varepsilon\cdot u\left(e^\varepsilon\right)\Big\}^2+\frac{1}{6}\Big\{\varepsilon\cdot u\left(e^\varepsilon\right)\Big\}^3+\varepsilon\end{aligned}$$

$$\begin{aligned}&I^*(Q_T,P;C_\varepsilon)\\&\quad\geq I^*(Q_T,Q_m^*;C_\varepsilon)+I^*(Q_m^*,P;C_\varepsilon)-\left(\left|U_m^*\right|+\left|\Psi(\beta_1,\cdots\beta_m^*)\right|\right)\cdot\left|Q_T(C_\varepsilon)-Q_m^*(C_\varepsilon)\right|\end{aligned}$$

$$
\begin{aligned}
&2(e^{-\varepsilon}-1)+\frac{1}{2}\left[\left\{l\left(e^{-\varepsilon}\right)\right\}^2\left(e^{-\varepsilon^*}\right)-\frac{1}{3}\varepsilon\cdot\{u\left(e^{\varepsilon}\right)\}^3\right]\varepsilon^2\\
&\qquad+\left(\left|U_m^*\right|+\left|\Psi(\beta_1,\cdots\beta_m^*)\right|\right)\left(e^{-\varepsilon}-1\right)\\
&=(e^{-\varepsilon}-1)\left\{2+\left|U_m^*\right|+\left|\Psi(\beta_1,\cdots,\beta_m)\right|\right\}\\
&\qquad+\frac{1}{2}\left[\left\{l(e^{-\varepsilon})\right\}^2\left(e^{-\varepsilon^*}\right)-\frac{1}{3}\varepsilon\cdot\left\{u\left(e^{\varepsilon}\right)\right\}^3\right]\varepsilon^2
\end{aligned}
$$

と与えられる.

以上の一連の証明により定理4.3.1の一連の不等式が証明された. □

ここまでは修正K-L情報量の評価を考えてきたが,一様近似の意味を明確に表現するための結果を示す. ここでは一様誤差に基づく評価を行なう.

系 4.3.1 今 C_ε上での確率 $Q_T(C_\varepsilon)$, $P(C_\varepsilon)$ が次の条件を満たすものとする:

$$\max\{P(C_\varepsilon),Q_T(C_\varepsilon)\}\geq e^{\varepsilon^\#}. \tag{4.3.9}$$

この時一様誤差 $D(Q_T,P;\mathbf{B}^{n\times n}):=D$に対して次の不等式が成り立つ:

$$
\begin{aligned}
D&\geq l(e^{-2\varepsilon})\cdot\varepsilon\cdot e^{\varepsilon^\#}\\
D&\leq u(e^{2\varepsilon})\cdot\varepsilon+e^{-2\varepsilon+\varepsilon^\#}.
\end{aligned}
$$

ここで $l(t)$, $u(t)$ は (A.1.6), (A.1.7) で与えられたものとし, $I_a^*(Q_T,P;C_\varepsilon)$ は

$$I_a^*(Q_T,P;C_\varepsilon):=\int_{C_\varepsilon}\left|\ln\frac{dQ_T}{dP}\right|dQ_T \tag{4.3.10}$$

で定義されたものとする.

証明

C_εの定義により Q_Tと Pの関係は

$$-2\varepsilon\leq\log\left(\frac{dQ_T}{dP}\right)=\log\left(\frac{dQ_T}{dQ_m^*}\right)+\log\left(\frac{dQ_m^*}{dP}\right)\leq 2\varepsilon \tag{4.3.11}$$

と与えられる．よって上記の設定の下では，

$$
\begin{aligned}
D \quad &\geq \frac{1}{2}\left(l\left(\inf_{C_\epsilon} \frac{dP}{dQ_T}\right) \cdot |I^*(Q_T, P; C_\epsilon)| + |Q_T(C_\epsilon) - P(C_\epsilon)| \right) && (4.3.12)\\
&\geq \frac{1}{2}|Q_T(C_\epsilon) - P(C_\epsilon)| \\
&\frac{1}{2}\left|\frac{\min\{Q_T(C_\epsilon), P(C_\epsilon)\}}{\max\{Q_T(C_\epsilon), P(C_\epsilon)\}} - 1\right| \cdot \max\{Q_T(C_\epsilon), P(C_\epsilon)\} && (4.3.13)\\
&\geq l(e^{-2\epsilon}) \cdot (2\varepsilon) \cdot e^{\epsilon^\#} \\
D \quad &\leq \frac{1}{2}u\left(\sup_{C_\epsilon} \frac{dP}{dQ_T}\right) |I_a^*(Q_T, P; C_\epsilon)| + \left\{1 - \frac{Q_T(C_\epsilon) + P(C_\epsilon)}{2}\right\} && (4.3.14)\\
&\leq \frac{1}{2}u\left(\sup_{C_\epsilon} \frac{dP}{dQ_T}\right) I_a^*(Q_T, P; C_\epsilon) + \left\{1 - \min\{Q_T(C_\epsilon), P(C_\epsilon\}\right\} && (4.3.15)\\
&\leq u(e^{2\epsilon}) \cdot \varepsilon + e^{-2\varepsilon + \epsilon^\#}
\end{aligned}
$$

を得る．

4.4 パラメータの微小変動に対する修正 K-L 情報量の揺動の変化

多変量一般指数型分布族に属する分布 dQ_m^* が，$u_i\Big(X\Big)(i=1,\cdots,m)$ の微小変動によってパラメータが $\{\beta_1,\cdots,\beta_m\} \to \{\beta_{1T},\cdots,\beta_{mT}\}$ と変化し，同じ分布族内の dQ_T に移行する場合を考える．このような設定は統計的にも物理的にも自然である．我々の関心は，この場合の二つの分布の変動を，パラメータの変化を用いてどう評価できるか，$u_i(X)(i=1,\cdots,m)$ の平均値を使ってどう評価できるかである．このことを以下で考察する．このために，dQ_m^*，dQ_T を少し扱い易い形で考える．

dQ_m^*，dQ_T およびそれらを誘導する基礎分布 dP が可測空間 $(R^{n\times n}, \mathbf{B})$ 上で定義される σ－有限測度 $\mu^{n\times n}$ に関して絶対連続とし，

$$
\begin{aligned}
&dP = P(A)d\mu^{n\times n}, dQ_m^* = q(X;\beta_1,\cdots,\beta_m)p(X)d\mu^{n\times n}, \\
&dQ_T = q(X;\beta_{1T},\cdots,\beta_{mT})p(X)d\mu^{n\times n} \\
&q(X;\theta_1,\cdots,\theta_m) = \exp\left[-tr\left\{\sum_{i=1}^m \theta_i^t u_i(X)\right\} - \Psi(\theta_1,\cdots,\theta_m)\right] \cdot p(X)
\end{aligned}
$$

と表す．ここで $q(X;\cdot)$ は考察対象の指数分布族の $\mu^{n\times n}$ に関する Radon-Nikodym 導関数を表し，$\theta_i(i=1,\cdots,m)\in R^{n\times n}$ で一般の正準パラメータ行列を代表させる．

ところで $q(X;\cdot)$ は存在が仮定されているがその型は明示されていない．そこで，以下で $\theta_i(i=1,\cdots,m)$ に連動させて $q(X;\theta_1,\cdots,\theta_m)=:q$ と記し，q がある可測集合 $K(\in\mathbf{B})$ 上で Kullback(1959) に準じた正則条件を満たすものとする．つまり

(a) 任意の $A\in\mathbf{B}$ と，その要素が非退化区間 $\min(\beta_{i\mu\nu},\beta_{iT\mu\nu})<\ \beta^*_{i\mu\nu}\ <\max(\beta_{i\mu\nu},\beta_{iT\mu\nu})$, $i=1,\cdots m;\mu,\nu=1,\cdots,n$ に属する任意の正準パラメータ行列 $\beta^*_i=(\beta^*_{i\mu\nu})[\beta^*_i\in(((\beta_i,\beta_{iT}))$ と以下で略記する$]$ に対して，$\dfrac{\partial}{\partial\theta_i}\ln p$, $\dfrac{\partial}{\partial\theta^t_i}\dfrac{\partial}{\partial\theta_i}\ln p$ が存在する．

(b) 任意の $\beta^*_i\in((\beta_i,\beta_{iT}))$ に対して，$\left|\dfrac{\partial\ln p}{\partial\beta_{i\alpha\beta}}\right|<L(A)$, $\left|\dfrac{\partial^2\ln p}{\partial\beta_{i\alpha\beta}\partial\beta_{i\alpha\beta_T}}\right|<M(A)$ が全ての $i=1,\cdots,m$ に対して成立するものとする．ここで $\|\cdot\|$ は次の意味での行列の Frobenius ノルムを表す (cf. Golub and van Loan (1989), Chatelin(1993)):

$$\|A\|=\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_ij\right|^2\right)^{1/2}=tr\left(A^tA\right)\ \text{for}\ A=(a_{ij})\in R^{n\times n}$$

また $L(A)$, $M(A)$ は K 上で可積分 $[\mu^{n\times n}]$ で，$\sup\limits_{A\in K}|L(A),M(A)/2|<H<\infty$ を満たし，H は θ_i, $i=1,\cdots,m$ と関数的に独立とする．

(c) パラメータ行列による微分と，A に関する積分の順序が交換可能とし，この時，分布 dQ_T と dQ^*_m に関連して

$$\begin{aligned}
\frac{\partial}{\partial\beta_{iT}}(-\Psi(\beta_{1T},\cdots,\beta_{mT}))&=\int_{R^{n\times n}}u_i(X)q(X;\beta_{1T},\cdots,\beta_{mT})d\mu^{n\times n}(X^t)\\
&=:U_{iT}(\beta_{1T},\cdots,\beta_{mT})\\
\frac{\partial}{\partial\beta_{iT}}\frac{\partial}{\partial\beta_{iT}}\Psi(\beta_{1T},\cdots,\beta_{mT})&=\int_{R^{n\times n}}(u_i(X)-U_{iT})^t(u_i(X)-U_{iT})qd\mu^{n\times n}(X^t)=:G_i\\
\frac{\partial}{\partial\beta_i}(-\Psi(\beta_1,\cdots,\beta_m))&=\int_{R^{n\times n}}u_i(X)q(X;\beta_1,\cdots,\beta_m)d\mu^{n\times n}(X^t)=:U_i(\beta_1,\cdots,\beta_m)
\end{aligned}$$

の諸量の存在と関係式が成り立つものとする．

これらの条件の下，正準パラメータが $\{\beta_1,\cdots,\beta_m\}\to\{\beta_{1T},\cdots,\beta_{mT}\}$ と変化したとき，$I^*(Q_T,Q^*_m;K)$ の変動がどうなるかを知ることは多変量一般指数型分布族の揺動の定量評

価という意味で, 統計理論および物理的観点からも大いに興味のあることである. このことについて, 次の命題を与える.

定理 4.4.1 上記の正則条件の下で次の不等式が成立する:

$$\left|\ln \frac{dQ_T}{dQ_m^*}\right| \le H \cdot \sum_{i=1}^{m} \max\left\{\|\beta_i - \beta_{iT}\|, \|\beta_i - \beta_{iT}\|^2\right\} \tag{4.4.1}$$

$$\begin{aligned} I^*(Q_T, Q_m^*; K) \le \sum_{i=1}^{m} \Big[& \|\beta_i - \beta_{iT}\| \cdot \|U_{iT}^* - U_{iT} \cdot Q_T^*\| \\ & + \frac{1}{2} tr\left\{G_i^* (\beta_i - \beta_{iT})^t (\beta_i - \beta_{iT})\right\} \cdot Q_T^* \Big] \end{aligned} \tag{4.4.2}$$

$$\begin{aligned} I^*(Q_T, Q_m^*; K) \ge \sum_{i=1}^{m} \Big[& - \|\beta_i - \beta_{iT}\| \cdot \|U_{iT}^* - U_{iT} \cdot Q_T^*\| \\ & + \frac{1}{2} tr\left\{G_i^* (\beta_i - \beta_{iT})^t (\beta_i - \beta_{iT})\right\} \cdot Q_T^* \Big]. \end{aligned} \tag{4.4.3}$$

ここで, Hは正則条件 (b) で定義した量, G_i^*, Q_T^*, U_{iT}^*はそれぞれ次の諸量を表す:

$$\begin{aligned} G_i &:= \left(\frac{\partial}{\partial \beta_{iT}} \frac{\partial}{\partial \beta_{iT}} \Psi(\beta_{1T}, \cdots, \beta_{mT})\right)_{\beta_{iT} = \beta^*} \\ Q_T^* &:= Q_T(K; \beta_{1T}, \cdots, \beta_{mT}) = \int_K q(X) d\mu^{n\times n}(X) \\ U_{iT}^* &:= U_{iT}^*(\beta_{1T}, \cdots, \beta_{mT}) = \int_K u_i(X) q(X) d\mu^{n\times n}(A). \end{aligned}$$

注 4.4 Barron and Sheu (1991) は, $\ln dQ_m^*$, $\ln dQ_T$が有界かつ線形独立な直交関数系で張られる線形空間に属するものとして, 上の命題に対応する一変量の場合の結果を得ている. そこでの有界性が彼らの証明の本質的な役割をしているが, 統計理論の観点からはその意味が理解しにくい. 本論文では, 彼らの接近法と違って, 情報量密度比 $\ln(dQ_T/dQ_m^*)$ に Taylor 展開を用いて K-L 情報量の近似につなげるという Kullback (1959) の接近法を多変量の場合に適用する. この方法では, $\{\beta_1, \cdots, \beta_m\} \to \{\beta_{1T}, \cdots, \beta_{mT}\}$ の変化量だけでなく, その係数としていわゆる Fisher 情報量が現れる利点がある. しかし Barron and Sheu の接近法にはこの利点が現れない.

証明. 正則条件の下, $\beta_i = \beta_{iT} + (\beta_i - \beta_{iT}) =: \beta_{iT} + \Delta\beta_{iT}$, $(i = 1, \cdots, m)$ として $dQ_m^* = q(A; \beta_{1T} + \Delta\beta_{1T}, \cdots, \beta_{mT} + \Delta\beta_{mT}) d\mu^{n\times n}(A)$ に Taylor 近似を適用して

$$\ln \frac{dQ_T}{dQ_m^*} = tr\left[\sum_{i=1}^{m} (\beta_i - \beta_{iT})^t \frac{\partial}{\partial \beta_i} \ln q + \frac{1}{2} \sum_{i=1}^{m} (\beta_i - \beta_{iT})^t (\beta_i - \beta_{iT}) \left(\frac{\partial}{\partial \beta_i^t} \frac{\partial}{\partial \beta_i} \ln q\right)_{\beta^*}\right] \tag{4.4.4}$$

を得る．ここで$\beta_i^* \in ((\beta_i, \beta_{iT}))$ は$\beta_i^*(i=1,\cdots,m)$ の各要素が，β_iとβ_{iT}の対応する要素の間の値を取るパラメータ行列であることを示す．即ち，$\beta^* = (\beta_i + \eta_1(\beta_1 - \beta_{iT}),\cdots,\beta_m + \eta_m(\beta_m - \beta_{mT}))$, $\eta_i = (\eta_{ij}) \in R^{n\times n} 0 < \eta_{ij} < 1, i, j = 1,\cdots,m$ を意味する．正則条件 (b) によって

$$\begin{aligned}
\left|\ln \frac{dQ_T}{dQ_m^*}\right| &\leq \sum_{i=1}^{m} \left\{\|\beta_i - \beta_{iT}\| \cdot L(A) + \frac{1}{2}\|\beta_i - \beta_{iT}\|^2 \cdot M(A)\right\} \\
&\leq \sup_{A\in R^{n\times n}} \{L(A), M(A)/2\} \cdot \sum_{i=1}^{m} \max\left\{\|\beta_i - \beta_{iT}\|, \|\beta_i - \beta_{iT}\|^2\right\}
\end{aligned}$$

と与えられる．正則条件 (b) の最後の条件により，式 (4.4.1) が従う．次に式 (4.4.2) を証明する．式 (4.4.4) を K上で dQ_Tで積分する．正則条件 (c) から

$$\begin{aligned}
tr\left[\int_K \left(\frac{\partial}{\partial \beta_{iT}} \ln q\right) q d\mu^{n\times n}\right] &= -tr(U_{iT}^* - U_{iT} \cdot Q_T^*) \\
tr\left[\int_K \left(\frac{\partial}{\partial \beta_{iT}^t} \frac{\partial}{\partial \beta_{iT}} \ln q\right)_{\beta_{iT}=\beta_i^*} \cdot q d\mu^{n\times n}\right] &= -tr G_i^* \cdot Q_T^*
\end{aligned}$$

であるから

$$\begin{aligned}
I^*(Q_T, Q_{\mathrm{m}}^*; K) &= \sum_{i=1}^{m} tr\left[(\beta_i - \beta_{iT})^t (U_{iT}^* - U_{iT} \cdot Q_T^*) + \frac{1}{2}\left\{(\beta_i - \beta_{iT})^t (\beta_i - \beta_{iT}) G_i^* \cdot Q_T^*\right\}\right] \\
&\leq \sum_{i=1}^{m} \left\{\|\beta_i - \beta_{iT}\| \cdot \|U_{iT}^* - U_{iT} \cdot Q_T^*\| + \frac{1}{2} tr\left\{G_i^* (\beta_i - \beta_{iT})^t (\beta_i - \beta_{iT})\right\} \cdot Q_t^*\right]
\end{aligned}$$

これにより不等式 (4.4.2) が証明された．不等式 (4.4.3) も同様に得られる． □

付録 A

誤差評価に有用な関数の近似展開

A.1 $|t-1|$ の近似展開

$\log t$ と$\left|t-1\right|$の不等式関係を次のように与える.

