

はじめに

このノートは統計解析と情報統計学の講義録である。

2021 年 4 月 9 日更新

2015 年 11 月 10 日更新

目次

はじめに	i
第 1 章 確率・確率変数・期待値	3
1 確率	3
1.1 試行・標本空間・事象	3
1.2 確率の定義	6
2 条件つき確率と独立性	10
3 確率変数	12
4 確率密度関数と確率関数	15
5 確率変数の期待値	17
6 積率と積率母関数	19
7 確率変数の変換	22
8 確率変数の不等式 1	27
9 演習問題	31
第 2 章 1 次元の確率分布の代表的モデル	41
1 離散型確率変数のモデル	41
2 連続型確率変数のモデル	45
3 演習問題	50
第 3 章 多次元の確率変数	53
1 同時分布と周辺分布	53
1.1 同時確率関数	55
1.2 同時確率密度関数	56
1.3 独立性	57
1.4 同時分布に関する期待値	59
2 共分散と相関係数	61

3	条件付き分布と独立性	66
3.1	離散型確率変数の場合	66
3.2	連続型確率変数の場合	68
3.3	独立性との関係	69
4	条件付き期待値	69
5	2次元の確率変数の変換	74
5.1	離散型確率ベクトルの場合	74
5.2	連続型の場合	76
6	多次元分布の代表的なモデル	78
6.1	二変量正規分布	78
7	演習問題	82
第4章	標本分布論と漸近分布論	93
1	標本分布論の枠組み	93
1.1	ランダム標本	93
1.2	統計量と標本分布	93
2	正規分布からのランダム標本	99
2.1	t 分布と F 分布	101
3	順序統計量	107
4	確率変数の列の収束について	112
5	大数の法則と中心極限定理	125
6	演習問題	128
第5章	データの縮約	135
1	十分統計量	135
1.1	十分統計量の定義	136
1.2	分解定理	138
2	演習問題	143
第6章	点推定法	145
1	点推定量の性質	145
1.1	近隣度 (closeness)	145
1.2	平均 2 乗誤差	146

1.3	推定量の一致性	149
2	最尤法	150
3	不偏推定と Cramér-Rao の定理	156
3.1	不偏推定量の分散の下限	157
4	十分性と完備性	162
5	ベイズ推定量	169
6	演習問題	169
第 7 章	検定法	173
1	検定論の枠組み	173
2	最強力検定とネイマン・ピアソンの補題	180
3	一様最強力検定	185
4	演習問題	185
第 8 章	信頼区間	187
1	信頼区間の定義	187
2	ピボット法	188
3	演習問題	193
付録 A	補遺	195
1	二項定理と多項定理	195
2	数と数列	195
2.1	数の集合	195
2.2	指数法則	196
2.3	数列	196
2.4	有限級数	197
2.5	無限級数	197
2.6	上極限と下極限	198
3	微積分学の復習	198
3.1	関数の極限	198
3.2	連続関数の定義	200
3.3	合成関数の極限値の性質	200
3.4	導関数の定義	200

3.5	合成関数の微分	201
3.6	逆関数の微分法	201
3.7	微積分学の基本定理	201
3.8	指数関数の級数展開	202
3.9	積分の計算	202
3.10	広義積分	202
4	偏微分とその計算法	205
4.1	偏導関数	205
4.2	合成関数の偏微分	206
4.3	2変数関数のテイラー展開	206
4.4	2変数関数の極大・極小	207
5	重積分	207
5.1	重積分の定義：リーマン和と積分可能性	207
5.2	面積確定な有界集合	208
5.3	重積分の性質	209
5.4	重積分の計算法：累次積分	210
5.5	重積分の変数変換	210
5.6	広義積分	211
6	積分記号下の微分	214

第1章 確率・確率変数・期待値

統計推測理論は、確率論を基礎として構築される。この章では、確率論の基本的な事柄について学ぶ。

1 確率

不確実性や変動というものは多くの状況では避けることができないものである。確率論はこれらの概念を数学的に定式化する道具である。

1.1 試行・標本空間・事象

実験とか観測とか調査とかを総称して試行という。試行を行ったときに起こり得る結果のすべての集まりを標本空間といい、 Ω と記すことにする。標本空間 Ω の点を標本点とか根元事象という。 Ω が有限または可算の場合を離散的という。

例 1.1 硬貨投げを考えよう。硬貨には表 (H) と裏 (T) があることから、試行のすべての集まりである標本空間は $\Omega = \{H, T\}$ となる。

部分集合 $A \subset \Omega$ を事象といい、 Ω を全事象、 \emptyset を空事象という。事象に対して、和、交わり (積)、直積、補集合、差などの演算を考える。

事象の記法 事象に対する記法は集合に対するものと基本的には同じである。 A と B を事象とする。 a が事象 A に含まれる標本点であることを $a \in A$ と書き、 a が事象 A に含まれる標本点でないことを $a \notin A$ と書く。また、1つの標本点も含まない事象を空事象といい、 \emptyset と書く。 A が B に含まれることを $A \subset B$ と書く。 A と B の共通部分を $A \cap B := \{\omega : \omega \in A \text{ かつ } \omega \in B\}$ 、 A と B の和事象を $A \cup B := \{\omega : \omega \in A \text{ また } \omega \in B\}$ 、 A に関する B の差集合を $A \setminus B := \{\omega : \omega \in A \text{ かつ } \omega \notin B\}$ と記す。 A の補事象を $A^c := \Omega \setminus A$ と書く。

和と共通部分に対して以下が成立する .

命題 1.1 (事象の演算に関する基本的性質) A, B, C を事象とする .

$$(i) \quad \begin{aligned} A \cup A &= A \\ A \cap A &= A. \end{aligned} \quad (\text{ベキ等律})$$

$$(ii) \quad \begin{aligned} A \cup B &= B \cup A \\ A \cap B &= B \cap A. \end{aligned} \quad (\text{交換律})$$

$$(iii) \quad \begin{aligned} (A \cup B) \cup C &= A \cup (B \cup C) \\ (A \cap B) \cap C &= A \cap (B \cap C). \end{aligned} \quad (\text{結合律})$$

$$(iv) \quad \begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C) \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C). \end{aligned} \quad (\text{分配法則})$$

$$(v) \quad \begin{aligned} (A \cup B)^c &= A^c \cap B^c \\ (A \cap B)^c &= A^c \cup B^c. \end{aligned} \quad (\text{ド・モルガンの法則})$$

証明 証明は略 . (「集合・位相」(佐久間一浩著, 共立) や「集合・位相入門」(松坂和夫著, 岩波) 等を参照のこと) □

一般に, 集合 Γ のそれぞれの元 γ に対して, 事象 A_γ が定まっているとき, この A_γ のなす事象の集合を「 Γ を添え字集合とする事象の族」といい,

$$\{A_\gamma : \gamma \in \Gamma\} \quad \{A_\gamma\}_{\gamma \in \Gamma}$$

と書く . 任意の濃度を持つ添え字集合 Γ に対して

$$\cup_{\gamma \in \Gamma} A_\gamma = \{\omega \in \Omega : \text{ある } \gamma \in \Gamma \text{ が存在して, } \omega \in A_\gamma\}$$

$$\cap_{\gamma \in \Gamma} A_\gamma = \{\omega \in \Omega : \text{すべての } \gamma \in \Gamma \text{ に対して, } \omega \in A_\gamma\}$$

とする .

例 1.2 たとえば, $\Omega = (0, 1]$ とし, $A_i = [1/i, 1]$, $i = 1, 2, \dots$ とする . したがって, $\Gamma = \mathbb{N}$ (自然数) である . このとき,

$$\begin{aligned} \cup_{i=1}^{\infty} A_i &= \cup_{i=1}^{\infty} [1/i, 1] \\ &= \{\omega \in (0, 1] : \text{ある } i \in \mathbb{N} \text{ が存在して, } \omega \in [1/i, 1]\} \\ &= \Omega. \end{aligned}$$

定義 1.1 $\{A_\gamma : \gamma \in \Gamma\}$ が互いに排反 (互いに素) であるとは, すべての $\alpha, \beta \in \Gamma$ に対して, $\alpha \neq \beta$ ならば, $A_\alpha \cap A_\beta = \emptyset$ が成り立つことである.

命題 1.2 (和と交わりおよびド・モルガンの法則の一般化) $\{A_\gamma : \gamma \in \Gamma\}$ を集合族¹とする. このとき,

- (i) $(\cup_{\gamma \in \Gamma} A_\gamma) \cap B = \cup_{\gamma \in \Gamma} (A_\gamma \cap B)$.
- (ii) $(\cap_{\gamma \in \Gamma} A_\gamma) \cup B = \cap_{\gamma \in \Gamma} (A_\gamma \cup B)$.
- (iii) $\{\cup_{\gamma \in \Gamma} A_\gamma\}^c = \cap_{\gamma \in \Gamma} A_\gamma^c$.
- (iv) $\{\cap_{\gamma \in \Gamma} A_\gamma\}^c = \cup_{\gamma \in \Gamma} A_\gamma^c$.

証明 証明は略.

□

定義 1.2 可算無限個の事象列 $\{A_n\}_{n=1}^\infty$ に対して, その上極限事象と下極限事象をそれぞれ

$$\begin{aligned}\overline{\lim}_{n \rightarrow \infty} A_n &= \cap_{m=1}^\infty \cup_{n=m}^\infty A_n \\ \underline{\lim}_{n \rightarrow \infty} A_n &= \cup_{m=1}^\infty \cap_{n=m}^\infty A_n\end{aligned}$$

で定める. 上極限事象は $\{A_n\}$ のうちの無限個が起きるという事象を, 下極限事象は $\{A_n\}$ のうちのある番号から先のすべてが起こるという事象を表す. すなわち, $\omega \in \overline{\lim}_{n \rightarrow \infty} A_n$ とは, 無限に多くの n に対して, $\omega \in A_n$ となることであり, $\omega \in \underline{\lim}_{n \rightarrow \infty} A_n$ とは, ある番号から先のすべての n に対して $\omega \in A_n$ である. 上極限と下極限が一致するとき, $\lim_{n \rightarrow \infty} A_n$ と書き, $\{A_n\}_{n=1}^\infty$ の極限事象という.

命題 1.3 つぎの関係式が成り立つ.

- (i) $\liminf_{n \rightarrow \infty} A_n \subset \limsup_{n \rightarrow \infty} A_n$.
- (ii) $\left(\liminf_{n \rightarrow \infty} A_n \right)^c = \limsup_{n \rightarrow \infty} A_n^c$.

¹集合のあつまり

$$(iii) \quad \left(\limsup_{n \rightarrow \infty} A_n \right)^c = \liminf_{n \rightarrow \infty} A_n^c .$$

$$(iv) \quad A_1 \subset A_2 \subset \cdots \text{ ならば, } \lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n .$$

$$(v) \quad A_1 \supset A_2 \supset \cdots \text{ ならば, } \lim_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n .$$

証明 証明は略 .

□

1.2 確率の定義

定義 1.3 Ω の部分集合の集まり \mathcal{F} は次の条件をみたすとき, Ω 上の有限加法族であるという .

$$(i) \quad \Omega \in \mathcal{F} .$$

$$(ii) \quad A \in \mathcal{F} \implies A^c \in \mathcal{F} .$$

$$(iii) \quad A_1, A_2, \dots, A_k \in \mathcal{F} \implies \bigcup_{i=1}^k A_i \in \mathcal{F} .$$

さらに,

$$(iii)' \quad A_1, A_2, \dots, A_k, \dots \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F} .$$

をみたすとき, Ω 上の完全加法族 (σ -加法族) という .

注意 1.1 $A_1, A_2, \dots, A_k, \dots \in \mathcal{F}$ ならば, $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$ である . Ω 上の完全加法族で最も大きなものは Ω のすべての部分集合からなる族であり, 最小の完全加法族は $\{\emptyset, \Omega\}$ である .

確率論において用いられる完全加法族の重要な例をいくつか考察する . まず, 一般的な定義から始めよう .

定義 1.4 \mathcal{F} は Ω の部分集合の任意の族とする . \mathcal{F} を含む最小の完全加法族を $\sigma(\mathcal{F})$ と表す . 言い換えれば, $\sigma(\mathcal{F})$ は \mathcal{F} を含む完全加法族全体の交わりである . Ω のすべての部分集合全体は完全加法族であることから, 少なくとも 1 つは \mathcal{F} を含む完全加法族は存在することに注意しよう .

例 1.3 $\Omega = \mathbb{R}$ とし, 次の部分集合族を考える .

- (i) $\mathcal{F}_1 = \{(a, b) : a, b \in \mathbb{R}, a < b\}$.
- (ii) $\mathcal{F}_2 = \{[a, b) : a, b \in \mathbb{R}, a < b\}$.
- (iii) $\mathcal{F}_3 = \{(a, b] : a, b \in \mathbb{R}, a < b\}$.
- (iv) $\mathcal{F}_4 = \{[a, b] : a, b \in \mathbb{R}, a < b\}$.
- (v) $\mathcal{F}_5 = \{(-\infty, a] : a \in \mathbb{R}\}$.
- (vi) $\mathcal{F}_6 = \{(-\infty, a) : a \in \mathbb{R}\}$.
- (vii) \mathcal{F}_7 は \mathbb{R} の開集合の族.
- (viii) \mathcal{F}_8 は \mathbb{R} の閉集合の族.

このとき,

$$\sigma(\mathcal{F}_1) = \sigma(\mathcal{F}_2) = \sigma(\mathcal{F}_3) = \sigma(\mathcal{F}_4) = \sigma(\mathcal{F}_5) = \sigma(\mathcal{F}_6) = \sigma(\mathcal{F}_7) = \sigma(\mathcal{F}_8)$$

である. この完全加法族を $\mathcal{B}(\mathbb{R})$ と記し, \mathbb{R} のボレル集合族という.

定義 1.5 \mathcal{F} を Ω 上の完全加法族としたとき, (Ω, \mathcal{F}) のことを可測空間という.

定義 1.6 任意の可測空間 (Ω, \mathcal{F}) に対して, 確率とは \mathcal{F} 上で定義された関数 \mathbb{P} で以下の条件をみたすものである.