補題 A.1.1 実数 $t>0,\ \alpha\ (\neq 0)$ に対して, 次の不等式が成り立つ.

$$\ln t = (t-1)\left(\frac{2}{t^{1/\alpha}+1}\right)^{\alpha} + O\left(\frac{\alpha^2(3-\alpha)}{3}\left(\frac{t^{1/\alpha}-1}{t^{1/\alpha}+1}\right)^3\right) \tag{A.1.1}$$

$$\ln t = (t-1)\left(\frac{2}{t^{1/3}+1}\right)^{3} + O\left(\frac{1}{5}\left(\frac{t^{1/3}-1}{t^{1/3}+1}\right)^5\right) \tag{A.1.2}$$

注 A.1 補題 A.1.1の結果により $\ln t$ の$\left|t-1\right|$による近似を考える場合, $t\to t^{1/3}$なる立方根変換を利用することにより, $\ln t$ のさらに優れた近似が求まる.

証明. $\ln t$ を変形して

$$\begin{aligned}
\ln t &= \alpha\ln t^{1/\alpha} = -\alpha\ln\left\{1+\left(1-t^{1/\alpha}\right)/t^{1/\alpha}\right\} \\
&= -\alpha\ln\left[\left\{1+\frac{\left(1-t^{1/\alpha}\right)/t^{1/\alpha}}{2+\left(1-t^{1-\alpha}\right)\,/\,t^{1/\alpha}}\right\}/\left\{1-\frac{\left(1-t^{1/\alpha}/t^{1/\alpha}\right.}{2+\left(1-t^{1/\alpha}\right)\,/\,t^{1/\alpha}}\right\}\right] \\
&= -\alpha\ln\left\{\left(1+\frac{1-t^{1/\alpha}}{1+t^{1/\alpha}}\right)\,/\,\left(1-\frac{1-t^{1/\alpha}}{1+t^{1/\alpha}}\right)\right\} =: -\alpha\ln\left(\frac{1+u}{1-u}\right)
\end{aligned}$$

ここで

$$u := \frac{1-t^{1/\alpha}}{1+t^{1/\alpha}},\ (-1<u<1)$$

とする．これを展開して次のように表される:

$$\ln t = -2\alpha\left(u + \frac{1}{3}u^3 + \frac{1}{5}u^5 + \frac{1}{7}u^7 + \cdots\right). \tag{A.1.3}$$

一方, $t-1$ の量を含ませた, 式 (A.1.1) の右辺の主要項は

$$\begin{aligned} 2^\alpha(t-1)\left(1+t^{1-\alpha}\right)^{-\alpha} &= \left(1 - \frac{1-t^{1/\alpha}}{1+t^{1-\alpha}}\right)^\alpha - \left(1 + \frac{1-t^{1/\alpha}}{1+t^{1/\alpha}}\right)^\alpha = (1-u)^\alpha - (1+u)^\alpha \\ &= -2\left\{\begin{pmatrix}\alpha\\1\end{pmatrix}u + \begin{pmatrix}\alpha\\3\end{pmatrix}u^3 + \begin{pmatrix}\alpha\\5\end{pmatrix}u^5 + \cdots\right\} \end{aligned} \tag{A.1.4}$$

と表現できる．そこで, 誤差を評価するために, $(A.1.3) - (A.1.4)$ を計算すると,

$$\begin{aligned} &\ln t - 2^\alpha(t-1)\left(1+t^{1-\alpha}\right)^{-\alpha} \\ &= -2\alpha\left[\frac{1}{3}\left\{1 - \frac{(\alpha-1)(\alpha-2)}{2!}\right\}u^3 + \frac{1}{5}\left\{1 - \frac{(\alpha-1)(\alpha-2)(\alpha-3)(\alpha-4)}{4!}\right\}u^5 + \cdots\right] \end{aligned}$$

と与えられる．従って, $\ln t$ の精密な近似を得るには右辺の u^3の係数を 0 と置けば良い．つまり

$$1 - \frac{(\alpha-1)(\alpha-2)}{2!} = 0$$

が成立する．$\alpha \neq 0$ と仮定したから, $\alpha = 3$ と設定すれば良い．これにより式 (A.1.2) が成立する． □

補題 A.1.2 実数 $t > 0$ に対して, 次の不等式が成り立つ:

$$l(t)|\ln t| \le |t-1| \le u(t)|\ln t|. \tag{A.1.5}$$

ここで

$$l^*(t) = \frac{210t^{1/3}(t^{1/3}+1)^7(1+t^{1/3}+t^{2/3})}{\begin{array}{c}35 + 1832t^{1/3} + 7796t^{2/3} + 18968t + 23378t^{4/3} \\ +18968t^{5/3} + 7796t^2 + 1832t^{7/3} + 35t^{8/3}\end{array}} \tag{A.1.6}$$

$$u^*(t) = \frac{35(t^{1/3}+1)^5(1+t^{1/3}+t^{2/3})(t^{1/3}+16t^{1/3}+1)}{\begin{array}{c}457 + 5014t^{1/3} + 14471t^{2/3} + 20596t \\ +14471t^{4/3} + 5014t^{5/3} + 457t^2\end{array}} \tag{A.1.7}$$

とする．

証明. 次のよく知られた展開を利用すると

$$\begin{aligned} S = \frac{1}{2x}\ln\frac{1+x}{1-x} \quad &= 1 + \frac{x^2}{3} + \frac{x^4}{5} + \frac{x^6}{7} + \frac{x^8}{9} + \frac{x^{10}}{11} + \cdots \\ &< 1 + \frac{x^2}{3} + \frac{x^4}{5} + \frac{x^6}{7} + \frac{x^8}{9}(1 + x^2 + x^4 + \cdots) \\ &= 1 + \frac{x^2}{3} + \frac{x^4}{5} + \frac{x^6}{7} + \frac{x^8}{9}\cdot\frac{1}{1-x^2} \\ &= \frac{945 - 630x^2 - 126x^4 - 54x^6 - 30x^8}{945(1-x^2)} \end{aligned}$$

と得られる．ここで $x = \frac{y-1}{y+1}$ とすると，$S = \frac{y+1}{2(y-1)}\ln y$ と表される．これより

$$\begin{aligned} \frac{\ln y}{y-1} &= \frac{2}{y+1}S < \frac{2}{y+1}\cdot\frac{945 - 630(\frac{y-1}{y+1})^2 - 126(\frac{y-1}{y+1})^4 - 54(\frac{y-1}{y+1})^6 - 30(\frac{y-1}{y+1})^8}{945(1-(\frac{y-1}{y+1})^2)} \\ &= \frac{2}{y+1}\cdot\frac{\begin{matrix}945(y+1)^8 - 630(y-1)^2(y+1)^6 - 126(y-1)^4(y+1)^4 \\ -54(y-1)^6(y+1)^2 - 30(y-1)^8\end{matrix}}{945\cdot\frac{4y}{(1+y)^2}\cdot(y+1)^8} \\ &= \frac{35 + 1832y + 7796y^2 + 18968y^3 + 23378y^4 + 18968y^5 + 7796y^6 + 1832y^7 + 35y^8}{630y(y+1)^7} \end{aligned}$$

と与えられることにより，

$$\frac{630y(y+1)^7}{\begin{matrix}35 + 1832y + 7796y^2 + 18968y^3 + 23378y^4 \\ +18968y^5 + 7796y^6 + 1832y^7 + 35y^8\end{matrix}}|\ln y| < |y-1|.$$

ここで $y = t^{1/3}$ とし，$1 + t^{1/3} + t^{2/3}$ をかける立方根変換を行なうことにより

$$l^*(t)|\ln t| < |t-1|$$

を得る．ここで

$$l^*(t) = \frac{210t^{1/3}(t^{1/3}+1)^7(1+t^{1/3}+t^{2/3})}{\begin{matrix}35 + 1832t^{1/3} + 7796t^{2/3} + 18968t + 23378t^{4/3} \\ +18968t^{5/3} + 7796t^2 + 1832t^{7/3} + 35t^{8/3}\end{matrix}}$$

とする．また $\log t$ の下界についても同様な方法により得られる．まず

$$S = \frac{1}{2x}\ln\frac{1+x}{1-x} \quad = 1 + \frac{x^2}{3} + \frac{x^4}{5} + \frac{x^6}{7} + \frac{x^8}{9} + \frac{x^{10}}{11} + \cdots$$

$$
\begin{aligned}
&> 1 + \frac{x^2}{3} + \frac{x^4}{5} + \frac{x^6}{7}(1 + \frac{7}{9}x^2 + (\frac{7}{9}x)^4 + \cdots) \\
&= 1 + \frac{x^2}{3} + \frac{x^4}{5} + \frac{x^6}{7} \cdots \frac{1}{1 - 7/9x^2} \\
&= \frac{945 - 420x^2 - 56x^4 - 12x^6}{105(9 - 7x^2)}
\end{aligned}
$$

を得ることにより

$$
\begin{aligned}
\frac{\ln y}{y-1} &= \frac{2}{y+1}S > \frac{2}{y+1} \cdot \frac{945 - 420(\frac{y-1}{y+1})^2 - 56(\frac{y-1}{y+1})^4 - 12(\frac{y-1}{y+1})^6}{105(9 - 7(\frac{y-1}{y+1})^2)} \\
&= \frac{2}{y+1} \cdot \frac{945(y+1)^6 - 420(y-1)^2(y+1)^4 - 56(y-1)^4(y+1)^2 - 12(y-1)^6}{105 \cdot \frac{2y^2+32y+2y}{(y+1)^2} \cdot (y+1)^6} \\
&= \frac{457 + 5014y + 14471y^2 + 20596y^3 + 14471y^4 + 5014y^5 + 457y^6}{105(y+1)^5(y^2+16y+1)}
\end{aligned}
$$

$$
|y-1| < \frac{105(y+1)^5(y^2+16y+1)}{457 + 5014y + 14471y^2 + 20596y^3 + 14471y^4 + 5014y^5 + 457y^6} |\ln y|
$$

と与えられる．ここで立方根変換を行なうことにより

$$
|t-1| < u^*(t)|\ln t|
$$

を得る．ここで

$$
u^*(t) = \frac{35(t^{1/3}+1)^5(1 + t^{1/3} + t^{2/3})(t^{1/3} + 16t^{1/3} + 1)}{457 + 5014t^{1/3} + 14471t^{2/3} + 20596t + 14471t^{4/3} + 5014t^{5/3} + 457t^2}.
$$

とする．これにより補題 A.1.2を得た． □

注 A.2 $\log t$ の不等式として次の関係がよく知られている：

$$
|\log t| \leq |t-1|.
$$

これは対数の線形近似を行なっているのだが，この補題はさらに精度のよい結果が得られた．

注 A.3 補題の証明の中の近似の精度を変えることにより逐次的に上下界を得ることができる．分母の次数が i 次である関数をそれぞれ $u_i(t)$, $l_i(t)$ とすると以下のように表される：

$$
u_2(t) := \frac{(t^{1/3}+1)(1 + t^{1/3} + t^{2/3})(1 + 8t^{1/3} + t^{2/3})}{11 + 38t^{1/3} + 11t^{2/3}}
$$

$$u_4(t) := \frac{5(t^{1/3}+1)^3(1+t^{1/3}+t^{2/3})(1+12t^{1/3}+t^{2/3})}{61+436t^{1/3}+686t^{2/3}+436t+61t^{4/3}}$$

$$u_6(t) := \frac{35(t^{1/3}+1)^5(1+t^{1/3}+t^{2/3})(t^{2/3}+16t^{1/3}+1)}{457+5014t^{1/3}+14471t^{2/3}+20596t+14471t^{4/3}+5014t^{5/3}+457t^2}$$

$$u_8(t) := \frac{315(t^{2/3}+20t^{1/3}+1)(t^{1/3}+1)^7(t^{2/3}+t^{1/3}+1)}{\begin{array}{c}4323+64344t^{1/3}+271380t^{2/3}+604776t+771474t^{4/3}\\ +604776t^{5/3}+271380t^2+64344t^{7/3}+4323t^{8/3}\end{array}}$$

$$l_4(t) := \frac{10t^{1/3}(t^{1/3}+1)^3(1+t^{1/3}+t^{2/3})}{3+68t^{1/3}+98t^{2/3}+68t+3t^{4/3}}$$

$$l_6(t) := \frac{70t^{1/3}(t^{1/3}+1)^5(1+t^{1/3}+t^{2/3})}{15+554t^{1/3}+1569t^{2/3}+2444t+1569t^{4/3}+554t^{5/3}+15t^2}$$

$$l_8(t) := \frac{210t^{1/3}(t^{1/3}+1)^7(1+t^{1/3}+t^{2/3})}{\begin{array}{c}35+1832t^{1/3}+7796t^{2/3}+18968t+23378t^{4/3}\\ +18968t^{5/3}+7796t^2+1832t^{7/3}+35t^{8/3}\end{array}}.$$

この $l_i(t)$, $u_i(t)$ を用いて，誤差の評価を行なったものが表 A.1 と表 A.2 である．*が付いているものは計算機の精度上，負の値が得られたものである．この結果を見てもわかるように i が大きくなるにつれてさらによい精度の近似を得ることができる．また $l_i(t)$, $u_i(t)$ の関数の動きを見るために特に $l_8(t)$ と $u_6(t)$ のグラフを与えたのが図 A.1 と図 A.2 である．この 2 つの関数が単調増加関数であることは，t に関する微分を実行することにより確かめ得る．この図からも単調性が保たれていることがわかる．

注 A.4 Matsunawa(1982) では $l_4(t)$, $u_2(t)$ を得ている．我々はこの結果を改良し，さらによい近似を与えた．

Table A.1: $|t-1|-l_i[t]|\ln t|$

項数	4	6	8
t=0.1	1.47×10^{-4}	1.11×10^{-5}	9.52×10^{-7}
t=0.2	1.61×10^{-5}	6.18×10^{-7}	2.71×10^{-8}
t=0.3	2.53×10^{-6}	5.53×10^{-8}	1.38×10^{-9}
t=0.4	4.27×10^{-7}	5.46×10^{-9}	7.99×10^{-11}
t=0.5	6.72×10^{-8}	4.95×10^{-10}	4.17×10^{-12}
t=0.6	8.66×10^{-9}	3.47×10^{-11}	1.59×10^{-13}
t=0.7	7.54×10^{-10}	1.48×10^{-12}	3.44×10^{-15}
t=0.8	3.02×10^{-11}	2.33×10^{-14}	2.78×10^{-17}
t=0.9	1.67×10^{-13}	0.00	$^{*}2.78 \times 10^{-17}$
t=1.0	0.00	0.00	0.00
t=1.1	9.18×10^{-14}	0.00	0.00
t=1.2	8.99×10^{-12}	4.55×10^{-15}	$^{*}5.55 \times 10^{-17}$
t=1.3	1.20×10^{-10}	1.27×10^{-13}	1.11×10^{-16}
t=1.4	7.09×10^{-10}	1.24×10^{-12}	2.61×10^{-15}
t=1.5	2.71×10^{-9}	6.86×10^{-12}	2.00×10^{-14}
t=1.6	7.88×10^{-9}	2.68×10^{-11}	1.05×10^{-13}
t=1.7	1.90×10^{-8}	8.23×10^{-11}	4.08×10^{-13}
t=1.8	4.01×10^{-8}	2.13×10^{-10}	1.29×10^{-12}
t=1.9	7.64×10^{-8}	4.82×10^{-10}	3.49×10^{-12}
t=2.0	1.34×10^{-7}	9.89×10^{-10}	8.33×10^{-12}