- (i) 任意の $A \in \mathcal{F}$ について, $\mathbb{P}(A) \geq 0$.
- (ii) $\mathbb{P}(\Omega) = 1$.
- (iii) $i = 1, 2, \dots$ について $A_i \in \mathcal{F}$ かつ $A_i \cap A_j = \emptyset (i \neq j)$ ならば,

$$\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

$\mathbb{P}(A)$ は事象 A の起きる確率とよばれる. $(\Omega, \mathcal{F}, \mathbb{P})$ を確率空間とよぶ.

注意 1.2 Ω が離散のとき, \mathcal{F} は Ω のすべての部分集合の集まりと通常する. また, $\Omega = \mathbb{R}$ のとき, \mathcal{F} として, \mathbb{R} のすべての部分集合の集まりを採用するのでなく, ボレル集合族 $\mathcal{B}(\mathbb{R})$ を採用するのが標準的である.

命題 1.4 (確率の性質) (i) $\mathbb{P}(\emptyset) = 0$.

(ii) 任意の $A \in \mathcal{F}$ に対し, $\mathbb{P}(A) \leq 1$.

(iii) 任意の $A \in \mathcal{F}$ に対し, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.

(iv) 任意の $A, B \in \mathcal{F}$ に対し, $\mathbb{P}(A \cap B^c) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$.

(v) $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

(vi) $A \subset B$ ならば, $\mathbb{P}(A) \leq \mathbb{P}(B)$.

(vii) $n = 1, 2, \dots$ に対し, $A_n \in \mathcal{F}$ ならば,

$$\mathbb{P}(\cup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

(viii) $A_n \subset A_{n+1}$ ならば,

$$\mathbb{P}(\cup_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n).$$

(ix) $A_n \supset A_{n+1}$ ならば,

$$\mathbb{P}(\cap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n).$$

(x) $A_n \supset A_{n+1}$ かつ $\cap_{n=1}^{\infty} A_n = \emptyset$ ならば,

$$\mathbb{P}(\cap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n) = 0.$$

証明 (i) 定義 1.6(iii) において, $A_1 = A_2 = \dots = \emptyset$ とおくと $\mathbb{P}(\emptyset) = \sum_{i=1}^{\infty} \mathbb{P}(\emptyset)$ となる. この式は $\mathbb{P}(\emptyset) = 0$ のときのみ成立する.

(ii) は 定義 1.6(ii) と (vi) からわかる.

(iii) 定義 1.6(iii) において, $A_1 = A, A_2 = A^c, A_3 = A_4 = \dots = \emptyset$ とおく. $A \cup A^c = \Omega$, (i) と定義 1.6(i) から $\mathbb{P}(A) + \mathbb{P}(A^c) = \mathbb{P}(\Omega) = 1$ となることよりわかる. また, $A_{n+1} = A_{n+2} = \dots = \emptyset$ とおくことより, 任意の正の整数 n に対して, $A_i \cap A_j = \emptyset (i \neq j)$ ならば,

$$\mathbb{P}(\cup_{i=1}^n A_i) = \sum_{i=1}^n \mathbb{P}(A_i) \quad (1.1)$$

が成立することもわかる.

(iv) $(A \cap B) \cup (A \cap B^c) = A$ と $(A \cap B) \cap (A \cap B^c) = \emptyset$ に注意して, $A_1 = A \cap B, A_2 = A \cap B^c$ として (1.1) を用いればよい.

(v) $A \cup B = A \cup (B \setminus (A \cap B))$ として, (1.1) を用いると $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \setminus (A \cap B))$ となる. しかし, $\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(B \setminus (A \cap B))$ である. これを代入すればよい.

(vi) $B = A \cup (B \setminus A)$ と定義 1.6(i) から $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geq \mathbb{P}(A)$ がわかる.

(vii) 各 n に対して, $B_n = A_n \setminus (\cup_{i=1}^{n-1} A_i)$ とおくと $B_n \subset A_n$ となる. さらに, B_1, B_2, \dots は互いに排反で $\cup_{n=1}^{\infty} B_n = \cup_{n=1}^{\infty} A_n$ がわかる. したがって, 定義 1.6(iii) と (vi) から

$$\mathbb{P}(\cup_{n=1}^{\infty} A_n) = \mathbb{P}(\cup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mathbb{P}(B_n) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n)$$

となる.

(viii) $A_0 = \emptyset$ とし, $B_n = A_n \setminus A_{n-1}$ とおく. B_1, B_2, \dots は互いに排反で $\cup_{n=1}^{\infty} B_n = \cup_{n=1}^{\infty} A_n$ となるので,

$$\begin{aligned} \mathbb{P}(\cup_{n=1}^{\infty} A_n) &= \mathbb{P}(\cup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mathbb{P}(B_n) = \lim_{n \rightarrow \infty} \sum_{i=1}^n \mathbb{P}(B_i) \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n [\mathbb{P}(A_i) - \mathbb{P}(A_{i-1})] = \lim_{n \rightarrow \infty} \mathbb{P}(A_n) \end{aligned}$$

からわかる.

(ix) $B_n = A_n^c$ とおくと $B_n \subset B_{n+1}$ となる. ド・モルガンの法則 (命題 1.2(iii)) から $\{\cup_{n=1}^{\infty} A_n\}^c = \cup_{n=1}^{\infty} A_n^c = \cup_{n=1}^{\infty} B_n$ になることに注意して, (iii) および

(ix) を用いると

$$\begin{aligned} 1 - \mathbb{P}(\cap_{n=1}^{\infty} A_n) &= \mathbb{P}(\{\cap A_n\}^c) = \mathbb{P}(\cup_{n=1}^{\infty} B_n) = \lim_{n \rightarrow \infty} \mathbb{P}(B_n) = 1 - \lim_{n \rightarrow \infty} \mathbb{P}(B_n^c) \\ &= 1 - \lim_{n \rightarrow \infty} \mathbb{P}(A_n) \end{aligned}$$

なることからわかる。

□

2 条件つき確率と独立性

$(\Omega, \mathcal{F}, \mathbb{P})$ を確率空間とし, $B \in \mathcal{F}$ かつ $\mathbb{P}(B) > 0$ とする. 事象 B 上の条件つき確率とは, 完全加法族 \mathcal{F} 上で定義された確率であって, 任意の $A \in \mathcal{F}$ について

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

で定める. $\mathbb{P}(A|B)$ を事象 B が与えられたときの事象 A の条件つき確率という.