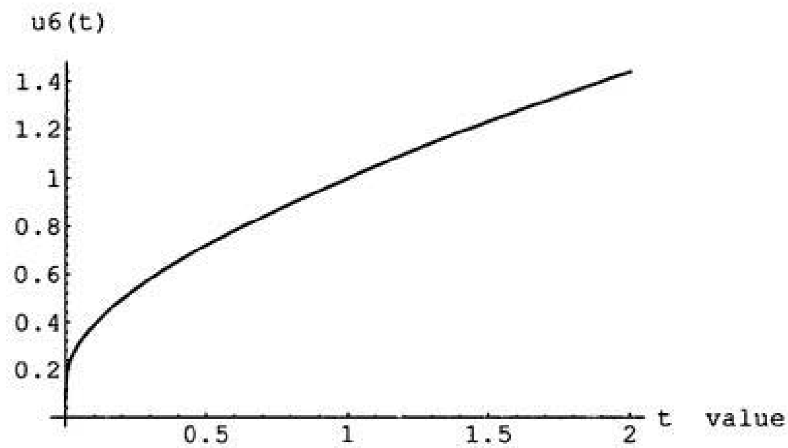

Figure A.1: $l_8(t)$ のグラフ

Table A.2: $u_i[t]|\ln t| - |t-1|$

項数	2	4	6	8
t=0.1	6.00×10^{-5}	3.32×10^{-6}	2.26×10^{-7}	1.74×10^{-8}
t=0.2	6.51×10^{-6}	1.81×10^{-7}	6.22×10^{-9}	2.44×10^{-10}
t=0.3	1.02×10^{-6}	1.60×10^{-8}	3.13×10^{-10}	6.99×10^{-12}
t=0.4	1.71×10^{-7}	1.57×10^{-9}	1.79×10^{-11}	2.35×10^{-13}
t=0.5	2.69×10^{-8}	1.42×10^{-10}	9.31×10^{-13}	6.99×10^{-15}
t=0.6	3.47×10^{-9}	9.94×10^{-12}	3.55×10^{-14}	1.67×10^{-16}
t=0.7	3.02×10^{-10}	4.23×10^{-13}	6.66×10^{-16}	*1.67×10^{-16}
t=0.8	1.21×10^{-11}	6.63×10^{-15}	*2.78×10^{-17}	*1.11×10^{-16}
t=0.9	6.70×10^{-14}	4.16×10^{-17}	4.16×10^{-17}	4.16×10^{-17}
t=1.0	0.00	0.00	0.00	0.00
t=1.1	3.67×10^{-14}	4.16×10^{-17}	0.00	0.00
t=1.2	3.60×10^{-12}	1.36×10^{-15}	5.55×10^{-17}	1.11×10^{-16}
t=1.3	4.79×10^{-11}	3.64×10^{-14}	5.55×10^{-17}	5.55×10^{-17}
t=1.4	2.84×10^{-10}	3.54×10^{-13}	2.78×10^{-16}	*2.22×10^{-16}
t=1.5	1.09×10^{-9}	1.96×10^{-12}	4.11×10^{-15}	*3.33×10^{-16}
t=1.6	3.16×10^{-9}	7.67×10^{-12}	2.26×10^{-14}	*8.88×10^{-16}
t=1.7	7.62×10^{-9}	2.36×10^{-11}	9.06×10^{-14}	0.00
t=1.8	1.61×10^{-8}	6.10×10^{-11}	2.89×10^{-13}	2.66×10^{-15}
t=1.9	3.06×10^{-8}	1.38×10^{-10}	7.80×10^{-13}	5.77×10^{-15}
t=2.0	5.39×10^{-8}	2.84×10^{-10}	1.86×10^{-12}	1.33×10^{-14}

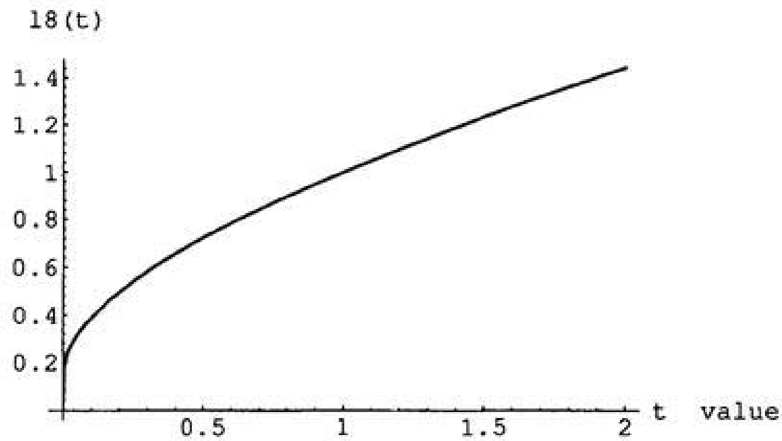

Figure A.2: $u_6(t)$ のグラフ

A.2 $\log(1+\frac{1}{u})$ の近似展開

補題 A.2.1 任意の $u \geq 0$, $v \geq 1$ に対して,

$$\log\left(1+\frac{1}{u}\right) \geq \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} - \frac{1}{3u^3(u+1)} \tag{A.2.1}$$

$$\log\left(1+\frac{1}{u}\right) \leq \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} - \frac{u+2}{6u^3(u+1)^2} \tag{A.2.2}$$

$$\leq \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} - \frac{1}{6u^3(u+1)} \tag{A.2.3}$$

$$\log\left(1-\frac{1}{v}\right) \geq -\frac{1}{v} - \frac{1}{2v^2} - \frac{1}{3v^3} - \frac{1}{3v^3(v-1)} \tag{A.2.4}$$

$$\log\left(1-\frac{1}{v}\right) \leq -\frac{1}{v} - \frac{1}{2v^2} - \frac{1}{3v^3} + \frac{1}{6v^2(v-1)^2} \tag{A.2.5}$$

注 A.5 式 (A.2.1)-(A.2.5) までの精度を視覚的に捉えるため, これらの関数をグラフに表す. この結果を見てもかなり良い近似が行なえている.

証明. 次の不等式 (cf. Matsunawa(1976)) を変形することにより, 補題の不等式を与える:

$$\log\left(1+\frac{1}{u}\right) \geq \frac{1}{u} - \frac{1}{2u(u+1)} - \frac{1}{6u^2(u+1)} \tag{A.2.6}$$

$$\log\left(1+\frac{1}{u}\right) \leq \frac{1}{u} - \frac{1}{2u(u+1)} - \frac{1}{6u(u+1)^2} \tag{A.2.7}$$

$$\log\left(1-\frac{1}{v}\right) \geq -\frac{1}{v-1} + \frac{1}{2v(v-1)} + \frac{1}{6v^2(v-1)} \tag{A.2.8}$$

$$\log\left(1-\frac{1}{v}\right) \leq -\frac{1}{v-1} + \frac{1}{2v(v-1)} + \frac{1}{6v(v-1)^2} \tag{A.2.9}$$

式 (A.2.6), (A.2.7) を書き換えると, $\log\left(1+\frac{1}{u}\right)$の上下界は次のように与えられる:

$$\begin{aligned}
\log\left(1+\frac{1}{u}\right) &\leq \frac{1}{u} - \frac{1}{2u(u+1)} - \frac{1}{6u(u+1)^2} \\
&= \frac{1}{u} - \frac{1}{2u^2} + \left(\frac{1}{2u^2} - \frac{1}{2u(u+1)}\right) - \frac{1}{6u(u+1)^2} \\
&= \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{2u^2(u+1)} - \frac{1}{6u(u+1)^2} \\
&= \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} - \left(\frac{1}{3u^3} - \frac{1}{2u^2(u+1)} + \frac{1}{6u(u+1)^2}\right) \\
&= \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} - \frac{u+2}{6u^3(u+1)^2}
\end{aligned}$$

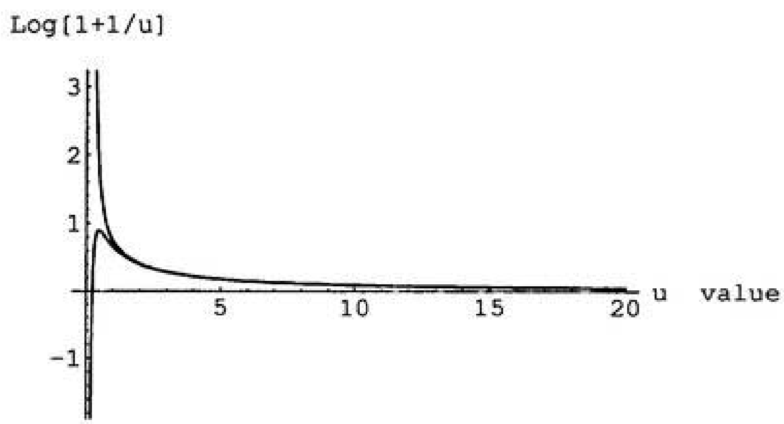

Figure A.3: $\log(1+1/u)$ の上下界

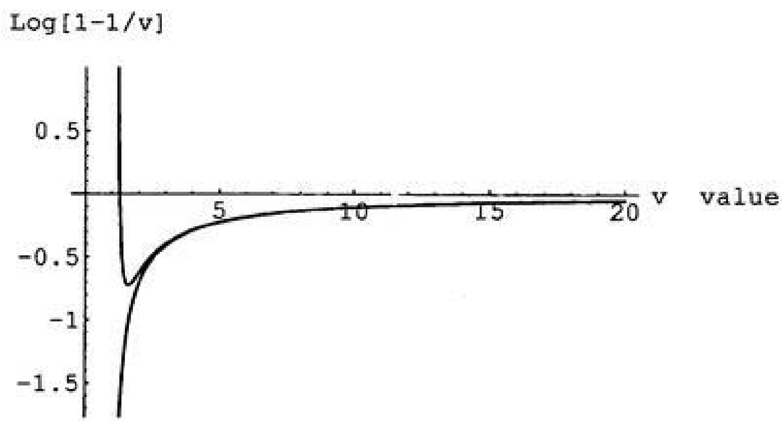

Figure A.4: $\log(1-1/u)$ の上下界

$$\leq \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} - \left(\inf_u \frac{u+2}{u+1}\right) \cdot \frac{1}{6u^3(u+1)}$$

$$\begin{aligned}
\log\left(1+\frac{1}{u}\right) &\geq \frac{1}{u} - \frac{1}{2u(u+1)} - \frac{1}{6u^2(u+1)} \\
&= \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{2u^2(u+1)} - \frac{1}{6u^2(u+1)} \\
&= \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^2(u+1)} \\
&= \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} + \left(\frac{1}{3u^2(u+1)} - \frac{1}{3u^3}\right) \\
&= \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} - \frac{1}{3u^3(u+1)}.
\end{aligned}$$

ここで上界については, $\inf_u \frac{u+2}{u+1} = 1$ より

$$\log\left(1+\frac{1}{u}\right) \leq \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} - \frac{1}{6u^3(u+1)}$$

と与えられる. また式 (A.2.8), (A.2.9) を書き換えると, $\log\left(1-\frac{1}{v}\right)$の上下界は

$$\begin{aligned}
\log\left(1-\frac{1}{v}\right) &\leq -\frac{1}{v-1} + \frac{1}{2v(v-1)} + \frac{1}{6v(v-1)^2} \\
&= -\frac{1}{v} + \left(\frac{1}{v-1} - \frac{1}{v-1} + \frac{1}{2v(v-1)}\right) + \frac{1}{6v(v-1)^2} \\
&= -\frac{1}{v} - \frac{1}{2v(v-1)} + \frac{1}{6v(v-1)^2} \\
&= -\frac{1}{v} - \frac{1}{2v^2} + \left(\frac{1}{2v^2} - \frac{1}{2v(v-1)}\right) + \frac{1}{6v(v-1)^2} \\
&= -\frac{1}{v} - \frac{1}{2v^2} - \frac{1}{2v^2(v-1)} + \frac{1}{6v(v-1)^2} \\
&= -\frac{1}{v} - \frac{1}{2v^2} - \frac{1}{3v^2(v-1)} + \left(\frac{1}{3v^2(v-1)} - \frac{1}{2v^2(v-1)}\right) + \frac{1}{6v(v-1)^2} \\
&= -\frac{1}{v} - \frac{1}{2v^2} - \frac{1}{3v^2(v-1)} - \frac{1}{6v^2(v-1)} + \frac{1}{6v(v-1)^2} \\
&= -\frac{1}{v} - \frac{1}{2v^2} - \frac{1}{3v^2(v-1)} + \frac{1}{6v^2(v-1)^2} \\
&\leq -\frac{1}{v} - \frac{1}{2v^2} - \frac{1}{3v^3} + \frac{1}{6v^2(v-1)^2} \\
\log\left(1-\frac{1}{v}\right) &\geq -\frac{1}{v-1} + \frac{1}{2v(v-1)} + \frac{1}{6v^2(v-1)} \\
&= -\frac{1}{v} - \frac{1}{2v^2} - \frac{1}{2v^2(v-1)} + \frac{1}{6v^2(v-1)}
\end{aligned}$$

$$
\begin{aligned}
&= -\frac{1}{v} - \frac{1}{2v^2} - \frac{1}{3v^2(v-1)} \\
&= -\frac{1}{v} - \frac{1}{2v^2} - \frac{1}{3v^3} + \left(\frac{1}{3v^3} - \frac{1}{3v^2(v-1)}\right) \\
&= -\frac{1}{v} - \frac{1}{2v^2} - \frac{1}{3v^3} - \frac{1}{3v^3(v-1)}
\end{aligned}
$$

と得られる．これにより式 (A.2.6), 式 (A.2.7), 式 (A.2.8), 式 (A.2.9) によって得られた不等式として次の不等式が得られる：

$$
\frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} - \frac{1}{3u^3(u+1)} \leq \log\left(1 + \frac{1}{u}\right) \leq \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^2(u+1)} - \frac{1}{6u^3(u+1)}
$$

$$
-\frac{1}{v} - \frac{1}{2v^2} - \frac{1}{3v^3} - \frac{1}{3v^3(v-1)} \leq \log\left(1 - \frac{1}{v}\right) \leq -\frac{1}{v} - \frac{1}{2v^2} - \frac{1}{3v^3} + \frac{1}{6v^2(v-1)^2}.
$$

これにより補題 A.2.1が証明された． □

補題 A.2.1での結果は精度の良い近似が行なえた．しかし多変量などの複雑な計算では少々難しさを伴う．よってもっと簡単な形でまとめる．

系 A.2.1 L, Mを有限な正数とする．この時, $u > L$, $u < -1 - M$をみたす任意の u に対して

$$
\ln\left(1 + \frac{1}{u}\right) = \frac{1}{u} - \frac{1}{2u^2} + \frac{1}{3u^3} - \frac{\theta}{u^4} \tag{A.2.10}
$$

と表現できる．ただし θ は

$$
\frac{1}{6}\left\{1 - \frac{1}{(L+1)^2}\right\} < \theta = \theta(L, M) < \max\left\{\frac{1}{3}, \frac{1}{4}\left(1 + \frac{1}{M}\right)\right\}
$$

を満たす実数である．

注 A.6 式 (A.2.10) の利点は, u の符号に関わらず不等式が成り立つということである．

証明. $u > 0$ の時, 式 (A.2.1), (A.2.2) の最後の項に対して

$$
\begin{aligned}
&-\frac{1}{3u^3(u+1)} \geq -\frac{1}{3} \\
&-\frac{u+2}{6u^4(u+1)^2} = -\frac{1}{6u^4} + \frac{1}{6u^4(u+1)^2} < -\frac{1}{6}\left\{1 - \frac{1}{(L+1)^2}\right\}
\end{aligned}
$$

を得る．よってこれらを組み合わせることにより

$$\ln\left(1+\frac{1}{u}\right)=\frac{1}{u}-\frac{1}{2u^2}+\frac{1}{3u^3}-\frac{\theta}{u^4} \tag{A.2.10}$$

と与えられる．ここで$\frac{1}{6}\left\{1-\frac{1}{(L+1)^2}\right\}<\theta<\frac{1}{3}$である．

次に $u<0$ の時の $\ln\left(1+\frac{1}{u}\right)$の上下界を考える．この時，$u<-1$ の時のみが意味がある．$0<-1/u<1$ であるから，Maclaulin 展開を用いることにより，

$$\begin{aligned}\ln\left(1+\frac{1}{u}\right)&=\ln\left(1-\left(-\frac{1}{u}\right)\right)\\&=-\left(-\frac{1}{u}\right)-\frac{1}{2}\left(-\frac{1}{u}\right)^2-\frac{1}{3}\left(-\frac{1}{u}\right)^3-\frac{1}{4}\left(-\frac{1}{u}\right)^4-\frac{1}{5}\left(-\frac{1}{u}\right)^5-\cdots\\&<\frac{1}{u}-\frac{1}{2u^2}+\frac{1}{3u^3}-\frac{1}{4u^4}\end{aligned}$$

を得る．また下界に関しては $u<-1-M$ (M: 任意の正数) に対して

$$\begin{aligned}\ln\left(1+\frac{1}{u}\right)&>\frac{1}{u}-\frac{1}{2u^2}+\frac{1}{3u^3}-\frac{1}{4u^4}\cdot\left(\frac{u}{u+1}\right)\\&>\frac{1}{u}-\frac{1}{2u^2}+\frac{1}{3u^3}-\frac{1}{4u^4}\cdot\left(1+\frac{1}{M}\right)\end{aligned}$$

を得る．これらを組み合わせて

$$\ln\left(1+\frac{1}{u}\right)=\frac{1}{u}-\frac{1}{2u^2}+\frac{1}{3u^3}+\frac{\theta}{u^4} \tag{A.2.11}$$

を得る．ここで$\frac{1}{4}<\theta<\frac{1}{4}\left(1+\frac{1}{M}\right)$である．式 (A.2.10), (A.2.11) を組み合わせることにより系 A.2.1が証明される．　□

付録 B

具体的な近似計算に要する諸量の計算

B.1 一変量正規分布に基づいたある領域における積分計算

$f_N(x)$ を平均μ, 分散σ^2の正規分布の密度関数とする．つまり

$$f_N(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

とする．今 a, b を $0 < a \leq \mu \leq b < \infty$ を満たす数とし, 領域 $A = \{x; a < x < b\}$ を考える．この時, 正規分布に基づいた領域 A における積分をそれぞれ

$$\mu^* = \int_A x f_N(x) dx \tag{B.1.1}$$

$$\sigma^{2*} = \int_A (x-\mu)^2 f_N(x) dx \tag{B.1.2}$$

$$J = \int_A \log x f_N(x) dx. \tag{B.1.3}$$

とする．本節ではμ^*, σ^{2*}, Jの定量評価について考える．なお, 本節では

$$\gamma(a,b) := \int_0^b \frac{1}{\Gamma(a)} x^{a-1} e^{-x}$$

を定義し, $\gamma(a,b)$ を用いた定量評価を考えている．

補題 B.1.1 正規変数 x が領域 A に入る確率 $Pr(X \in A)$ は次の様に表される．

$$Pr(X \in A) = \frac{1}{2\sqrt{\pi}} \left\{ \Gamma(\frac{1}{2}) \left(\gamma(\frac{1}{2}, a^*) + \gamma(\frac{1}{2}, b^*) \right) \right\}$$

また正規分布に基づいた近似主領域での積分 μ^*, σ^{2*}は次の様に表される:

$$\mu^* = \frac{1}{2\sqrt{\pi}}\left\{\mu\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)-\sqrt{2\sigma^2}\Gamma(1)\left(\gamma(1,a^*)-\gamma(1,b^*)\right)\right\}$$
$$\sigma^{2*} = \frac{\sigma^2\Gamma(\frac{3}{2})}{\sqrt{\pi}}\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right).$$

ここで $a^* = \dfrac{(a-\mu)^2}{2\sigma^2}$, $b^* = \dfrac{(b-\mu)^2}{2\sigma^2}$とする.

証明. $\dfrac{(x-\mu)^2}{2\sigma^2} = y$とした変換を考えると, $x \leq \mu$と $x \geq \mu$では違ったヤコービアンを持っていて, それぞれ

$$dx = \frac{\sqrt{2\sigma^2}}{2}y^{-1/2}dy \quad (x \geq \mu)$$
$$dx = -\frac{\sqrt{2\sigma^2}}{2}y^{-1/2}dy \quad (x \leq \mu)$$

と与えられる. これにより領域 A に入る確率 $Pr(x \in A)$ は

$$\begin{aligned} Pr(x \in A) &= \frac{1}{2\sqrt{\pi}}\left\{\int_0^{a^*} y^{-1/2}e^{-y}dy + \int_0^{b^*} y^{-1/2}e^{-y}dy\right\} \\ &= \frac{1}{2\sqrt{\pi}}\left\{\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)\right\} \end{aligned}$$

と得られる. またμ^*は

$$\begin{aligned} \mu^* &= -\int_{a^*}^0 \left(\mu-\sqrt{2\sigma^2 y}\right) f_N(\mu-\sqrt{2\sigma^2 y})\frac{\sqrt{2\alpha}}{2}y^{-1/2}dy \\ &\quad + \int_0^{b^*} \left(\mu+\sqrt{2\sigma^2 y}\right) f_N(\mu+\sqrt{2\sigma^2 y})\frac{\sqrt{2\alpha}}{2}y^{-1/2}dy \\ &= \frac{1}{2\sqrt{\pi}}\left\{\int_0^{a^*}\left(\mu-\sqrt{2\sigma^2 y}\right)y^{-1/2}e^{-y}dy + \int_0^{b^*}\left(\mu+\sqrt{2\sigma^2 y}\right)y^{-1/2}e^{-y}dy\right\} \\ &= \frac{1}{2\sqrt{\pi}}\left\{\mu\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)-\sqrt{2\sigma^2}\Gamma(1)\left(\gamma(1,a^*)-\gamma(1,b^*)\right)\right\} \end{aligned}$$

と与えられる. またσ^{2*}は

$$\begin{aligned} \sigma^{2*} &= -\int_{a^*}^0 2\sigma^2 y f_N(\mu-\sqrt{2\sigma^2 y})\frac{\sqrt{2\alpha}}{2}y^{-1/2}dy + \int_0^{b^*} 2\sigma^2 y f_N(\mu+\sqrt{2\sigma^2 y})\frac{\sqrt{2\alpha}}{2}y^{-1/2}dy \\ &= \frac{2\sigma^2}{2\sqrt{\pi}}\left\{\int_0^{a^*} y^{1/2}e^{-y}dy + \int_0^{b^*} y^{1/2}e^{-y}dy\right\} \\ &= \frac{\sigma^2\Gamma(\frac{3}{2})}{\sqrt{\pi}}\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right) \end{aligned}$$

と与えられる. これにより補題 B.1.1が証明される. □

補題 B.1.2 log関数の正規分布に基づいた積分 Jの上下の評価は次のように与えられる:

$$
\begin{aligned}
J \leq \quad & \frac{1}{2\sqrt{\pi}}\Biggl\{\log\mu\cdot\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)-\frac{\sqrt{2\sigma^2}}{\mu}\Gamma(1)\left(\gamma(1,a^*)-\gamma(1,b^*)\right)\\
& -\frac{\sigma^2}{\mu^2}\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)-\frac{(2\sigma^2)^{3/2}}{3\mu^3}\Gamma(2)\left(\gamma(2,a^*)-\gamma(2,b^*)\right)\\
& +\frac{2(\sigma^2)^2}{3\mu^4}\Gamma(\frac{5}{2})\left(\frac{\mu^2}{a^2}\gamma(\frac{5}{2},a^*)-\frac{\mu}{b}\gamma(\frac{5}{2},b^*)\right)\Biggr\}\\
J \geq \quad & \frac{1}{2\sqrt{\pi}}\Biggl\{\log\mu\cdot\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)-\frac{\sqrt{2\sigma^2}}{\mu}\Gamma(1)\left(\gamma(1,a^*)-\gamma(1,b^*)\right)\\
& -\frac{\sigma^2}{\mu^2}\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)-\frac{(2\sigma^2)^{3/2}}{3\mu^3}\Gamma(2)\left(\gamma(2,a^*)-\gamma(2,b^*)\right)\\
& -\frac{4(\sigma^2)^2}{3\mu^4}\Gamma(\frac{5}{2})\left(\frac{\mu}{a}\gamma(\frac{5}{2},a^*)+\gamma(\frac{5}{2},b^*)\right)\Biggr\}.
\end{aligned}
$$

ここで $a^*=\dfrac{(\mu-a)^2}{2\sigma^2}$, $b^*=\dfrac{(b-\mu)^2}{2\sigma^2}$ とし, $\gamma(s,t)$ は

$$
\gamma(s,t)=\int_0^t\frac{1}{\Gamma(s)}x^{s-1}e^{-x}dx
$$

で定義されたものとする.

証明. 補題 B.1.1で得たヤコービアンを用いると, Jは

$$
\begin{aligned}
J = \quad & -\int_{a^*}^{0}\log\left(\mu-\sqrt{2\sigma^2y}\right)f_N(\mu-\sqrt{2\sigma^2y})\frac{\sqrt{2\alpha}}{2}y^{-1/2}dy\\
& +\int_0^{b^*}\log\left(\mu+\sqrt{2\sigma^2y}\right)f_N(\mu+\sqrt{2\sigma^2y})\frac{\sqrt{2\alpha}}{2}y^{-1/2}dy\\
= \quad & \int_0^{a^*}\log\left(\mu-\sqrt{2\sigma^2y}\right)f_N(\mu-\sqrt{2\sigma^2y})\frac{\sqrt{2\alpha}}{2}y^{-1/2}dy\\
& +\int_0^{b^*}\log\left(\mu+\sqrt{2\sigma^2y}\right)f_N(\mu+\sqrt{2\sigma^2y})\frac{\sqrt{2\alpha}}{2}y^{-1/2}dy
\end{aligned}
$$

と与えられる. ここで $a^*=\dfrac{(\mu-a)^2}{2\sigma^2}$, $b^*=\dfrac{(b-\mu)^2}{2\sigma^2}$とする. 今 J_1, J_2を

$$
\begin{aligned}
J_1 \quad & =\int_0^{a^*}\log\left(\mu-\sqrt{2\sigma^2y}\right)f_N(\mu-\sqrt{2\sigma^2y})\frac{\sqrt{2\alpha}}{2}y^{-1/2}dy\\
& =\frac{1}{2\sqrt{\pi}}\int_0^{a^*}\log\left(\mu-\sqrt{2\sigma^2y}\right)y^{-1/2}e^{-y}dy\\
J_2 \quad & =\int_0^{b^*}\log\left(\mu+\sqrt{2\sigma^2y}\right)f_N(\mu+\sqrt{2\sigma^2y})\frac{\sqrt{2\alpha}}{2}y^{-1/2}dy\\
& =\frac{1}{2\sqrt{\pi}}\int_0^{b^*}\log\left(\mu+\sqrt{2\sigma^2y}\right)y^{-1/2}e^{-y}dy
\end{aligned}
$$

と定義する．この時, 不等式 (A.2.1), 式 (A.2.3) を用いると, $\log\left(\mu-\sqrt{2\sigma^2 y}\right)$の上下界は次のように与えられる:

$$\log\left(\mu-\sqrt{2\sigma^2 y}\right) \le \log\mu - \frac{\sqrt{2\sigma^2 y}}{\mu} - \frac{2\sigma^2 y}{2\mu^2} - \frac{(2\sigma^2 y)^{3/2}}{3\mu^3} + \frac{(2\sigma^2 y)^2}{6\mu^3(\mu-\sqrt{2\sigma^2 y})}$$
$$\log\left(\mu-\sqrt{2\sigma^2 y}\right) \ge \log\mu - \frac{\sqrt{2\sigma^2 y}}{\mu} - \frac{2\sigma^2 y}{2\mu^2} - \frac{(2\sigma^2 y)^{3/2}}{3\mu^3} - \frac{(2\sigma^2 y)^2}{3\mu^3(\mu-\sqrt{2\sigma^2 y})}.$$

これにより J_1の上界は

$$\begin{aligned}
&J_1 \\
&\le \frac{1}{2\sqrt{\pi}}\int_0^{a^*}\left\{\log\mu - \frac{\sqrt{2\sigma^2}}{\mu}y^{1/2} - \frac{\sigma^2}{\mu^2}y - \frac{(2\sigma^2)^{3/2}}{3\mu^3}y^{3/2}\right. \\
&\qquad\left. + \frac{2(\sigma^2)^2}{3\mu^2(\mu-\sqrt{2\sigma^2 y})^2}y^2\right\} y^{-1/2}e^{-y}dy \\
&\le \frac{1}{2\sqrt{\pi}}\int_0^{a^*}\left\{\log\mu\cdot y^{-1/2} - \frac{\sqrt{2\sigma^2}}{\mu} - \frac{\sigma^2}{\mu^2}y^{1/2} - \frac{(2\sigma^2)^{3/2}}{3\mu^3}y + \frac{2(\sigma^2)^2}{3\mu^2a^2}y^{3/2}\right\}e^{-y}dy \\
&= \frac{1}{2\sqrt{\pi}}\left\{\log\mu\cdot\Gamma(\frac{1}{2})\gamma(\frac{1}{2},a^*) - \frac{\sqrt{2\sigma^2}}{\mu}\Gamma(1)\gamma(1,a^*) - \frac{\sigma^2}{\mu^2}\Gamma(\frac{3}{2})\gamma(\frac{3}{2},a^*)\right. \\
&\left. - \frac{(2\sigma^2)^{3/2}}{3\mu^3}\Gamma(2)\gamma(2,a^*) + \frac{2(\sigma^2)^2}{3\mu^2a^2}\Gamma(\frac{5}{2})\gamma(\frac{5}{2},a^*)\right\}
\end{aligned} \tag{B.1.4}$$

と与えられ, 逆に J_1の下界は

$$\begin{aligned}
&J_1 \\
&\ge \frac{1}{2\sqrt{\pi}}\int_0^{a^*}\left\{\log\mu - \frac{\sqrt{2\sigma^2}}{\mu}y^{1/2} - \frac{\sigma^2}{\mu^2}y - \frac{(2\sigma^2)^{3/2}}{3\mu^3}y^{3/2}\right. \\
&\qquad\left. - \frac{(2\sigma^2)^2}{3\mu^3(\mu-\sqrt{2\sigma^2 y})}y^2\right\} y^{-1/2}e^{-y}dy \\
&\ge \frac{1}{2\sqrt{\pi}}\int_0^{a^*}\left\{\log\mu\cdot y^{-1/2} - \frac{\sqrt{2\sigma^2}}{\mu} - \frac{\sigma^2}{\mu^2}y^{1/2} - \frac{(2\sigma^2)^{3/2}}{3\mu^3}y - \frac{4(\sigma^2)^2}{3\mu^3a}y^{3/2}\right\}e^{-y}dy \\
&= \frac{1}{2\sqrt{\pi}}\left\{\log\mu\cdot\Gamma(\frac{1}{2})\gamma(\frac{1}{2},a^*) - \frac{\sqrt{2\sigma^2}}{\mu}\Gamma(1)\gamma(1,a^*) - \frac{\sigma^2}{\mu^2}\Gamma(\frac{3}{2})\gamma(\frac{3}{2},a^*)\right. \\
&\left. - \frac{(2\sigma^2)^{3/2}}{3\mu^3}\Gamma(2)\gamma(2,a^*) - \frac{4(\sigma^2)^2}{3\mu^3a}\Gamma(\frac{5}{2})\gamma(\frac{5}{2},a^*)\right\}
\end{aligned} \tag{B.1.5}$$

を得る．次に J_2も同じ方法で上下界を与える．まず不等式 (A.2.4), (A.2.5) を用いると, $\log(\mu+\sqrt{2\sigma^2 y})$ の上下界は次のように与えられる:

$$\log(\mu+\sqrt{2\sigma^2 y}) \le \log\mu + \frac{\sqrt{2\sigma^2 y}}{\mu} - \frac{2\sigma^2 y}{2\mu^2} + \frac{(2\sigma^2 y)^{3/2}}{3\mu^3} - \frac{(2\sigma^2 y)^2}{6\mu^3(\mu+\sqrt{2\sigma^2 y})}$$

$$\log(\mu+\sqrt{2\sigma^2 y}) \geq \log\mu + \frac{\sqrt{2\sigma^2 y}}{\mu} - \frac{2\sigma^2 y}{2\mu^2} + \frac{(2\sigma^2 y)^{3/2}}{3\mu^3} - \frac{(2\sigma^2 y)^2}{3\mu^3(\mu+\sqrt{2\sigma^2 y})}.$$

これにより J_2の上界は

$$\begin{aligned}
J_2 & \\
&\leq \frac{1}{2\sqrt{\pi}}\int_0^{b^*}\left\{\log\mu + \frac{\sqrt{2\sigma^2}}{\mu}y^{1/2} - \frac{\sigma^2}{\mu^2}y + \frac{(2\sigma^2)^{3/2}}{3\mu^3}y^{3/2}\right. \\
&\qquad\left. + \frac{2(\sigma^2)^2}{3\mu^3(\mu+\sqrt{2\sigma^2 y})}y^2\right\} y^{-1/2}e^{-y}dy \\
&\leq \frac{1}{2\sqrt{\pi}}\int_0^{b^*}\left\{\log\mu\cdot y^{-1/2} + \frac{\sqrt{2\sigma^2}}{\mu} - \frac{\sigma^2}{\mu^2}y^{1/2} + \frac{(2\sigma^2)^{3/2}}{3\mu^3}y - \frac{2(\sigma^2)^2}{3\mu^4}y^{3/2}\right\} e^{-y}dy \\
&= \frac{1}{2\sqrt{\pi}}\left\{\log\mu\cdot\Gamma(\frac{1}{2})\gamma(\frac{1}{2},b^*) + \frac{\sqrt{2\sigma^2}}{\mu}\Gamma(1)\gamma(1,b^*) - \frac{\sigma^2}{\mu^2}\Gamma(\frac{3}{2})\gamma(\frac{3}{2},b^*)\right. \\
&\left. + \frac{(2\sigma^2)^{3/2}}{3\mu^3}\Gamma(2)\gamma(2,b^*) - \frac{2(\sigma^2)^2}{3\mu^3 b}\Gamma(\frac{5}{2})\gamma(\frac{5}{2},b^*)\right\} \qquad \text{(B.1.6)}
\end{aligned}$$

と与えられ, 逆に J_2の下界は

$$\begin{aligned}
J_2 & \\
&\geq \frac{1}{2\sqrt{\pi}}\int_0^{b^*}\left\{\log\mu + \frac{\sqrt{2\sigma^2}}{\mu}y^{1/2} - \frac{\sigma^2}{\mu^2}y + \frac{(2\sigma^2)^{3/2}}{3\mu^3}y^{3/2}\right. \\
&\qquad\left. - \frac{(2\sigma^2)^2}{3\mu^3(\mu+\sqrt{2\sigma^2 y})}y^2\right\} y^{-1/2}e^{-y}dy \\
&\geq \frac{1}{2\sqrt{\pi}}\int_0^{b^*}\left\{\log\mu\cdot y^{-1/2} + \frac{\sqrt{2\sigma^2}}{\mu} - \frac{\sigma^2}{\mu^2}y^{1/2} + \frac{(2\sigma^2)^{3/2}}{3\mu^2 b}y - \frac{4(\sigma^2)^2}{3\mu^4}y^{3/2}\right\} e^{-y}dy \\
&= \frac{1}{2\sqrt{\pi}}\left\{\log\mu\cdot\Gamma(\frac{1}{2})\gamma(\frac{1}{2},b^*) + \frac{\sqrt{2\sigma^2}}{\mu}\Gamma(1)\gamma(1,b^*) - \frac{\sigma^2}{\mu^2}\Gamma(\frac{3}{2})\gamma(\frac{3}{2},b^*)\right. \\
&\left. + \frac{(2\sigma^2)^{3/2}}{3\mu^3}\Gamma(2)\gamma(2,b^*) - \frac{4(\sigma^2)^2}{3\mu^4}\Gamma(\frac{5}{2})\gamma(\frac{5}{2},b^*)\right\} \qquad \text{(B.1.7)}
\end{aligned}$$