命題 1.5 $\mathbb{P}(\cdot|B)$ は確率測度である. すなわち,

- (i) 任意の $A \in \mathcal{F}$ に対し, $\mathbb{P}(A|B) \geq 0$
- (ii) $\mathbb{P}(\Omega|B) = 1$
- (iii) $i = 1, 2, \dots$ に対し, $A_i \in \mathcal{F}$ かつ $A_i \cap A_j = \emptyset (i \neq j)$ ならば,

$$\mathbb{P}(\cup_{i=1}^{\infty} A_i|B) = \sum_{i=1}^{\infty} \mathbb{P}(A_i|B)$$

証明 条件付確率と確率の定義 1.3 よりわかる。

□

Ω を互いに排反な事象 A_1, A_2, \dots, A_n に分割されるとする. すなわち, $A_i \cap A_j = \emptyset (i \neq j)$ かつ $\cup_{i=1}^n A_i = \Omega$ である. $i = 1, 2, \dots, n$ に対し, $\mathbb{P}(A_i) > 0$ のとき, 任意の事象 B の確率は条件つき確率を用いてつぎのように表現できる.

$$\mathbb{P}(B) = \sum_{i=1}^n \mathbb{P}(A_i) \mathbb{P}(B|A_i)$$

これを全確率の法則という。

事象 A_1, A_2, \dots, A_n は互いに排反で Ω を分割するとする。事象 B が起こる前と後のことを事前と事後とよぶことにする。事象 A_i に対して $\mathbb{P}(A_i)$ は事象 B が起こる前に与えられる確率であるから事前確率といい、 $\mathbb{P}(A_i|B)$ は B が起こった後で与えられる確率であるから事後確率という。

定理 1.1 (ベイズの法則) B と $A_i (i = 1, 2, \dots, n)$ を $\mathbb{P}(B) > 0, \mathbb{P}(A_i) > 0$ なる事象とする。事後確率は事前確率を用いてつぎのように表現できる。

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(A_i)\mathbb{P}(B|A_i)}{\sum_{i=1}^n \mathbb{P}(A_i)\mathbb{P}(B|A_i)}$$

証明 全確率の法則を用いればよい。 □

定義 1.7 2つの事象 A, B が確率的に独立であるとは

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

が成り立つことをいう。独立でないとき、従属という。

注意 1.3 事象 A と B は独立で $\mathbb{P}(A) > 0, \mathbb{P}(B) > 0$ とする。このとき、

$$\mathbb{P}(B|A) = \mathbb{P}(B), \quad \mathbb{P}(A|B) = \mathbb{P}(A)$$

である。

定義 1.8 事象 A_1, A_2, \dots, A_n が独立であるとは、すべての $k \in \{2, 3, \dots, n\}$ と $i_1 < i_2 < \dots < i_k, i_l \in \{1, 2, \dots, n\}, l = 1, 2, \dots, k$, に対し

$$\mathbb{P}(\cap_{l=1}^k A_{i_l}) = \prod_{l=1}^k \mathbb{P}(A_{i_l})$$

が成り立つことである。すべての $A_i, A_j, i, j = 1, 2, \dots, n, i \neq j$ に対し

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$$

が成り立つとき、対独立という。

注意 1.4 対独立であっても、独立でない場合がある。

定義 1.9 \mathcal{F}_1 と \mathcal{F}_2 を完全加法族 \mathcal{F} の部分完全加法族とする。このとき、 \mathcal{F}_1 と \mathcal{F}_2 が独立であるとは、すべての $A \in \mathcal{F}_1$ と $B \in \mathcal{F}_2$ に対し

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

が成り立つことをいう。

3 確率変数

$(\Omega, \mathcal{F}, \mathbb{P})$ を確率空間とし、 X を Ω から \mathbb{R} への写像とする。 X によるボレル集合 $B \in \mathcal{B}(\mathbb{R})$ の逆像は Ω の部分集合で $X^{-1}(B)$ と記し

$$X^{-1}(B) = \{\omega \in \Omega \mid X(\omega) \in B\}$$

で定義する。

命題 1.6 (逆像の性質) $B, B', \{B_\gamma : \gamma \in \Gamma\}$ はボレル集合とする。このとき

- (i) $B \subset B'$ ならば、 $X^{-1}(B) \subset X^{-1}(B')$.
- (ii) $X^{-1}(\cup_{\gamma \in \Gamma} B_\gamma) = \cup_{\gamma \in \Gamma} X^{-1}(B_\gamma)$
 $X^{-1}(\cap_{\gamma \in \Gamma} B_\gamma) = \cap_{\gamma \in \Gamma} X^{-1}(B_\gamma)$.
- (iii) B と B' が互いに排反ならば、 $X^{-1}(B)$ と $X^{-1}(B')$ も互いに排反である。
- (iv) $X^{-1}(B^c) = \{X^{-1}(B)\}^c$.

証明 集合・位相入門 (松坂和夫, 岩波) 等を参照.) □

定義 1.10 写像 $X : \Omega \rightarrow \mathbb{R}$ が確率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 上の (実) 確率変数であるとは、任意の $B \in \mathcal{B}(\mathbb{R})$ に対して、 $X^{-1}(B) \in \mathcal{F}$ をみたすときをいう。

注意 1.5 すべての $a \in \mathbb{R}$ に対して、 $\{\omega \in \Omega : X(\omega) \leq a\} \in \mathcal{F}$ をみたすことと同値である。また、すべての $a \in \mathbb{R}$ に対して、 $\{X < a\} \in \mathcal{F}$ とも同値である。

確率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ 上の確率変数 X により $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度 P_X がつぎのように定まる：任意の $B \in \mathcal{B}(\mathbb{R})$ に対して

$$P_X(B) = \mathbb{P}(X \in B) = \mathbb{P}(\omega \in X^{-1}(B)) = \mathbb{P}(X^{-1}(B))$$

P_X のことを X の分布という。

X_1, X_2, \dots, X_p が $(\Omega, \mathcal{F}, \mathbb{P})$ 上の確率変数のとき, $\mathbf{X} = (X_1, X_2, \dots, X_p)'$ を p -次元確率変数という。

命題 1.7 (確率変数の性質) X, Y および $\{X_n\}_{n=1}^{\infty}$ は $(\Omega, \mathcal{F}, \mathbb{P})$ 上の確率変数 (列) とする。

- (i) すべての $a, b \in \mathbb{R}$ に対し, $aX + bY$ も確率変数である。
- (ii) $\max\{X, Y\}$ も $\min\{X, Y\}$ も確率変数である。
- (iii) XY も確率変数である。
- (iv) 各 $\omega \in \Omega$ に対し, $Y(\omega) \neq 0$ ならば, X/Y も確率変数である。
- (v) $\sup_n X_n, \inf_n X_n, \limsup_{n \rightarrow \infty} X_n, \liminf_{n \rightarrow \infty} X_n$ も確率変数である。

証明 (i) を示すために, $X + Y$ と aX も確率変数であることを示す。各 t に対して,

$$\{X + Y < t\} = \cup_{r \in \mathbb{Q}} (\{X < r\} \cap \{Y < t - r\}) \in \mathcal{F}$$

となる。ただし, \mathbb{Q} は有理数とする。つぎに, aX も確率変数であることを示す。 $a > 0$ の場合, $\{aX \leq t\} = \{X \leq t/a\} \in \mathcal{F}$ 。 $a < 0$ の場合, $\{aX \leq t\} = \{X \geq t/a\} \in \mathcal{F}$ 。

(ii) を示すためには,

$$\{\max\{X, Y\} \leq t\} = \{X \leq t\} \cap \{Y \leq t\}$$

と

$$\{\min\{X, Y\} \leq t\} = \{X \leq t\} \cup \{Y \leq t\}$$

に注意すればよい。

(iii) を示すためには、

$$\{X^2 \leq t\} = \{-\sqrt{t} \leq X \leq \sqrt{t}\} = \{X \leq \sqrt{t}\} \setminus \{X \leq -\sqrt{t}\}$$