と与えられる. J_1と J_2は, $J = J_1 + J_2$となるように定義されたもので, Jの上下界は式(B.1.4), (B.1.5), (B.1.6), (B.1.7) の上下の評価を用いると次のように表される:

$$\begin{aligned}
J \leq \frac{1}{2\sqrt{\pi}}&\left\{\log\mu\cdot\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right) - \frac{\sqrt{2\sigma^2}}{\mu}\Gamma(1)\left(\gamma(1,a^*)-\gamma(1,b^*)\right)\right. \\
&- \frac{\sigma^2}{\mu^2}\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right) - \frac{(2\sigma^2)^{3/2}}{3\mu^3}\Gamma(2)\left(\gamma(2,a^*)-\gamma(2,b^*)\right)
\end{aligned}$$

$$+\frac{2(\sigma^2)^2}{3\mu^4}\Gamma(\frac{5}{2})\left(\frac{\mu^2}{a^2}\gamma(\frac{5}{2},a^*)-\frac{\mu}{b}\gamma(\frac{5}{2},b^*)\right)\Big\}$$

$$J \geq \frac{1}{2\sqrt{\pi}}\Big\{\log\mu\cdot\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)-\frac{\sqrt{2\sigma^2}}{\mu}\Gamma(1)\left(\gamma(1,a^*)-\gamma(1,b^*)\right)$$
$$-\frac{\sigma^2}{\mu^2}\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)-\frac{(2\sigma^2)^{3/2}}{3\mu^3}\Gamma(2)\left(\gamma(2,a^*)-\gamma(2,b^*)\right)$$
$$-\frac{4(\sigma^2)^2}{3\mu^3}\Gamma(\frac{5}{2})\left(\frac{1}{a}\gamma(\frac{5}{2},a^*)+\frac{1}{\mu}\gamma(\frac{5}{2},b^*)\right)\Big\}.$$

これにより補題が証明された. □

系 B.1.1 a, bを$0<a\leq\mu\leq b<1$を満たす数とし,

$$J'=\int_A\log(1-x)f_N(x)dx. \tag{B.1.8}$$

を考える. この時J'の上下界は次のように与えられる:

$$J' \leq \frac{1}{2\sqrt{\pi}}\Big\{\log(1-\mu)\cdot\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)-\frac{\sqrt{2\sigma^2}}{1-\mu}\Gamma(1)\left(\gamma(1,b^*)-\gamma(1,a^*)\right)$$
$$-\frac{\sigma^2}{(1-\mu)^2}\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)-\frac{(2\sigma^2)^{3/2}}{3(1-\mu)^3}\Gamma(2)\left(\gamma(2,b^*)-\gamma(2,a^*)\right)$$
$$+\frac{2(\sigma^2)^2}{3(1-\mu)^4}\Gamma(\frac{5}{2})\left(\frac{(1-\mu)^2}{(1-b)^2}\gamma(\frac{5}{2},b^*)-\frac{1-\mu}{1-a}\gamma(\frac{5}{2},a^*)\right)\Big\}$$

$$J' \geq \frac{1}{2\sqrt{\pi}}\Big\{\log(1-\mu)\cdot\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)-\frac{\sqrt{2\sigma^2}}{1-\mu}\Gamma(1)\left(\gamma(1,b^*)-\gamma(1,a^*)\right)$$
$$-\frac{\sigma^2}{(1-\mu)^2}\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)-\frac{(2\sigma^2)^{3/2}}{3(1-\mu)^3}\Gamma(2)\left(\gamma(2,b^*)-\gamma(2,a^*)\right)$$
$$-\frac{4(\sigma^2)^2}{3(1-\mu)^4}\Gamma(\frac{5}{2})\left(\frac{1-\mu}{1-b}\gamma(\frac{5}{2},b^*)+\gamma(\frac{5}{2},a^*)\right)\Big\}.$$

証明. $x=1-z$という変換を考えると, J'は

$$\begin{aligned} J' &= \int_a^b \log(1-x)f_N(x)dx \\ &= \int_{1-b}^{1-a}\log z\, f_N(1-z)dz \end{aligned}$$

と表される。ここで$f_N(1-z)$は平均, 分散がそれぞれ$1-\mu$, σ^2の正規分布の密度関数である. これと補題 B.1.2によりJ'の上下界が得られる. これにより系 B.1.1が証明される. □

B.2 ガンマ分布の正規近似

$f_g(x)$ を

$$f_g(x) = \frac{1}{\Gamma(\alpha)} \beta^\alpha x^{\alpha-1} e^{-\beta x}$$

で定義される一般のガンマ分布の密度関数とする．ここで$\alpha > 0, \beta > 0$とする．本節では B.1節で与えた諸量の計算を基にして修正 K-L 情報量に基づいたガンマ分布の正規近似の一般形について導出する．一般形の形で導出することによりこれらを含む問題が統合的に見通せる可能性がある．今, 近似される平均μ, 分散σ^2の正規分布の密度関数を

$$f_N(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

とする．$I^*(f_N, f_g; A)$ を修正 K-L 情報量とすると

$$I^*(f_N, f_g; A) = \int_A \log \frac{f_N(x)}{f_g(x)} f_N(x) dx.$$

と表される．この時, 修正 K-L 情報量 $I^*(f_N, f_g; A)$ に対して次の定理が成り立つ．

定理 B.2.1 ガンマ分布の正規近似に関する修正 K-L 情報量 $I^*(f_N, f_g; A)$ の上下界の評価は次のように表される:

$$\begin{aligned}
I^*(f_N, f_g; A) \le \\
&\frac{1}{2\sqrt{\pi}} \left\{ \left[\log \frac{\Gamma(\alpha)}{\sqrt{2\pi\sigma^2}\beta^\alpha} + \beta\mu - (\alpha-1)\log\mu \right] \Gamma(\frac{1}{2}) \left(\gamma(\frac{1}{2}, a^*) + \gamma(\frac{1}{2}, b^*) \right) \right. \\
&- \left[\beta\sqrt{2\sigma^2} - \frac{(\alpha-1)\sqrt{2\sigma^2}}{\mu} \right] \Gamma(1) \Big(\gamma(1, a^*) - \gamma(1, b^*) \Big) \\
&- \left(1 - \frac{(\alpha-1)\sigma^2}{\mu^2} \right) \Gamma(\frac{3}{2}) \left(\gamma(\frac{3}{2}, a^*) + \gamma(\frac{3}{2}, b^*) \right) \\
&+ \frac{(2\sigma^2)^{3/2}(\alpha-1)}{3\mu^3} \Gamma(2) \Big(\gamma(2, a^*) - \gamma(2, b^*) \Big) \\
&\left. + \frac{4(\sigma^2)^2(\alpha-1)}{3\mu^4} \Gamma(\frac{5}{2}) \left(\frac{\mu}{a} \gamma(\frac{5}{2}, a^*) + \gamma(\frac{5}{2}, b^*) \right) \right\}
\end{aligned}$$

$$I^*(f_N, f_g; A) \ge$$

$$
\begin{aligned}
&\frac{1}{2\sqrt{\pi}}\left\{\left[\log\frac{\Gamma(\alpha)}{\sqrt{2\pi\sigma^2}\beta^\alpha}+\beta\mu-(\alpha-1)\log\mu\right]\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)\right.\\
&-\left[\beta\sqrt{2\sigma^2}-\frac{(\alpha-1)\sqrt{2\sigma^2}}{\mu}\right]\Gamma(1)\Big(\gamma(1,a^*)-\gamma(1,b^*)\Big)\\
&-\left(1-\frac{(\alpha-1)\sigma^2}{\mu^2}\right)\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)\\
&+\frac{(2\sigma^2)^{3/2}(\alpha-1)}{3\mu^3}\Gamma(2)\Big(\gamma(2,a^*)-\gamma(2,b^*)\Big)\\
&\left.-\frac{2(\sigma^2)^2(\alpha-1)}{3\mu^4}\Gamma(\frac{5}{2})\left(\frac{\mu^2}{a^2}\gamma(\frac{5}{2},a^*)-\frac{\mu}{b}\gamma(\frac{5}{2},b^*)\right)\right\}.
\end{aligned}
$$

ここで $a^*=\dfrac{(a-\mu)^2}{2\sigma^2}$, $b^*=\dfrac{(b-\mu)^2}{2\sigma^2}$とする.

証明. $I^*(f_N,f_g;A)$ は

$$
\begin{aligned}
I^*(f_N,f_g;A) &= \int_A\log\frac{f_N(x)}{f_g(x)}f_N(x)dx\\
&= \int_A\left\{\log\frac{\Gamma(\alpha)}{\sqrt{2\pi\sigma^2}\beta^\alpha}+\beta x-\frac{(x-\mu)^2}{2\sigma^2}-(\alpha-1)\log x\right\}\\
&= \log\frac{\Gamma(\alpha)}{\sqrt{2\pi\sigma^2}\beta^\alpha}Pr(x\in A)+\beta\mu^*-\frac{\sigma^{*2}}{2\sigma^2}-(\alpha-1)J
\end{aligned}
$$

と書ける. ここでμ^*, σ^{*2}と Jはそれぞれ (B.1.1), (B.1.2), (B.1.3) で定義されたものとする. 補題 B.1.1と補題 B.1.2では, μ^*, σ^{*2} と Jの上下界をそれぞれ不完全ガンマ関数を用いて評価した. 補題 B.1.1と補題 B.1.2での結果を用いると, 修正 K-L 情報量の上界は

$$
\begin{aligned}
&I^*(f_N,f_g;A)=\log\frac{\Gamma(\alpha)}{\sqrt{2\pi\sigma^2}\beta^\alpha}Pr(x\in A)+\beta\mu^*-\frac{\sigma^{*2}}{2\sigma^2}-(\alpha-1)J\\
&\le\frac{1}{2\sqrt{\pi}}\left\{\log\frac{\Gamma(\alpha)}{\sqrt{2\pi\sigma^2}\beta^\alpha}\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)\right.\\
&+\beta\left[\mu\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)-\sqrt{2\sigma^2}\Big(\gamma(1,a^*)-\gamma(1,b^*)\Big)\right]-\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)\\
&-(\alpha-1)\left[\log\mu\cdot\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)-\frac{\sqrt{2\sigma^2}}{\mu}\Gamma(1)\Big(\gamma(1,a^*)-\gamma(1,b^*)\Big)\right.\\
&-\frac{\sigma^2}{\mu^2}\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)-\frac{(2\sigma^2)^{3/2}}{3\mu^3}\Gamma(2)\Big(\gamma(2,a^*)-\gamma(2,b^*)\Big)\\
&\left.\left.-\frac{4(\sigma^2)^2}{3\mu^4}\Gamma(\frac{5}{2})\left(\frac{\mu}{a}\gamma(\frac{5}{2},a^*)+\gamma(\frac{5}{2},b^*)\right)\right]\right\}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{2\sqrt{\pi}} \left\{ \left[\log \frac{\Gamma(\alpha)}{\sqrt{2\pi\sigma^2}\beta^\alpha} + \beta\mu - (\alpha-1)\log\mu \right] \Gamma(\tfrac{1}{2}) \left(\gamma(\tfrac{1}{2}, a^*) + \gamma(\tfrac{1}{2}, b^*) \right) \right. \\
&- \left[\beta\sqrt{2\sigma^2} - \frac{(\alpha-1)\sqrt{2\sigma^2}}{\mu} \right] \Gamma(1) \Big(\gamma(1, a^*) - \gamma(1, b^*) \Big) \\
&- \left(1 - \frac{\sigma^2(\alpha-1)}{\mu^2} \right) \Gamma(\tfrac{3}{2}) \left(\gamma(\tfrac{3}{2}, a^*) + \gamma(\tfrac{3}{2}, b^*) \right) \\
&+ \frac{(2\sigma^2)^{3/2}(\alpha-1)}{3\mu^3} \Gamma(2) \Big(\gamma(2, a^*) - \gamma(2, b^*) \Big) \\
&\left. + \frac{4(\sigma^2)^2(\alpha-1)}{3\mu^4} \Gamma(\tfrac{5}{2}) \left(\frac{\mu}{a} \gamma(\tfrac{5}{2}, a^*) + \gamma(\tfrac{5}{2}, b^*) \right) \right\}
\end{aligned}
$$

と与えられる．また逆に修正 K-L 情報量の下界は

$$
\begin{aligned}
I^*(f_N, f_g; A) &\geq \frac{1}{2\sqrt{\pi}} \left\{ \log \frac{\Gamma(\alpha)}{\sqrt{2\pi\sigma^2}\beta^\alpha} \Gamma(\tfrac{1}{2}) \left(\gamma(\tfrac{1}{2}, a^*) + \gamma(\tfrac{1}{2}, b^*) \right) \right. \\
&+ \beta \left[\mu\Gamma(\tfrac{1}{2}) \left(\gamma(\tfrac{1}{2}, a^*) + \gamma(\tfrac{1}{2}, b^*) \right) - \sqrt{2\sigma^2} \Big(\gamma(1, a^*) - \gamma(1, b^*) \Big) \right] - \left(\gamma(\tfrac{3}{2}, a^*) + \gamma(\tfrac{3}{2}, b^*) \right) \\
&- (\alpha-1) \left[\log\mu \cdot \Gamma(\tfrac{1}{2}) \left(\gamma(\tfrac{1}{2}, a^*) + \gamma(\tfrac{1}{2}, b^*) \right) - \frac{\sqrt{2\sigma^2}}{\mu} \Gamma(1) \Big(\gamma(1, a^*) - \gamma(1, b^*) \Big) \right. \\
&- \frac{\sigma^2}{\mu^2} \Gamma(\tfrac{3}{2}) \left(\gamma(\tfrac{3}{2}, a^*) + \gamma(\tfrac{3}{2}, b^*) \right) - \frac{(2\sigma^2)^{3/2}}{3\mu^3} \Gamma(2) \Big(\gamma(2, a^*) - \gamma(2, b^*) \Big) \\
&\left.\left. + \frac{2(\sigma^2)^2}{3\mu^4} \Gamma(\tfrac{5}{2}) \left(\frac{\mu^2}{a^2} \gamma(\tfrac{5}{2}, a^*) - \frac{\mu}{b} \gamma(\tfrac{5}{2}, b^*) \right) \right] \right\} \\
&= \frac{1}{2\sqrt{\pi}} \left\{ \left[\log \frac{\Gamma(\alpha)}{\sqrt{2\pi\sigma^2}\beta^\alpha} + \beta\mu - (\alpha-1)\log\mu \right] \Gamma(\tfrac{1}{2}) \left(\gamma(\tfrac{1}{2}, a^*) + \gamma(\tfrac{1}{2}, b^*) \right) \right. \\
&- \left[\beta\sqrt{2\sigma^2} - \frac{(\alpha-1)\sqrt{2\sigma^2}}{\mu} \right] \Gamma(1) \Big(\gamma(1, a^*) - \gamma(1, b^*) \Big) \\
&- \left(1 - \frac{(\alpha-1)\sigma^2}{\mu^2} \right) \Gamma(\tfrac{3}{2}) \left(\gamma(\tfrac{3}{2}, a^*) + \gamma(\tfrac{3}{2}, b^*) \right) \\
&+ \frac{(2\sigma^2)^{3/2}(\alpha-1)}{3\mu^3} \Gamma(2) \Big(\gamma(2, a^*) - \gamma(2, b^*) \Big) \\
&\left. - \frac{2(\sigma^2)^2(\alpha-1)}{3\mu^4} \Gamma(\tfrac{5}{2}) \left(\frac{\mu^2}{a^2} \gamma(\tfrac{5}{2}, a^*) - \frac{\mu}{b} \gamma(\tfrac{5}{2}, b^*) \right) \right\}
\end{aligned}
$$

と与えられる．これにより定理 B.2.1が証明された． □

定理 B.2.1は, 任意の平均, 分散を持つ正規分布に対する近似を修正 K-L 情報量を用いて評価している．近似分布の平均, 分散は, 様々な取り方があるが, 本論文では慣例に従い, 基

の分布と揃えた平均, 分散を持つ分布への近似を考える. この時, 次の結果を得る.

系 B.2.1 正規分布の平均と分散を$\mu = \dfrac{\alpha}{\beta}$, $\sigma^2 = \dfrac{\alpha}{\beta^2}$ととする. この時の修正 K-L 情報量 $I^*(f_N, f_g; A)$ の上下界はそれぞれ

$$
\begin{aligned}
&I^*(f_N, f_g; A) \le \frac{1}{2}\left[\frac{1}{12\alpha} - \underline{R}(\alpha)\right]\left(\gamma(\frac{1}{2}, a^*) + \gamma(\frac{1}{2}, b^*)\right) \\
&-\frac{1}{\sqrt{2\pi\alpha}}\left(\gamma(1, a^*) - \gamma(1, b^*)\right) - \frac{1}{4\alpha}\left(\gamma(\frac{3}{2}, a^*) + \gamma(\frac{3}{2}, b^*)\right) \\
&+\frac{\sqrt{2}(\alpha-1)}{3\alpha^{3/2}\sqrt{\pi}}\left(\gamma(2, a^*) - \gamma(2, b^*)\right) + \frac{\alpha-1}{2\alpha^2}\left(\frac{\alpha}{\beta a}\gamma(\frac{5}{2}, a^*) + \gamma(\frac{5}{2}, b^*)\right) \\
&I^*(f_N, f_g; A) \ge \frac{1}{2}\left[\frac{1}{12\alpha} - \overline{R}(\alpha)\right]\left(\gamma(\frac{1}{2}, a^*) + \gamma(\frac{1}{2}, b^*)\right) \\
&-\frac{1}{\sqrt{2\pi\alpha}}\left(\gamma(1, a^*) - \gamma(1, b^*)\right) - \frac{1}{4\alpha}\left(\gamma(\frac{3}{2}, a^*) + \gamma(\frac{3}{2}, b^*)\right) \\
&+\frac{\sqrt{2}(\alpha-1)}{3\alpha^{3/2}\sqrt{\pi}}\left(\gamma(2, a^*) - \gamma(2, b^*)\right) - \frac{\alpha-1}{4\alpha^2}\left(\frac{\alpha^2}{\beta^2 a^2}\gamma(\frac{5}{2}, a^*) - \frac{\alpha}{\beta b}\gamma(\frac{5}{2}, b^*)\right)
\end{aligned}
$$

と与えられる. ここで $a^* = \dfrac{(\alpha - \beta a)^2}{2\alpha}$, $b^* = \dfrac{(\beta b - \alpha)^2}{2\alpha}$とし, $\underline{R}(s)$, $\overline{R}(s)$ は

$$
\begin{aligned}
\underline{R}(s) &= \frac{1}{360(s-1)(s+1)} - \frac{1}{120s^2(s-1)(s+1)} \\
\overline{R}(s) &= \frac{1}{360(s-1)(s+1)} + \frac{11}{480s^2(s-1)(s+1)}.
\end{aligned}
$$

とする.

証明. 定理 B.2.1の結果に$\mu = \dfrac{\alpha}{\beta}$, $\sigma^2 = \dfrac{\alpha}{\beta^2}$ と置き, さらにガンマ関数の不等式評価として,

$$
\begin{aligned}
\log\Gamma(s) &\le \log\sqrt{2\pi} + \left(s - \frac{1}{2}\right)\log s - s + \frac{1}{12s} - \underline{R}(s) \\
\log\Gamma(s) &\ge \log\sqrt{2\pi} + \left(s - \frac{1}{2}\right)\log s - s + \frac{1}{12s} - \overline{R}(s)
\end{aligned}
$$

を用いると, 系 B.2.1が証明される. □

最後にガンマ分布の特別な場合として, χ^2分布の正規近似に関する結果を与える. χ^2分布の正規近似はよく知られていて, いろいろな応用分野で用いられている. 次の結果は, この近似を修正 K-L 情報量を用いた評価を与えているが, 今までの近似理論とは別の角度からこの近似を捉えているので, 大変興味のある結果である.

定理 B.2.2 $\alpha = \frac{n}{2}$, $\beta = \frac{1}{2}$とすると, f_gは自由度 n のχ^2分布に従う. 自由度 n のχ^2分布をf_{χ^2}とする. この時, χ^2分布の正規近似に関する修正 K-L 情報量 $I^*(f_N, f_{\chi^2}; A)$ の上下界は

$$
\begin{aligned}
I^*(f_N, f_{\chi^2}; A) \leq \frac{1}{2\sqrt{\pi}} \Bigg\{ & \left[\log \frac{\Gamma(\frac{n}{2})2^{n/2}}{\sqrt{2\pi\sigma^2}} + \frac{\mu}{2} - \frac{n-2}{2}\log\mu\right] \Gamma(\frac{1}{2}) \left(\gamma(\frac{1}{2}, a^*) + \gamma(\frac{1}{2}, b^*)\right) \\
& - \left[\sqrt{\frac{\sigma^2}{2}} - \frac{(n-2)\sqrt{2\sigma^2}}{2\mu}\right] \Gamma(1)\Big(\gamma(1, a^*) - \gamma(1, b^*)\Big) \\
& - \left(1 - \frac{(n-2)\sigma^2}{2\mu^2}\right) \Gamma(\frac{3}{2}) \left(\gamma(\frac{3}{2}, a^*) + \gamma(\frac{3}{2}, b^*)\right) \\
& + \frac{\sqrt{2}(\sigma^2)^{3/2}(n-2)}{3\mu^3} \Gamma(2)\Big(\gamma(2, a^*) - \gamma(2, b^*)\Big) \\
& + \frac{2(\sigma^2)^2(n-2)}{3\mu^4} \Gamma(\frac{5}{2}) \left(\frac{\mu}{a}\gamma(\frac{5}{2}, a^*) + \gamma(\frac{5}{2}, b^*)\right) \Bigg\} \\
I^*(f_N, f_{\chi^2}; A) \geq \frac{1}{2\sqrt{\pi}} \Bigg\{ & \left[\log \frac{\Gamma(\frac{n}{2})2^{n/2}}{\sqrt{2\pi\sigma^2}} + \frac{\mu}{2} - \frac{n-2}{2}\log\mu\right] \Gamma(\frac{1}{2}) \left(\gamma(\frac{1}{2}, a^*) + \gamma(\frac{1}{2}, b^*)\right) \\
& - \left[\sqrt{\frac{\sigma^2}{2}} - \frac{(n-2)\sqrt{2\sigma^2}}{2\mu}\right] \Gamma(1)\Big(\gamma(1, a^*) - \gamma(1, b^*)\Big) \\
& - \left(1 - \frac{(n-2)\sigma^2}{2\mu^2}\right) \Gamma(\frac{3}{2}) \left(\gamma(\frac{3}{2}, a^*) + \gamma(\frac{3}{2}, b^*)\right) \\
& + \frac{\sqrt{2}(\sigma^2)^{3/2}(n-2)}{3\mu^3} \Gamma(2)\Big(\gamma(2, a^*) - \gamma(2, b^*)\Big) \\
& - \frac{(\sigma^2)^2(n-2)}{3\mu^4} \Gamma(\frac{5}{2}) \left(\frac{\mu^2}{a^2}\gamma(\frac{5}{2}, a^*) - \frac{\mu}{b}\gamma(\frac{5}{2}, b^*)\right) \Bigg\}
\end{aligned}
$$

と与えられる.

系 B.2.2 χ^2分布の正規近似に関して平均, 分散をχ^2分布にそろえた正規分布, つまり $\mu = n$, $\sigma^2 = 2n$ を考える時, 修正 K-L 情報量 $I^*(f_N, f_{\chi^2}; A)$ の上下界は

$$
\begin{aligned}
I^*(f_N, f_{\chi^2}; A) \leq & \frac{1}{2}\left[\frac{1}{6n} - \underline{R}(\frac{n}{2})\right] \left(\gamma(\frac{1}{2}, a^*) + \gamma(\frac{1}{2}, b^*)\right) \\
& - \frac{1}{\sqrt{\pi n}}\Big(\gamma(1, a^*) - \gamma(1, b^*)\Big) - \frac{1}{2n}\left(\gamma(\frac{3}{2}, a^*) + \gamma(\frac{3}{2}, b^*)\right) \\
& + \frac{2(n-2)}{3n^{3/2}\sqrt{\pi}}\Big(\gamma(2, a^*) - \gamma(2, b^*)\Big) + \frac{n-2}{2n^2}\left(\frac{n}{a}\gamma(\frac{5}{2}, a^*) + \gamma(\frac{5}{2}, b^*)\right) \\
I^*(f_N, f_{\chi^2}; A) \geq & \frac{1}{2}\left[\frac{1}{6n} - \overline{R}(\frac{n}{2})\right] \left(\gamma(\frac{1}{2}, a^*) + \gamma(\frac{1}{2}, b^*)\right) \\
& - \frac{1}{\sqrt{\pi n}}\Big(\gamma(1, a^*) - \gamma(1, b^*)\Big) - \frac{1}{2n}\left(\gamma(\frac{3}{2}, a^*) + \gamma(\frac{3}{2}, b^*)\right)
\end{aligned}
$$

$$+\frac{2(n-2)}{3n^{3/2}\sqrt{\pi}}\Big(\gamma(2,a^*)-\gamma(2,b^*)\Big)-\frac{n-2}{2n^2}\left(\frac{n^2}{a^2}\gamma(\frac{5}{2},a^*)-\frac{\mu}{b}\gamma(\frac{5}{2},b^*)\right)$$

と与えられる.

B.3 ベータ分布の正規近似

3.3.1節の例 2 で与えているベータ分布の正規近似についての証明をまとめる. f_β, f_Nを

$$f_\beta(x)=\frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}$$
$$f_N(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

で定義されるパラメータ (α,β) のベータ分布, パラメータ (μ,σ^2) の正規分布とする.

定理 B.3.1 ベータ分布の正規近似に関する修正 K-L 情報量 $I^*(f_N,f_\beta;A)$ の上下界の評価は次のように与えられる:

$$I^*(f_N,f_\beta;A)\le\frac{1}{2\sqrt{\pi}}\left\{\left[\log\frac{B(\alpha,\beta)}{\sqrt{2\pi\sigma^2}}-(\alpha-1)\log\mu-(\beta-1)\log(1-\mu)\right]\Gamma(\frac{1}{2})\right.$$
$$\Big(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\Big)+\sqrt{2\sigma^2}\left(\frac{\alpha-1}{\mu}-\frac{\beta-1}{1-\mu}\right)\Gamma(1)\Big(\gamma(1,a^*)-\gamma(1,b^*)\Big)$$
$$-\left(1-\frac{(\alpha-1)\sigma^2}{\mu^2}-\frac{(\beta-1)\sigma^2}{(1-\mu)^2}\right)\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)$$
$$+\frac{(2\sigma^2)^{3/2}}{3}\left[\frac{\alpha-1}{\mu^3}-\frac{\beta-1}{(1-\mu)^3}\right]\Gamma(2)\Big(\gamma(2,a^*)-\gamma(2,b^*)\Big)$$
$$\left.+\frac{4(\sigma^2)^2}{3}\Gamma(\frac{5}{2})\left[\left(\frac{\beta-1}{(1-\mu)^4}+\frac{\alpha-1}{\mu^3a}\right)\gamma(\frac{5}{2},a^*)+\left(\frac{\beta-1}{(1-\mu)^3(1-b)}+\frac{\alpha-1}{\mu^4}\right)\gamma(\frac{5}{2},b^*)\right]\right\}$$
$$I^*(f_N,f_\beta;A)\ge\frac{1}{2\sqrt{\pi}}\left\{\left[\log\frac{B(\alpha,\beta)}{\sqrt{2\pi\sigma^2}}-(\alpha-1)\log\mu-(\beta-1)\log(1-\mu)\right]\Gamma(\frac{1}{2})\right.$$
$$\Big(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\Big)+\sqrt{2\sigma^2}\left(\frac{\alpha-1}{\mu}-\frac{\beta-1}{1-\mu}\right)\Gamma(1)\Big(\gamma(1,a^*)-\gamma(1,b^*)\Big)$$
$$-\left(1-\frac{(\alpha-1)\sigma^2}{\mu^2}-\frac{(\beta-1)\sigma^2}{(1-\mu)^2}\right)\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)$$
$$+\frac{(2\sigma^2)^{3/2}}{3}\left[\frac{\alpha-1}{\mu^3}-\frac{\beta-1}{(1-\mu)^3}\right]\Gamma(2)\Big(\gamma(2,a^*)-\gamma(2,b^*)\Big)$$
$$-\frac{2(\sigma^2)^2}{3}\Gamma(\frac{5}{2})\left[\left(-\frac{\beta-1}{(1-\mu)^3(1-a)}+\frac{\alpha-1}{\mu^2a^2}\right)\gamma(\frac{5}{2},a^*)\right.$$

$$+\left(\frac{\beta-1}{(1-\mu)^2(1-b)^2}-\frac{\alpha-1}{\mu^3 b}\right)\gamma(\frac{5}{2},b^*)\Bigg]\Bigg\}.$$

ここで $a^*=\dfrac{(a-\mu)^2}{2\sigma^2}$, $b^*=\dfrac{(b-\mu)^2}{2\sigma^2}$ とする.

証明. $I^*(f_N,f_\beta;A)$ は

$$\begin{aligned}
I^*(f_N,f_\beta;A)&=\int_A \log\frac{f_N(x)}{f_\beta(x)}f_N(x)dx\\
&=\int_A\left\{\log\frac{B(\alpha,\beta)}{\sqrt{2\pi\sigma^2}}-\frac{(x-\mu)^2}{2\sigma^2}-(\alpha-1)\log x-(\beta-1)\log(1-x)\right\}f_N(x)dx\\
&=\log\frac{B(\alpha,\beta)}{\sqrt{2\pi\sigma^2}}Pr(x\in A)-\frac{\sigma^{*2}}{2\sigma^2}-(\alpha-1)J-(\beta-1)J'
\end{aligned}$$

と書くことができる. ここでσ^{*2}, J, J'はそれぞれ式 (B.1.2), (B.1.3), (B.1.8) で定義されたものとする. 補題 B.1.1, 補題 B.1.2と系 B.1.1では, σ^{*2}と J, J'の上下界を不完全ガンマ関数を用いて評価した. 補題 B.1.1, 補題 B.1.2と系 B.1.1での結果を用いると修正 K-L 情報量の上界は

$$\begin{aligned}
I^*(f_N,f_\beta;A)&\le \log\frac{B(\alpha,\beta)}{\sqrt{2\pi\sigma^2}}\cdot\frac{1}{2\sqrt{\pi}}\left\{\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)\right\}\\
&-\frac{\Gamma(\frac{3}{2})}{2\sqrt{\pi}}\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)\\
&-\frac{\alpha-1}{2\sqrt{\pi}}\Bigg\{\log\mu\cdot\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)-\frac{\sqrt{2\sigma^2}}{\mu}\Gamma(1)\Big(\gamma(1,a^*)-\gamma(1,b^*)\Big)\\
&-\frac{\sigma^2}{\mu^2}\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)-\frac{(2\sigma^2)^{3/2}}{3\mu^3}\Gamma(2)\Big(\gamma(2,a^*)-\gamma(2,b^*)\Big)\\
&-\frac{4(\sigma^2)^2}{3\mu^4}\Gamma(\frac{5}{2})\left(\frac{\mu}{a}\gamma(\frac{5}{2},a^*)+\gamma(\frac{5}{2},b^*)\right)\Bigg\}\\
&-\frac{\beta-1}{2\sqrt{\pi}}\Bigg\{\log(1-\mu)\cdot\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)-\frac{\sqrt{2\sigma^2}}{1-\mu}\Gamma(1)\Big(\gamma(1,b^*)-\gamma(1,a^*)\Big)\\
&-\frac{\sigma^2}{(1-\mu)^2}\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)-\frac{(2\sigma^2)^{3/2}}{3(1-\mu)^3}\Gamma(2)\Big(\gamma(2,b^*)-\gamma(2,a^*)\Big)\\
&-\frac{4(\sigma^2)^2}{3(1-\mu)^4}\Gamma(\frac{5}{2})\left(\frac{1-\mu}{1-b}\gamma(\frac{5}{2},b^*)+\gamma(\frac{5}{2},a^*)\right)\Bigg\}\\
&=\frac{1}{2\sqrt{\pi}}\Bigg\{\left[\log\frac{B(\alpha,\beta)}{\sqrt{2\pi\sigma^2}}-(\alpha-1)\log\mu-(\beta-1)\log(1-\mu)\right]\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)\\
&+\sqrt{2\sigma^2}\left(\frac{\alpha-1}{\mu}-\frac{\beta-1}{1-\mu}\right)\Gamma(1)\Big(\gamma(1,a^*)-\gamma(1,b^*)\Big)
\end{aligned}$$

$$
\begin{aligned}
&-\left(1-\frac{(\alpha-1)\sigma^2}{\mu^2}-\frac{(\beta-1)\sigma^2}{(1-\mu)^2}\right)\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)\\
&+\frac{(2\sigma^2)^{3/2}}{3}\left[\frac{\alpha-1}{\mu^3}-\frac{\beta-1}{(1-\mu)^3}\right]\Gamma(2)\Big(\gamma(2,a^*)-\gamma(2,b^*)\Big)\\
&+\frac{4(\sigma^2)^2}{3}\Gamma(\frac{5}{2})\left[\left(\frac{\beta-1}{(1-\mu)^4}+\frac{\alpha-1}{\mu^3 a}\right)\gamma(\frac{5}{2},a^*)\right.\\
&\left.+\left(\frac{\beta-1}{(1-\mu)^3(1-b)}+\frac{\alpha-1}{\mu^4}\right)\gamma(\frac{5}{2},b^*)\right]\Bigg\}
\end{aligned}
$$