から X^2 は確率変数になることに注意する。さらに、

$$XY = \frac{1}{2}\{(X+Y)^2 - (X-Y)^2\}$$

から証明される。

(iv) を示すためには、

$$\{\sup_n X_n \leq t\} = \bigcap_n \{X_n \leq t\}, \quad \{\inf_n X_n \geq t\} = \bigcap_n \{X_n \geq t\}$$

から $\sup_n X_n$ と $\inf_n X_n$ は確率変数であることがわかる。さらに、

$$\limsup_n X_n = \inf_n \sup_{k \geq n} X_k, \quad \liminf_n X_n = \sup_n \inf_{k \geq n} X_k$$

から (iv) は示された。 □

定義 1.11 $(\Omega, \mathcal{F}, \mathbb{P})$ 上の確率変数 X の分布関数を

$$F_X(x) = P_X((-\infty, x]) = \mathbb{P}(X \leq x) = \mathbb{P}(\omega \in \Omega : X(\omega) \leq x)$$

で定義する。ただし、 x は任意に実数である。

命題 1.8 (分布関数の性質) (i) (単調性) $x, y \in \mathbb{R}$ とする。このとき、 $x < y$ ならば、 $F_X(x) \leq F_X(y)$ 。

(ii) (右連続性) $\lim_{y \rightarrow x+0} F_X(y) = F_X(x)$ 。

(iii) 任意の $x \in \mathbb{R}$ に対して、 $0 \leq F_X(x) \leq 1$ で $F_X(-\infty) = \lim_{x \rightarrow -\infty} F_X(x) = 0$ 、 $F_X(\infty) = \lim_{x \rightarrow \infty} F_X(x) = 1$ である。

証明 (i) $\{\omega \in \Omega : X(\omega) \in (-\infty, x]\} \subset \{\omega \in \Omega : X(\omega) \in (-\infty, y]\}$ に注意して命題 1.4(vi) を用いると

$$\begin{aligned} F_X(x) &= P_X((-\infty, x]) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in (-\infty, x]\}) \\ &\leq \mathbb{P}(\{\omega \in \Omega : X(\omega) \in (-\infty, y]\}) \\ &= P_X((-\infty, y]) = F_X(y). \end{aligned}$$

(ii) $\{x_n\}$ を単調列で $x_n \rightarrow x$ とする. 命題 1.4(ix) より

$$\begin{aligned} F_X(x) &= \mathbb{P}(\{\omega \in \Omega : X(\omega) \in (-\infty, x]\}) \\ &= \lim_{n \rightarrow \infty} \mathbb{P}(\{\omega \in \Omega : X(\omega) \in (-\infty, x_n]\}) = \lim_{n \rightarrow \infty} F_X(x_n). \end{aligned}$$

(iii) $n \rightarrow \infty$ に対して, $(-\infty, -n] \downarrow \emptyset$ と $(-\infty, n] \uparrow \mathbb{R}$ に注意する. あとは (ii) と同様にすればよい. \square

確率変数 X の分布関数が $F_X(\cdot)$ である場合「確率変数 X は分布 $F_X(\cdot)$ に従う」といい, “ $X \sim F_X$ ” と書くことにする.

注意 1.6 分布関数 $F(\cdot)$ が与えられたとき, $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度 μ で

$$F(x) = \mu((-\infty, x]), \quad x \in \mathbb{R}$$

となるものが一意的に存在することが知られている.

X と Y を確率変数としたとき, X と Y の分布が等しいとは, 任意の $A \in \mathcal{B}(\mathbb{R})$ に対し

$$P_X(A) = P_Y(A)$$

が成り立つ²ときをいう.

X と Y の分布が等しいために必要十分条件は, すべての $x \in \mathbb{R}$ に対し, $F_X(x) = F_Y(x)$ が成り立つことである.

4 確率密度関数と確率関数

定義 1.12 $F_X(\cdot)$ を確率変数 X の分布関数とする. X が連続型確率変数であるとは, F_X が \mathbb{R} 上の連続関数のときをいう. また, X が離散型確率変数

²これは $P(X \in A) = P(Y \in A)$ である.

であるとは, F_X が \mathbb{R} 上の階段関数のときをいう.

定義 1.13 X を離散型確率変数とする. X の確率関数を

$$f_X(x) = \mathbb{P}(X = x), \quad \text{すべての } x$$

で定める.

注意 1.7 確率変数 X が離散型のとき, $\{x \in \mathbb{R} : f_X(x) > 0\}$ は高々可算個であることを示すことができる. また, $f_X(x) > 0$ なる点に対し,

$$f_X(x) = \mathbb{P}(X \leq x) - \mathbb{P}(X < x) = F_X(x) - \lim_{y \rightarrow x-0} F_X(y)$$

となる. $S := \{x \in \mathbb{R} : f_X(x) > 0\}$ としたとき,

$$p_i = f_X(x_i), \quad x_i \in S, \quad i = 1, 2, \dots$$

を離散型確率変数 X の分布とよぶことにする. p_i と $x_i, i = 1, 2, \dots$ を表にまとめたものを確率分布表という.

命題 1.9 (確率関数の性質) (i) $0 \leq f_X(x) \leq 1$

$$(ii) \quad \sum_{x \in S} f_X(x) = 1$$

証明 定義よりわかる. □

定義 1.14 X を連続型確率変数とし, $F_X(\cdot)$ をその分布関数とする. \mathbb{R} 上の非負値関数 $f_X(\cdot)$ で任意の $x \in \mathbb{R}$ に対して

$$F_X(x) = \int_{-\infty}^x f_X(y) dy$$

をみたすものが存在するとき, $f_X(\cdot)$ を X の確率密度関数という.

X は密度 $f_X(\cdot)$ を持つとか分布 $f_X(\cdot)$ に従うという. さらに, $F_X(\cdot)$ が \mathbb{R} 上で微分可能ならば

$$f_X(x) = \frac{d}{dx} F_X(x)$$

となる.

注意 1.8 X が連続型確率変数ならば、任意の $x \in \mathbb{R}$ において、 $\mathbb{P}(X = x) = 0$ となる。すなわち、 $\mathbb{P}(\omega \in \Omega : X(\omega) = 0) = 0$ である。なぜならば、任意の $\epsilon > 0$ に対して

$$\begin{aligned} \mathbb{P}(X = x) &\leq \mathbb{P}(x - \epsilon < X \leq x) \\ &= \mathbb{P}(X \leq x) - \mathbb{P}(X \leq x - \epsilon) \\ &= F_X(x) - F_X(x - \epsilon) \end{aligned}$$

となり、 $\epsilon \downarrow 0$ とすれば、 $F_X(\cdot)$ の連続性より

$$\mathbb{P}(X = x) \leq \lim_{\epsilon \downarrow 0} \{F_X(x) - F_X(x - \epsilon)\} = 0$$

となる。

命題 1.10 (確率密度関数の性質) (i) $f_X(x) \geq 0$.

(ii) $\int_{-\infty}^{\infty} f_X(x) dx = 1$.