と与えられる．また逆に修正 K-L 情報量の下界は

$$
\begin{aligned}
&I^*(f_N,f_\beta;A)\geq\log\frac{B(\alpha,\beta)}{\sqrt{2\pi\sigma^2}}\cdot\frac{1}{2\sqrt{\pi}}\left\{\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)\right\}\\
&-\frac{\Gamma(\frac{3}{2})}{2\sqrt{\pi}}\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)\\
&-\frac{\alpha-1}{2\sqrt{\pi}}\left\{\log\mu\cdot\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)-\frac{\sqrt{2\sigma^2}}{\mu}\Gamma(1)\Big(\gamma(1,a^*)-\gamma(1,b^*)\Big)\right.\\
&-\frac{\sigma^2}{\mu^2}\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)-\frac{(2\sigma^2)^{3/2}}{3\mu^3}\Gamma(2)\Big(\gamma(2,a^*)-\gamma(2,b^*)\Big)\\
&\left.+\frac{2(\sigma^2)^2}{3\mu^4}\Gamma(\frac{5}{2})\left(\frac{\mu^2}{a^2}\gamma(\frac{5}{2},a^*)-\frac{\mu}{b}\gamma(\frac{5}{2},b^*)\right)\right\}\\
&-\frac{\beta-1}{2\sqrt{\pi}}\left\{\log(1-\mu)\cdot\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)-\frac{\sqrt{2\sigma^2}}{1-\mu}\Gamma(1)\Big(\gamma(1,b^*)-\gamma(1,a^*)\Big)\right.\\
&-\frac{\sigma^2}{(1-\mu)^2}\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)-\frac{(2\sigma^2)^{3/2}}{3(1-\mu)^3}\Gamma(2)\Big(\gamma(2,b^*)-\gamma(2,a^*)\Big)\\
&\left.+\frac{2(\sigma^2)^2}{3(1-\mu)^4}\Gamma(\frac{5}{2})\left(\frac{(1-\mu)^2}{(1-b)^2}\gamma(\frac{5}{2},b^*)-\frac{1-\mu}{1-a}\gamma(\frac{5}{2},a^*)\right)\right\}\\
&=\frac{1}{2\sqrt{\pi}}\left\{\left[\log\frac{B(\alpha,\beta)}{\sqrt{2\pi\sigma^2}}-(\alpha-1)\log\mu-(\beta-1)\log(1-\mu)\right]\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)\right.\\
&+\sqrt{2\sigma^2}\left(\frac{\alpha-1}{\mu}-\frac{\beta-1}{1-\mu}\right)\Gamma(1)\Big(\gamma(1,a^*)-\gamma(1,b^*)\Big)\\
&-\left(1-\frac{(\alpha-1)\sigma^2}{\mu^2}-\frac{(\beta-1)\sigma^2}{(1-\mu)^2}\right)\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)\\
&+\frac{(2\sigma^2)^{3/2}}{3}\left[\frac{\alpha-1}{\mu^3}-\frac{\beta-1}{(1-\mu)^3}\right]\Gamma(2)\Big(\gamma(2,a^*)-\gamma(2,b^*)\Big)\\
&-\frac{2(\sigma^2)^2}{3}\Gamma(\frac{5}{2})\left[\left(-\frac{\beta-1}{(1-\mu)^3(1-a)}+\frac{\alpha-1}{\mu^2a^2}\right)\gamma(\frac{5}{2},a^*)\right.\\
&\left.\left.+\left(\frac{\beta-1}{(1-\mu)^2(1-b)^2}-\frac{\alpha-1}{\mu^3 b}\right)\gamma(\frac{5}{2},b^*)\right]\right\}
\end{aligned}
$$

と与えられる．これにより定理 B.3.1が証明された． □

ベータ変数の平均，分散はそれぞれ $E[x] = \dfrac{\alpha}{\alpha+\beta}$, $Var[x] = \dfrac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$ であることが知られている．今，$\alpha+\beta=n$, $\alpha=k_1 n$, $\beta=k_2 n$ とし，ガンマ分布における正規近似の例題と同様に正規分布の平均，分散をベータ変数にそろえたものを考えた時，次の結果が得られる．

系 B.3.1 $\mu = \dfrac{\alpha}{\alpha+\beta} = k_1$, $\sigma^2 = \dfrac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2} = \dfrac{k_1k_2}{n+1}$ とした時，修正 K-L 情報量 $I^*(f_N, f_\beta; A)$ は

$$\underline{I}^*(f_N, f_\beta; A) \le I^*(f_N, f_\beta; A) \le \overline{I}^*(f_N, f_\beta; A)$$

を満たす．ここで

$$\begin{aligned}
&\overline{I}^*(f_N, f_\beta; A) = \frac{1}{2}\left[\frac{1}{2}\log\frac{n+1}{n} + \frac{1}{12n}\left(\frac{1}{k_1}+\frac{1}{k_2}-1\right) + \overline{R}(n) - \underline{R}(k_1 n) - \underline{R}(k_2 n)\right] \\
&\left(\gamma(\frac{1}{2}, a^*) + \gamma(\frac{1}{2}, b^*)\right) + \frac{1}{\sqrt{n+1}}\cdot\frac{k_1-k_2}{\sqrt{2\pi k_1k_2}}\left(\gamma(1,a^*)-\gamma(1,b^*)\right) \\
&-\frac{1}{n+1}\cdot\frac{1-k_1k_2}{4k_1k_2}\left(\gamma(\frac{3}{2}, a^*) + \gamma(\frac{3}{2}, b^*)\right) \\
&+\left[\frac{n}{(n+1)^{3/2}}\cdot\frac{\sqrt{2}(k_2-k_1)}{\sqrt{\pi k_1k_2}} + \frac{1}{(n+1)^{3/2}}\cdot\frac{\sqrt{2}(k_1^3-k_2^3)}{3\sqrt{\pi}(k_1k_2)^{3/2}}\right]\left(\gamma(2,a^*)-\gamma(2,b^*)\right) \\
&+\frac{n}{2(n+1)^2}\left[\left(\frac{k_1^2}{k_2}+\frac{k_2^2}{a}\right)\gamma(\frac{5}{2}, a^*) + \left(\frac{k_1^2}{1-b}+\frac{k_2^2}{k_1}\right)\gamma(\frac{5}{2}, b^*)\right] \\
&-\frac{1}{2(n+1)^2}\left[\left(\frac{k_1^2}{k_2^2}+\frac{k_2^2}{k_1 a}\right)\gamma(\frac{5}{2}, a^*) + \left(\frac{k_1^2}{k_2(1-b)}+\frac{k_2^2}{k_1^2}\right)\gamma(\frac{5}{2}, b^*)\right] \\
&\underline{I}^*(f_N, f_\beta; A) = \frac{1}{2}\left[\frac{1}{2}\log\frac{n+1}{n} + \frac{1}{12n}\left(\frac{1}{k_1}+\frac{1}{k_2}-1\right) + \underline{R}(n) - \overline{R}(k_1 n) - \overline{R}(k_2 n)\right] \\
&\left(\gamma(\frac{1}{2}, a^*) + \gamma(\frac{1}{2}, b^*)\right) + \frac{1}{\sqrt{n+1}}\cdot\frac{k_1-k_2}{\sqrt{2\pi k_1k_2}}\left(\gamma(1,a^*)-\gamma(1,b^*)\right) \\
&-\frac{1}{n+1}\cdot\frac{1-k_1k_2}{4k_1k_2}\left(\gamma(\frac{3}{2}, a^*) + \gamma(\frac{3}{2}, b^*)\right) \\
&+\left[\frac{n}{(n+1)^{3/2}}\cdot\frac{\sqrt{2}(k_2-k_1)}{\sqrt{\pi k_1k_2}} + \frac{1}{(n+1)^{3/2}}\cdot\frac{\sqrt{2}(k_1^3-k_2^3)}{3\sqrt{\pi}(k_1k_2)^{3/2}}\right]\left(\gamma(2,a^*)-\gamma(2,b^*)\right) \\
&+\frac{n}{2(n+1)^2}\left[\left(-\frac{k_1^2}{1-a}+\frac{k_1k_2^2}{a^2}\right)\gamma(\frac{5}{2}, a^*) + \left(\frac{k_1^2k_2}{(1-b)^2}-\frac{k_2^2}{b}\right)\gamma(\frac{5}{2}, b^*)\right] \\
&-\frac{1}{2(n+1)^2}\left[\left(-\frac{k_1^2}{k_2(1-a)}+\frac{k_2^2}{a^2}\right)\gamma(\frac{5}{2}, a^*) + \left(\frac{k_1^2}{(1-b)^2}-\frac{k_2^2}{k_1 b}\right)\gamma(\frac{5}{2}, b^*)\right]
\end{aligned}$$

とする．また $a^* = \dfrac{(k_1-a)^2}{2k_1k_2}(n+1)$, $b^* = \dfrac{(b-k_1)^2}{k_1k_2}(n+1)$ とし，$\underline{R}(s)$, $\overline{R}(s)$ は

$$\underline{R}(s) = \frac{1}{360(s-1)(s+1)} - \frac{1}{120s^2(s-1)(s+1)}$$
$$\overline{R}(s) = \frac{1}{360(s-1)(s+1)} + \frac{11}{480s^2(s-1)(s+1)}.$$

で与えられたものとする．

証明． 定理 B.3.1で与えた結果に $\mu = k_1$, $\sigma^2 = \dfrac{k_1k_2}{(n+1)}$ を代入し，次の不等式

$$\log\Gamma(s) \le \log\sqrt{2\pi} + \left(s-\frac{1}{2}\right)\log s - s + \frac{1}{12s} - \underline{R}(s)$$
$$\log\Gamma(s) \ge \log\sqrt{2\pi} + \left(s-\frac{1}{2}\right)\log s - s + \frac{1}{12s} - \overline{R}(s),$$

を用いると，修正 K-L 情報量の上界は

$$\begin{aligned}
I^*(f_N, f_\beta; A) \le& \frac{1}{2\sqrt{\pi}}\left\{\left[\log\frac{B(k_1n,k_2n)}{\sqrt{2\pi\frac{k_1k_2}{n+1}}} - (k_1n-1)\log k_1n - (k_2n-1)\log k_2n\right]\right.\\
&\cdot\Gamma(\tfrac{1}{2})\left(\gamma(\tfrac{1}{2},a^*)+\gamma(\tfrac{1}{2},b^*)\right) + \sqrt{\frac{2k_1k_2}{n+1}}\left(\frac{k_1n-1}{k_1n}-\frac{k_2n-1}{k_2n}\right)\Gamma(1)\Big(\gamma(1,a^*)-\gamma(1,b^*)\Big)\\
&-\left[1-\frac{k_1k_2}{n+1}\left(\frac{k_1n-1}{{k_1}^2}+\frac{k_2n-1}{k_2^2}\right)\right]\Gamma(\tfrac{3}{2})\left(\gamma(\tfrac{3}{2},a^*)+\gamma(\tfrac{3}{2},b^*)\right)\\
&+\frac{(2k_1k_2)^{3/2}}{3(n+1)^{3/2}}\left[\frac{k_1n-1}{{k_1}^3}-\frac{k_2n-1}{k_2^3}\right]\Gamma(2)\Big(\gamma(2,a^*)-\gamma(2,b^*)\Big)\\
&+\frac{4(k_1k_2)^2}{3(n+1)^2}\Gamma(\tfrac{5}{2})\left[\left(\frac{k_2n-1}{k_2^4}+\frac{k_1n-1}{{k_1}^3a}\right)\gamma(\tfrac{5}{2},a^*)\right.\\
&\left.\left.+\left(\frac{k_2n-1}{k_2^3(1-b)}+\frac{k_1n-1}{{k_1}^4}\right)\gamma(\tfrac{5}{2},b^*)\right]\right\}\\
\le& \frac{1}{2}\left[\frac{1}{2}\log\frac{n+1}{n}+\frac{1}{12n}\left(\frac{1}{k_1}+\frac{1}{k_2}-1\right)+\overline{R}(n)-\underline{R}(k_1n)-\underline{R}(k_2n)\right]\\
&\left(\gamma(\tfrac{1}{2},a^*)+\gamma(\tfrac{1}{2},b^*)\right)+\frac{k_1-k_2}{\sqrt{2\pi(n+1)k_1k_2}}\Big(\gamma(1,a^*)-\gamma(1,b^*)\Big)\\
&-\frac{{k_1}^2+k_2^2+k_1k_2}{4(n+1)k_1k_2}\left(\gamma(\tfrac{3}{2},a^*)+\gamma(\tfrac{3}{2},b^*)\right)\\
&+\frac{\sqrt{2}((k_1k_2^3-{k_1}^3k_2)n+{k_1}^3-k_2^3)}{3\sqrt{\pi}(n+1)^{3/2}(k_1k_2)^{3/2}}\Big(\gamma(2,a^*)-\gamma(2,b^*)\Big)\\
&+\frac{(k_1k_2)^2}{2(n+1)^2}\left[\left(\frac{k_2n-1}{k_2^4}+\frac{k_1n-1}{{k_1}^3a}\right)\gamma(\tfrac{5}{2},a^*)\right.
\end{aligned}$$

$$+\left(\frac{k_2n-1}{k_2^3(1-b)}+\frac{k_1n-1}{{k_1}^4}\right)\gamma(\frac{5}{2},b^*)\Bigg]$$

と与えられる．また逆に修正 K-L 情報量の下界は

$$I^*(f_N,f_\beta;A)\geq\frac{1}{2\sqrt{\pi}}\left\{\left[\log\frac{B(k_1n,k_2n)}{\sqrt{2\pi\frac{k_1k_2}{n+1}}}-(k_1-1)\log k_1n-(k_2-1)\log k_2n\right]\right.$$
$$\cdot\Gamma(\frac{1}{2})\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)+\sqrt{\frac{2k_1k_2}{n+1}}\left(\frac{k_1n-1}{k_1}-\frac{k_2n-1}{k_2}\right)\Gamma(1)\Big(\gamma(1,a^*)-\gamma(1,b^*)\Big)$$
$$-\left[1-\frac{k_1k_2}{n+1}\left(\frac{k_1n-1}{{k_1}^2}+\frac{k_2n-1}{k_2^2}\right)\right]\Gamma(\frac{3}{2})\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)$$
$$+\frac{(2k_1k_2)^{3/2}}{3(n+1)^{3/2}}\left[\frac{k_1n-1}{{k_1}^3}-\frac{k_2n-1}{k_2^3}\right]\Gamma(2)\Big(\gamma(2,a^*)-\gamma(2,b^*)\Big)$$
$$-\frac{2(k_1k_2)^2}{3(n+1)^2}\Gamma(\frac{5}{2})\left[\left(\frac{k_2n-1}{k_2^4}+\frac{k_1n-1}{{k_1}^2a^2}\right)\gamma(\frac{5}{2},a^*)\right.$$
$$\left.\left.+\left(\frac{k_2n-1}{k_2^2(1-b)^2}+\frac{k_1-n1}{{k_1}^4}\right)\gamma(\frac{5}{2},b^*)\right]\right\}$$
$$\geq\frac{1}{2}\left[\frac{1}{2}\log\frac{n+1}{n}+\frac{1}{12n}\left(\frac{1}{k_1}+\frac{1}{k_2}-1\right)+\underline{R}(n)-\overline{R}(k_1n)-\overline{R}(k_2n)\right]$$
$$\left(\gamma(\frac{1}{2},a^*)+\gamma(\frac{1}{2},b^*)\right)+\frac{k_1-k_2}{\sqrt{2\pi(n+1)k_1k_2}}\Big(\gamma(1,a^*)-\gamma(1,b^*)\Big)$$
$$-\frac{{k_1}^2+k_2^2+k_1k_2}{4(n+1)k_1k_2}\left(\gamma(\frac{3}{2},a^*)+\gamma(\frac{3}{2},b^*)\right)$$
$$+\frac{\sqrt{2}((k_1k_2^3-{k_1}^3k_2)n+{k_1}^3-k_2^3)}{3\sqrt{\pi}(n+1)^{3/2}(k_1k_2)^{3/2}}\Big(\gamma(2,a^*)-\gamma(2,b^*)\Big)$$
$$-\frac{(k_1k_2)^2}{4(n+1)^2}\left[\left(-\frac{k_2n-1}{k_2^3(1-a)}+\frac{k_1n-1}{{k_1}^2a^2}\right)\gamma(\frac{5}{2},a^*)\right.$$
$$\left.+\left(\frac{k_2n-1}{k_2^2(1-b)^2}-\frac{k_1n-1}{{k_1}^3b}\right)\gamma(\frac{5}{2},b^*)\right]$$

と与えられる．これにより系 B.3.1が証明された． □

B.4 ordered Dirichlet 密度関数の近似展開

本節では (3.3.13) 式で定義した ordered Dirichlet 分布の密度関数 $h_n(z_{(k)})$ の展開を考える．これは Mosteller (1946) により大まかな極限理論での評価が行なわれている．これに対し，我々の結果は系 A.2.1で与えられる log 関数と Stirling の誤差の精緻な評価を用いるこ

とによりより精緻な定量評価を与えることができる．今

$$Q^0_{n(k)} = \left\{ z_{(k)} = (z_1, \cdots, z_k); 0 \equiv z_0 < z_1 < \cdots < z_k < z_{k+1} \equiv 1 \right\}$$

と定義する．またある与えられた正数 L, M に対して

$$Q^{L,M}_{n(k)} = \left\{ z_{(k)} = (z_1, \cdots, z_k);\ \begin{array}{l} 1 < \frac{z_i - z_{i-1}}{l_{ni} - l_{ni-1}} < 1 + \frac{1}{L} \text{ or} \\ \frac{M}{1+M} < \frac{z_i - z_{i-1}}{l_{ni} - l_{ni-1}} < 1, i = 0, 1, \cdots, k+1 \end{array} \right\},$$

を定義する．ここで 3.3.2節と同様に $l_{ni} = n_i/(n+1)$, $i = 1, \cdots, k$, $l_{n0} = 0$, $l_{nk+1} = 1$, $0 \equiv n_0 < n_1 < n_2 < \cdots < n_k < n_{k+1} \equiv n+1$ とする．3.3.2節と同じ記号の下で次の結果が得られる:

定理 B.4.1 $z_{(k)} \in Q^0_{n(k)} \cap Q^{L,M}_{n(k)} =: A^{L,M}_{n(k)}$ に対して, ordered Dirichlet 分布の密度関数は正規密度関数 g_n を用いて

$$h_n(z_{(k)}) = \underline{g_n(z_{(k)})} \cdot \left(1 + \frac{1}{n+1}\right)^{-k/2} \tag{B.4.1}$$
$$\cdot \exp \left(\begin{array}{l} \frac{1}{12(n+1)} \left(1 - \sum_{i=1}^{k+1} \frac{1}{l_{ni+1} - l_{ni}}\right) - R(n+1) + \sum_{i=1}^{k+1} R(d_i + 1) \\ + \frac{1}{2(n+2)} (z_{n(k)} - l_{n(k)})^t L^{-1}_{n(k)} (z_{n(k)} - l_{n(k)}) \\ - \sum_{i=1}^{k+1} \frac{(z_i - l_{ni}) - (z_{i-1} - l_{ni-1})}{l_{ni} - l_{ni-1}} \\ + \sum_{i=1}^{k+1} \frac{1}{3}(n+1) \frac{\{(z_i - l_{ni}) - (z_{i-1} - l_{ni-1})\}^3}{(l_{ni} - l_{ni-1})^2} \\ + \sum_{i=1}^{k+1} \frac{1}{2} \left(\frac{(z_i - l_{ni}) - (z_{i-1} - l_{ni-1})}{l_{ni} - l_{ni-1}} \right)^2 \\ + \sum_{i=1}^{k+1} \theta(n+1) \frac{\{(z_i - l_{ni}) - (z_{i-1} - l_{ni-1})\}^4}{(l_{ni} - l_{ni-1})^3} \\ - \frac{1}{3} \sum_{i=1}^{k+1} \left(\frac{(z_i - l_{ni}) - (z_{i-1} - l_{ni-1})}{l_{ni} - l_{ni-1}} \right)^3 \\ - \theta \sum_{i=1}^{k+1} \left(\frac{(z_i - l_{ni}) - (z_{i-1} - l_{ni-1})}{l_{ni} - l_{ni-1}} \right)^4 \end{array} \right)$$

と表される．ここで $z_{(k)} \in Q^0_{n(k)} \cap Q^{L,M}_{n(k)}$ に対して

$$g^*_n(Z_{(k)}) = (2\pi)^{-k/2} \cdot \left| L_{n(k)} \right|^{-k/2} \cdot \exp \left[-\frac{1}{2} \left(z_{(k)} - l_{n(k)} \right)^t L^{-1}_{n(k)} \left(z_{(k)} - l_{n(k)} \right) \right] \tag{B.4.2}$$

$$R(x) = \sum_{i=2}^{\infty} \frac{a_{i+1}}{x(x+1)\cdots(x+i)} \tag{B.4.3}$$
$$a_r = \frac{1}{r}\int_0^1 t(1-t)(2-t)\cdots(r-1-t)\left(\frac{1}{2}-t\right)dt,$$

とする．また与えられた正数 L, M と任意の正の整数 n に対し，

$$\frac{1}{6}\left\{1 - \frac{1}{(L+1)^2}\right\} < \theta = \theta(n; L, M) < \max\left\{\frac{1}{3}, \frac{1}{4}\left(1+\frac{1}{M}\right)\right\}. \tag{B.4.4}$$

が成り立つものとする．

注 B.1 定理 B.4.1に於ける $g_n^*(z_{(k)})$ は (3.3.14) 式で定義した $g_n^*(z_{(k)})$ と同じ関数の形をしているが，定義されている領域が違う．また関数 $R(x)$ は後に与えられる (B.4.7) 式によって評価することができる．

証明.

$$t_i = \sqrt{n+1}(z_i - l_{ni}),\ (i = 0, 1, \cdots, k+1) \tag{B.4.5}$$

と置く．この変換によるヤコービアンは $dz_i = (n+1)^{1/2}dt_i$ と与えられるので

$$\begin{aligned}
&\left\{n! \Big/ \prod_{i=1}^{k+1} d_i!\right\} dz_1 \cdots dz_k \\
&= \frac{(n+1)!n_1(n_2-n_1)\cdots(n_k-n_{k-1})(n_{k+1}-n_k)}{(n+1)n_1!(n_2-n_1)!\cdots(n_k-n_{k-1})!(n_{k+1}-n_k)!} \cdot \prod_{i=1}^{k}(n+1)^{-1/2}dt_i \\
&= \frac{(n+1)!}{(n+1)n_1!(n_2-n_1)!\cdots(n_k-n_{k-1})!(n_{k+1}-n_k)!} \\
&\cdot \prod_{i=1}^{k+1}(n_i - n_{i-1})^{1/2} \cdot (n+1)^{1/2} \prod_{i=1}^{k+1}(l_{ni} - l_{ni-1})^{1/2} \cdot dt_1 \cdots dt_k
\end{aligned}$$

と表される．また任意の正数 x に対して次の表現

$$\Gamma(m+1) = \sqrt{2\pi}m^{m+1/2}\exp\left\{-m + \frac{1}{12m} - R(m)\right\} \tag{B.4.6}$$

が成り立つ．ここで

$$\frac{1}{360m(m+1)(m+2)} < R(m) < \frac{1}{360m(m+1)(m+2)} + \frac{1}{32m^2(m+1)(m+2)} \tag{B.4.7}$$

とする.(cf. Matsunawa(1977).) この様にいくつかの計算の後に

$$
\begin{aligned}
&\left\{n! \Big/ \prod_{i=1}^{k+1} d_i!\right\} dz_1 \cdots dz_k \\
&= \frac{1}{(2\pi)^{k/2}} \cdot \frac{(n+1)^{n+3/2}\exp\left(-(n+1)+\dfrac{1}{12(n+1)}-R(n+1)\right)}{(n+1)\displaystyle\prod_{i=1}^{k+1}\left[(d_i+1)^{d_i+1}\exp\left(-(d_i+1)+\dfrac{1}{12(d_i+1)}-R(d_i+1)\right)\right]} \\
&\quad \cdot \frac{\displaystyle\prod_{i=1}^{k+1}(d_i+1)^{1/2}\cdot\prod_{i=1}^{k+1}(l_{ni}-l_{ni-1})^{1/2}}{\displaystyle\prod_{i=1}^{k+1}(d_i+1)^{1/2}} \cdot dt_1 \cdots dt_k \\
&= \frac{\displaystyle\prod_{i=1}^{k+1}(l_{ni}-l_{ni-1})^{-(n_i-n_{i-1}-1/2)}}{(2\pi)^{k/2}} \\
&\quad \cdot \exp\left(\frac{1}{12}\left(\frac{1}{n+1}-\sum_{i=1}^{k+1}\frac{1}{d_i+1}\right)-R(n+1)+\sum_{i=1}^{k+1}R(d_i+1)\right)\cdot dt_1\cdots dt_k
\end{aligned} \tag{B.4.8}
$$

を得る. また, 次の項の評価を行なう.

$$
\begin{aligned}
&\prod_{i=1}^{k+1}(z_i - z_{i-1})^{d_i} \\
&\quad = \prod_{i=1}^{k+1}(l_{ni}-l_{ni-1})^{n_i-n_{i-1}-1}\cdot\prod_{i=1}^{k+1}\left\{1+\frac{t_i-t_{i-1}}{\sqrt{n+1}(l_{ni}-l_{ni-1})}\right\}^{(n+1)(l_{ni}-l_{ni-1})-1}
\end{aligned}
$$

まず初めに

$$
w_i = w_i(n) = \frac{\sqrt{n+1}(l_{ni}-l_{ni-1})}{t_i-t_{i-1}} \tag{B.4.9}
$$

と置く. この時与えられた正数 L, M と $w_i > L$, $w_i < -1-M$ $(i=1,\cdots,k+1)$ を満たす任意の正数 w_i に対して

$$
\begin{aligned}
\ln\left\{1+\frac{t_i-t_{i-1}}{\sqrt{n+1}(l_{ni}-l_{ni-1})}\right\} &= \left[\frac{t_i-t_{i-1}}{\sqrt{n+1}(l_{ni}-l_{ni-1})}-\frac{1}{2}\left\{\frac{t_i-t_{i-1}}{\sqrt{n+1}(l_{ni}-l_{ni-1})}\right\}^2\right. \\
&\qquad \left. +\frac{1}{3}\left\{\frac{t_i-t_{i-1}}{\sqrt{n+1}(l_{ni}-l_{ni-1})}\right\}^3+\theta\left\{\frac{t_i-t_{i-1}}{\sqrt{n+1}(l_{ni}-l_{ni-1})}\right\}^4\right]
\end{aligned}
$$

と表される. このように

$$
\prod_{i=1}^{k+1}\left\{1+\frac{t_i-t_{i-1}}{\sqrt{n+1}(l_{ni}-l_{ni-1})}\right\}^{(n+1)(l_{ni}-l_{ni-1})-1}
$$

$$
\begin{aligned}
&= \prod_{i=1}^{k+1} \exp\left[\ln\left\{1+\frac{t_i - t_{i-1}}{\sqrt{n+1}(l_{ni}-l_{ni-1})}\right\}^{(n+1)(l_{ni}-l_{ni-1})-1}\right] \\
&= \exp\left[\sqrt{n+1}\sum_{i=1}^{k+1}(t_i - t_{i-1}) - \frac{1}{2}\sum_{i=1}^{k+1}\frac{(t_i-t_{i-1})^2}{l_{ni}-l_{ni-1}}\right] \\
&\cdot \exp\left[+\frac{1}{3\sqrt{n+1}}\sum_{i=1}^{k+1}\frac{(t_i-t_{i-1})^3}{(l_{ni}-l_{ni-1})^2} + \frac{\theta}{n+1}\sum_{i=1}^{k+1}\frac{(t_i-t_{i-1})^4}{(l_{ni}-l_{ni-1})^3}\right] \\
&\cdot \exp\left[-\frac{1}{\sqrt{n+1}}\sum_{i=1}^{k+1}\frac{t_i-t_{i-1}}{l_{ni}-l_{ni-1}} + \frac{1}{2(n+1)}\sum_{i=1}^{k+1}\left(\frac{t_i-t_{i-1}}{l_{ni}-l_{ni-1}}\right)^2\right] \\
&\cdot \exp\left[-\frac{1}{3(n+1)^{3/2}}\sum_{i=1}^{k+1}\left(\frac{t_i-t_{i-1}}{l_{ni}-l_{ni-1}}\right)^3 - \frac{\theta}{(n+1)^2}\sum_{i=1}^{k+1}\left(\frac{t_i-t_{i-1}}{l_{ni}-l_{ni-1}}\right)^4\right] \\
&= \exp\left[-\frac{1}{2}\sum_{i=1}^{k+1}\frac{(t_i-t_{i-1})^2}{l_{ni}-l_{ni-1}}\right] \qquad \text{(B.4.10)} \\
&\cdot \exp\left[-\frac{1}{\sqrt{n+1}}\sum_{i=1}^{k+1}\left\{\frac{t_i-t_{i-1}}{l_{ni}-l_{ni-1}} - \frac{(t_i-t_{i-1})^3}{3(l_{ni}-l_{ni-1})^2}\right\}\right] \\
&\cdot \exp\left[+\frac{1}{n+1}\sum_{i=1}^{k+1}\left\{\frac{1}{2}\left(\frac{t_i-t_{i-1}}{l_{ni}-l_{ni-1}}\right)^2 + \theta\frac{(t_i-t_{i-1})^4}{(l_{ni}-l_{ni-1})^3}\right\}\right] \\
&\cdot \exp\left[-\frac{1}{3(n+1)^{3/2}}\sum_{i=1}^{k+1}\left(\frac{t_i-t_{i-1}}{l_{ni}-l_{ni-1}}\right)^3\right] \\
&\cdot \exp\left[-\frac{\theta}{(n+1)^2}\sum_{i=1}^{k+1}\left(\frac{t_i-t_{i-1}}{l_{ni}-l_{ni-1}}\right)^4\right]
\end{aligned}
$$

を得る．よって (B.4.9) 式, (B.4.10) 式と t_iを z_i $(i = 0, 1, \cdots, k+1)$ に変換することと (3.3.15) 式の定義により

$$
\begin{aligned}
h_n(z_{(k)})dz_1\cdots dz_k &= \frac{\prod_{i=1}^{k+1}(l_{ni}-l_{ni-1})^{-1/2}}{(2\pi)^{k/2}} \qquad \text{(B.4.11)} \\
&\cdot \exp\left[-\frac{1}{2}\left\{\sum_{i=1}^{k+1}\frac{l_{ni+1}-l_{ni-1}}{(l_{ni+1}-l_{ni})(l_{ni}-l_{ni-1})}t_i^2 - 2\sum_{i=1}^{k+1}\frac{1}{l_{ni}-l_{ni-1}}t_i t_{i-1}\right\}\right] \\
&\cdot \exp\left[\frac{1}{12}\left(\frac{1}{n+1} - \sum_{i=1}^{k+1}\frac{1}{d_i+1}\right) - R(n+1) + \sum_{i=1}^{k+1}R(d_i+1)\right] \\
&\cdot \exp\left[-\frac{1}{\sqrt{n+1}}\sum_{i=1}^{k+1}\left\{\frac{t_i-t_{i-1}}{l_{ni}-l_{ni-1}} - \frac{(t_i-t_{i-1})^3}{3(l_{ni}-l_{ni-1})^2}\right\}\right] \\
&\cdot \exp\left[+\frac{1}{n+1}\sum_{i=1}^{k+1}\left\{\frac{1}{2}\left(\frac{t_i-t_{i-1}}{l_{ni}-l_{ni-1}}\right)^2 + \theta\frac{(t_i-t_{i-1})^4}{(l_{ni}-l_{ni-1})^3}\right\}\right]
\end{aligned}
$$

$$\cdot \exp\left[-\frac{1}{3(n+1)^{3/2}}\sum_{i=1}^{k+1}\left(\frac{t_i - t_{i-1}}{l_{ni} - l_{ni-1}}\right)^3\right]$$
$$\cdot \exp\left[-\frac{\theta}{(n+1)^2}\sum_{i=1}^{k+1}\left(\frac{t_i - t_{i-1}}{l_{ni} - l_{ni-1}}\right)^4\right] \cdot dt_1 \cdots dt_k$$

が得られる．これにより (B.4.1) 式が証明された．

参考文献

[1] Barron, A. R. and Sheu, C. H. (1991). Approximation of density functions by sequential families. *Ann. Statist.* **19**, 1347-1369.

[2] Boltzmann, L. (1887). Über die Beziehung zwischen dem zwischen dem zweiten Hauptsatze der mechanishen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht, *Winer Berichte* **76**, 373-435.

[3] Kagan, A. M. (1963). On the theory of Fisher's amount of information. *Dokl. Akad. Nauk, SSSR* **151**, 277-278.

[4] Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *Ann. Math. Statist.* **22**, 79-86.

[5] Kullback, S. (1952). Certain inequalities in information theory and the Cramer-Rao inequality *Ann. Math. Statist.* **25**. 745-751.

[6] Kullback, S. (1959). *Information Theory and Statistics.* Wiley, New York.

[7] Matsunawa, T. (1976). Some inequalities based on inverse factorial series, *Ann. Inst. Statist. Math.*, **28**, 291-305.

[8] Matsunawa, T. (1977). Approximations to the probabilities of binomial and multinomial random variables and chisquare type statistics, *Ann. Inst. Statist. Math.*, **29**, 333-358.

[9] Matsunawa, T. (1982). Uniform ϕ-equivalence of probability distributions based on information and related measure of discrepancy. *Ann. Inst. Statist. Math.* **34**, A,1-17.

[10] Matsunawa, T. (1986). Modified information criteria for a uniform approximate equivalence of probability distributions. *Ann. Inst. Statist. Math.* **38**, A,205-222.

[11] 松縄 規 (1994). 分布の起源 - ノンパラメトリックな統計的不確定性関係と統計基礎方程式 -, 統計数理, **42**, 197-214.

[12] 松縄 規 (1995). 分布の発展 - ルジャンドル変換と正準情報量規準 -, 統計数理, **43**, 293-311.

[13] Matusita, K. (1955). Decision rules, based on the distance, for problems of fit, two samples, and estimation. *Ann. Inst. Statist. Math.* **26**, 631-640.

[14] Mosteller, F. (1946). On some useful "inefficient" statistics, *Ann. Math. Statist.*, **17**, 377-408.

[15] Neyman, J. (1937). Smooth test for goodness of fit. *Scand. Actuar. J.* **20**, 149-199.

[16] Reiss, R.D. (1974). The asymptotic normality and asymptotic expansions for the joint distribution of several order statistics, *Colloquia Math. Soc. Janos Bolyai*. 297-340.

[17] Walker, A. M. (1968). A note on the asymptotic distribution of sample quantiles, *J. Roy. Statist. Soc.,* **B.30**, 570-575.

[18] Weiss, L. (1969). The asymptotic joint distribution of an increasing number of sample quantiles, *Ann. Inst. Statist. Math.* **21**. 257-263.

[19] Wilks, S. S. (1962). *Mathematical Statistics.* Wiley, New York.

[20] 山田 智哉, 松縄 規 (1998). 近似主領域での修正情報量に基づく確率分布間の一様近似と多変量一般指数型分布族の摇動の定量評価への応用, 統計数理, **46**, 461-476

[21] Yamada, T and Matsunawa, T. Quantitative approximation to the ordered Dirichlet distribution under varying basic probability spaces. *Ann. Inst. Statist. Math.* (submitted)