(iii) 任意のボレル集合 $B \in \mathcal{B}(\mathbb{R})$ に対して

$$\mathbb{P}(X \in B) = \int_B f_X(x) dx.$$

証明 定義よりわかる。

□

5 確率変数の期待値

定義 1.15 確率変数 X は確率関数または密度関数 $f_X(\cdot)$ を持つとき、 X の期待値を

$$\mathbb{E}[X] = \begin{cases} \sum_x x f_X(x), & \text{(離散型)} \\ \int_{-\infty}^{\infty} x f_X(x) dx, & \text{(連続型)} \end{cases}$$

で定義する。ただし、離散型の場合は $\sum_x |x| f_X(x) < \infty$ のとき、連続型の場合は $\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty$ のとき、 X の期待値を定義することにする。期待値が定義されるとき、 X の期待値が存在するという。

一般に、 \mathbb{R} 上のボレル可測関数³ $g(\cdot)$ に対し、 $g(X)$ の期待値を

$$\mathbb{E}[g(X)] = \begin{cases} \sum_x g(x)f_X(x), & (\text{離散型}) \\ \int_{-\infty}^{\infty} g(x)f_X(x) dx, & (\text{連続型}) \end{cases}$$

で定める。 $g(X)$ の期待値の存在は X の期待値の存在と同様に定める。

例 1.4 X の確率分布は

X の取る値	0	1	合計
確率	1/2	1/2	1

とする。このとき、

$$\mathbb{E}[X] = 0 \times f_X(0) + 1 \times f_X(1) = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2}$$

となる。

例 1.5 X は確率密度関数

$$f_X(x) = \begin{cases} 1, & (0 \leq x \leq 1) \\ 0, & (\text{その他}) \end{cases}$$

を持つとする。このとき、

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x \cdot 1 dx = \frac{1}{2}$$

例 1.6 X は確率密度関数

$$f_X(x) = \frac{1}{\pi(1+x^2)}$$

を持つとする。このとき、 X の期待値は存在しない。なぜならば、 $M \geq 1$ に対して

$$\mathbb{E}[|X|] \geq \int_1^M x \frac{1}{\pi(1+x^2)} dx \geq \int_1^M \frac{x}{\pi(2x^2)} dx = \frac{1}{2\pi} \int_1^M \frac{dx}{x}$$

よって

$$\mathbb{E}[|X|] \geq \lim_{M \rightarrow \infty} \frac{1}{2\pi} \int_1^M \frac{dx}{x} = \infty$$

よりわかる。

³任意のボレル集合 $B \in \mathcal{B}(\mathbb{R})$ に対し、 $\{x \in \mathbb{R} : g(x) \in B\} \in \mathcal{B}(\mathbb{R})$ が成り立つような関数である。

命題 1.11 (期待値の性質) (i) 定数 c に対して, $\mathbb{E}[c] = c$

(ii) ふたつのボレル可測関数 $h(\cdot)$ と $g(\cdot)$ および定数 a, b に対して

$$\mathbb{E}[ag(X) + bh(X)] = a\mathbb{E}[g(X)] + b\mathbb{E}[h(X)]$$

(iii) $h(x) \geq 0$ ならば, $\mathbb{E}[h(X)] \geq 0$

(iv) $|\mathbb{E}[h(X)]| \leq \mathbb{E}[|h(X)|]$

(v) X が非負値確率変数⁴のとき, $\mathbb{E}[X] = 0$ ならば, $P(X = 0) = 1$ である.

ただし, 上記において, いずれの期待値も存在するものと仮定する.

証明 (i) から (iv) は積分の性質からわかる.

(v) $\mathbb{P}(X > 0) > 0$ と仮定する. 分布関数の右連続性より, ある $\epsilon > 0$ が存在して, $\mathbb{P}(X > \epsilon) > 0$ となる. しかし, $X \geq \epsilon \mathbb{1}\{X > \epsilon\}$ より

$$0 = \mathbb{E}[X] \geq \epsilon \mathbb{P}(X > \epsilon) > 0$$

となり矛盾.

□

6 積率と積率母関数

定義 1.16 正の各整数 n に対して, $\mathbb{E}[|X|^n] < \infty$ のとき,

$$\mu'_n = \mathbb{E}[X^n]$$

を X の n 次の積率という. さらに,

$$\mu_n = \mathbb{E}[(X - \mu)^n]$$

を中心まわりの n 次の積率という. ただし, $\mu = \mathbb{E}[X]$ である.

定義 1.17 X の中心まわりの 2 次の積率を分散といい, $\text{VAR}[X]$ と書く. すなわち, $\text{VAR}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$ である.

⁴ $P(X \geq 0) = 1$ である. 一般には, $\mathbb{E}[X] = 0$ であっても, $P(X = 0) = 1$ ではないことに注意する.

注意 1.9 $\mathbb{E}[X^2] < \infty$ ならば, $\mathbb{E}[|X|] < \infty$ であることに注意せよ. これは Hölder の不等式からわかる. 以下では直接的に確認する. 簡単なために, X を連続型確率変数とし, 確率密度関数 $f_X(x)$ を持つとして議論を進める.

$$\begin{aligned}\mathbb{E}[|X|] &= \int_{-\infty}^{\infty} |x|f_X(x) dx \\ &= \int_{-\infty}^{-1} |x|f_X(x) dx + \int_{-1}^1 |x|f_X(x) dx + \int_1^{\infty} |x|f_X(x) dx\end{aligned}$$

となる. $x \leq -1$ または $x \geq 1$ のときは, $|x|f_X(x) \leq x^2 f_X(x)$ となり, $-1 \leq x \leq 1$ のときは $|x|f_X(x) \leq f_X(x)$ となることに注意すれば,

$$\begin{aligned}\mathbb{E}[|X|] &\leq \int_{-\infty}^{-1} x^2 f_X(x) dx + \int_{-1}^1 f_X(x) dx + \int_1^{\infty} x^2 f_X(x) dx \\ &\leq \int_{-\infty}^{\infty} x^2 f_X(x) dx + \int_{-\infty}^{\infty} f_X(x) dx = \mathbb{E}[X^2] + 1 < \infty\end{aligned}$$

がわかる.

定理 1.2 (分散の性質) (i) $\text{VAR}[a + bX] = b^2 \text{VAR}[X]$. ただし, a, b は定数である.

$$(ii) \quad \text{VAR}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2$$

$$(iii) \quad \text{VAR}[X] = 0 \implies P(X = \mathbb{E}[X]) = 1$$

証明 (i) 分散の定義と命題 1.11(ii) より

$$\begin{aligned}\text{VAR}[a + bX] &= \mathbb{E}\{[a + bX - \mathbb{E}(a + bX)]^2\} = \mathbb{E}\{[b(X - \mathbb{E}(X))]^2\} \\ &= b^2 \mathbb{E}\{(X - \mathbb{E}(X))^2\} = b^2 \text{VAR}[X].\end{aligned}$$

$$(ii) \quad \text{VAR}[X] = \mathbb{E}\{(X - \mathbb{E}(X))^2\} = \mathbb{E}[X^2 - 2\mathbb{E}(X)X + \{\mathbb{E}(X)\}^2] = \mathbb{E}(X^2) - \{\mathbb{E}(X)\}^2.$$

(iii) $\mathbb{E}\{(X - \mathbb{E}(X))^2\} = 0$ と命題 1.11(v) から

$$\mathbb{P}\{(X - \mathbb{E}(X))^2 = 0\} = 1 \iff \mathbb{P}\{X - \mathbb{E}(X) = 0\} = 1.$$

□

定義 1.18 ある正の数 t_0 が存在して、すべての $t \in (-t_0, t_0)$ に対し、 e^{tX} の期待値が存在するならば、 X の積率母関数を

$$M_X(t) = \mathbb{E}[e^{tX}], \quad t \in (-t_0, t_0)$$

で定義する。このような t_0 が存在しないとき、 X の積率母関数は存在しないという。

定理 1.3 (積率母関数の性質) $M_X(t)$ を X の積率母関数とする。このとき、正の整数 n に対して

$$\mathbb{E}[X^n] = \left. \frac{d^n}{dt^n} M_X(t) \right|_{t=0}$$

となる。

証明 X が連続型の場合のみを示す。微分記号と積分記号の交換が可能であると仮定すれば、

$$\begin{aligned} \frac{d}{dt} M_X(t) &= \frac{d}{dt} \int_{-\infty}^{\infty} e^{tx} f_X(x) dx \\ &= \int_{-\infty}^{\infty} \left(\frac{d}{dt} e^{tx} \right) f_X(x) dx \\ &= \int_{-\infty}^{\infty} (x e^{tx} f_X(x)) dx = \mathbb{E}[X e^{tX}] \end{aligned}$$

したがって

$$\left. \frac{d}{dt} M_X(t) \right|_{t=0} = \mathbb{E}[X e^{tX}] \Big|_{t=0} = \mathbb{E}[X]$$

□

例 1.7 X の確率分布は

X の取る値	0	1	合計
確率	1/2	1/2	1

とする。このとき、 $t \in \mathbb{R}$ に対し

$$M_X(t) = \mathbb{E}[e^{tX}] = 1 \times f_X(0) + e^t \times f_X(1) = \frac{1}{2}(1 + e^t)$$

となる．よって

$$\mathbb{E}[X] = \left. \frac{d}{dt} M_X(t) \right|_{t=0} = \left. \frac{1}{2} e^t \right|_{t=0} = \frac{1}{2}$$

となる． X の分散は

$$\text{VAR}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = (0 - \frac{1}{2})^2 \times f_X(0) + (1 - \frac{1}{2})^2 \times f_X(1) = \frac{1}{4}$$

となる．一方，

$$\mathbb{E}[X^2] = \left. \frac{d^2}{dt^2} M_X(t) \right|_{t=0} = \left. \frac{1}{2} e^t \right|_{t=0} = \frac{1}{2}$$

より

$$\text{VAR}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

となる．

命題 1.12 確率変数 X と Y の分布関数を $F_X(\cdot)$ と $F_Y(\cdot)$ とする．ある正の数 t_0 が存在して，すべての $t \in (-t_0, t_0)$ に対し， X と Y の積率母関数 $M_X(t)$ と $M_Y(t)$ が存在して， $M_X(t) = M_Y(t)$ ならば，すべての $x \in \mathbb{R}$ に対して $F_X(x) = F_Y(x)$ である．

証明 証明は略．

□

7 確率変数の変換

確率変数 X は分布関数 $F_X(\cdot)$ を持つとする． $g(\cdot)$ をボレル可測関数としたとき， $Y = g(X)$ も確率変数になり， Y の分布 P_Y はつぎのように定まる：任意のボレル集合 $A \in \mathcal{B}(\mathbb{R})$ に対して

$$P_Y(A) = \mathbb{P}(Y \in A) = \mathbb{P}(g(X) \in A)$$

また， Y の分布関数は

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(g(X) \leq y)$$

となる． P_Y は P_X と g に依存するので， $F_Y(\cdot)$ は $F_X(\cdot)$ と g に依存するのである．さらに，存在するならば， $F_Y(\cdot)$ の確率密度関数 $f_Y(\cdot)$ も求めることができれば便利である．

以下の例では， $F_X(\cdot)$ は微分可能とする．

例 1.8 X の線形変換を考える．すなわち， $Y = a + bX$ である．ここで， a, b は定数で $b \neq 0$ とする． $b > 0$ のとき

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(a + bX \leq y) = \mathbb{P}\left(X \leq \frac{y-a}{b}\right) = F_X\left(\frac{y-a}{b}\right)$$

また， $b < 0$ のとき

$$F(y) = \mathbb{P}(Y \leq y) = \mathbb{P}\left(X \geq \frac{y-a}{b}\right) = 1 - \mathbb{P}\left(X \leq \frac{y-a}{b}\right) = 1 - F_X\left(\frac{y-a}{b}\right)$$

となる．したがって

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{1}{|b|}f_X\left(\frac{y-a}{b}\right)$$

となる．線形関数は狭義単調関数である．

例 1.9 X を正值確率変数とする．すなわち， $\mathbb{P}(X \leq 0) = 0$ である． $Y = \log X$ とおく．このとき， $y \in \mathbb{R}$ に対して

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X \leq e^y) = F_X(e^y)$$

となる．したがって

$$f_Y(y) = \frac{d}{dy}F_Y(y) = e^y f_X(e^y)$$

となる． $g(x) = \log x$ も \mathbb{R}^+ 上では狭義単調関数である．

例 1.10 $Y = X^2$ とおく．これは \mathbb{R} 上の単調関数ではない．しかし， \mathbb{R}^+ 上と \mathbb{R}^- 上では単調である． $y > 0$ に対して，

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(\omega \in \Omega : (X(\omega))^2 \leq y) \\ &= \mathbb{P}(\omega \in \Omega : (X(\omega))^2 \leq y, X(\omega) > 0) \\ &\quad + \mathbb{P}(\omega \in \Omega : (X(\omega))^2 \leq y, X(\omega) < 0) + \mathbb{P}(\omega \in \Omega : X(\omega) = 0) \\ &= \mathbb{P}(0 < X < \sqrt{y}) + \mathbb{P}(-\sqrt{y} \leq X < 0) + \mathbb{P}(X = 0) \\ &= \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \end{aligned}$$

となる．したがって，

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y})$$

となる． $y \leq 0$ に対しては $F_Y(y) = 0$ となるので， $f_Y(y) = 0$ となることは容易にわかる．よって

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y}), & (y > 0), \\ 0, & (\text{その他}) \end{cases}$$

となる．

これらの例を踏まえて，一般の $Y = g(X)$ の場合を考える．そのために

$$\mathcal{X} = \{x : f_X(x) > 0\}, \quad \mathcal{Y} = \{y : y = g(x), x \in \mathcal{X}\}$$

とおく． \mathcal{X} と \mathcal{Y} を X と Y の台とよぶ．たとえば， $Y = X^2$ のとき， $\mathcal{X} = \mathbb{R}$ ならば， $\mathcal{Y} = [0, \infty)$ となる．

いま，関数 $g : \mathcal{X} \rightarrow \mathcal{Y}$ を考える．もし， g が一対一上への写像ならば，逆写像 $g^{-1} : \mathcal{Y} \rightarrow \mathcal{X}$ は定義できる．しかし，一般には，逆像で定義する．すなわち，

$$g^{-1}(y) = \{x \in \mathcal{X} : g(x) = y\}$$

さらに，ボレル集合 A に対し

$$g^{-1}(A) = \{x \in \mathcal{X} : g(x) \in A\}$$

とする．すると

$$\mathbb{P}(Y \in A) = \mathbb{P}(\omega \in \Omega : g(X(\omega)) \in A) = \mathbb{P}(X \in g^{-1}(A))$$

となる．しかし， Y の確率密度関数を求めるためには，追加の条件が必要となる．

命題 1.13 X を連続型確率変数とし， f_X をその確率密度関数とし， \mathcal{X} 上で連続とする．

(i) 逆関数 $g^{-1}(\cdot)$ は \mathcal{Y} 上で定義され, 連続微分可能とする. このとき

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|, & (y \in \mathcal{Y}), \\ 0, & (\text{その他}) \end{cases}$$

となる.

(ii) $\{\mathcal{X}_i\}_{i=0}^k$ を \mathcal{X} の分割⁵とし, $\mathbb{P}(X \in \mathcal{X}_0) = 0$ とし, $f_X(\cdot)$ は各 $\mathcal{X}_i, i = 1, 2, \dots, k$ 上で連続とする. さらに, 各 $\mathcal{X}_i, i = 1, 2, \dots, k$, 上で定義された関数 $g_i(\cdot)$ が存在して, $x \in \mathcal{X}_i$ に対し, $g(x) = g_i(x)$ が成立し, $g_i(x)$ は \mathcal{X}_i 上で狭義単調関数とし, $g_i^{-1}(\cdot)$ は $\mathcal{Y} \setminus g(\mathcal{X}_0)$ 上で定義され, 連続微分可能とする. このとき,

$$f_Y(y) = \begin{cases} \sum_{i=1}^k f_X(g_i^{-1}(y)) \left| \frac{d}{dy} g_i^{-1}(y) \right|, & (y \in \mathcal{Y} \setminus g(\mathcal{X}_0)), \\ 0, & (\text{その他}) \end{cases}$$

となる.

証明 (i) g が狭義単調増加のとき, $y \in \mathcal{Y}$ に対し,

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X \leq g^{-1}(y)) = F_X(g^{-1}(y))$$

となる. よって

$$f_Y(y) = \frac{d}{dy} F_Y(y) = f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$$

となる. また, g が狭義単調減少のとき,

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y))$$

となる. よって,

$$f_Y(y) = \frac{d}{dy} F_Y(y) = f_X(g^{-1}(y)) \left(-\frac{d}{dy} g^{-1}(y) \right)$$

となる. g が単調増加のときは $(d/dy)g^{-1}(y) > 0$, g が単調減少のときは $(d/dy)g^{-1}(y) < 0$ に注意すればよい.

⁵すなわち, $\mathcal{X}_i \cap \mathcal{X}_j = \emptyset (i \neq j)$ かつ $\mathcal{X} = \cup_{i=0}^k \mathcal{X}_i$

(ii) $y \in \mathcal{Y}$ に対し,

$$F_Y(y) = \mathbb{P}(Y \leq y) = \sum_{i=1}^k \mathbb{P}(g(X) \leq y, X \in \mathcal{X}_i) = \sum_{i=1}^k \mathbb{P}(X \in g_i^{-1}((-\infty, y]))$$

となる. 最右辺の各項に対して, (i) の議論を適用すれば, (ii) は証明される. \square

例 1.11 確率変数 X は確率密度関数

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad -\infty < x < \infty$$

を持つとする. $Y = X^2$ としたとき, Y の確率密度関数を求めよう. 関数 $g(x) = x^2$ は $(-\infty, 0)$ 上と $(0, \infty)$ 上のそれぞれで単調関数である. $\mathcal{Y} = (0, \infty)$ として, 命題 1.13 を利用するために,

$$\begin{aligned} \mathcal{X}_0 &= \{0\}, \\ \mathcal{X}_1 &= (-\infty, 0), & g_1(x) &= x^2, & g_1^{-1}(y) &= -\sqrt{y} \\ \mathcal{X}_2 &= (0, \infty), & g_2(x) &= x^2, & g_2^{-1}(y) &= \sqrt{y} \end{aligned}$$

とおく. Y の確率密度関数は

$$\begin{aligned} f_Y(y) &= \frac{1}{\sqrt{2\pi}} e^{-(-\sqrt{y})^2/2} \left| -\frac{1}{2\sqrt{y}} \right| + \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{y})^2/2} \left| \frac{1}{2\sqrt{y}} \right| \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{y}} e^{-y/2}, \quad 0 < y < \infty \end{aligned}$$

となる.

例 1.12 確率変数 X は連続で分布関数 $F_X(\cdot)$ を持つとし, $Y = F_X(X)$ とおく. このとき, Y は確率密度関数

$$f_Y(y) = \begin{cases} 1 & (0 < y < 1), \\ 0 & (\text{その他}) \end{cases} \quad (1.2)$$

を持つ.

$F_X(x)$ は x に関して狭義単調増加として証明⁶をする. $0 \leq F_X(x) \leq 1$ であり, $\mathbb{P}(F_X(X) \leq 0) = 0$ と $\mathbb{P}(F_X(X) \geq 1) = 0$ となるので, $y \leq 0$ のときは

⁶この命題は $F_X(x)$ が x に関して単調増加のときに成立するが, $F_X^{-1}(x)$ の定義と証明やや難しくなるので簡単な場合のみの証明を与える.

$\mathbb{P}(Y \leq y) = 0$ である。 $0 < y < 1$ に対して,

$$\begin{aligned} \mathbb{P}(Y \leq y) &= \mathbb{P}(F_X(X) \leq y) \\ &= \mathbb{P}(F_X^{-1}(F_X(X)) \leq F_X^{-1}(y)) \\ &= \mathbb{P}(X \leq F_X^{-1}(y)) \\ &= F_X(F_X^{-1}(y)) \\ &= y \end{aligned}$$

となる。 $y \geq 1$ に対しては $\mathbb{P}(Y \leq y) = 1$ となる。したがって,

$$F_Y(y) = \begin{cases} 0, & (y \leq 0), \\ y, & (0 < y < 1), \\ 1, & (y \geq 1) \end{cases}$$

となる。 $0 < y < 1$ 上で $F_Y(y)$ は微分可能で導関数は 1 となる。また, $y = 0$ における $F_Y(y)$ の導関数を左微分で, $y = 1$ における $F_Y(y)$ の導関数を右微分で定めると Y の確率密度関数は (1.2) となることがわかる。

8 確率変数の不等式 1

定理 1.4 (マルコフの不等式) 非負値確率変数 X が有限の期待値 $\mathbb{E}[X] < \infty$ をもつとき, 任意の正数 $a > 0$ に対して,

$$\mathbb{P}\{X \geq a\} \leq \frac{1}{a} \mathbb{E}[X]$$

が成立する。

証明 X が連続型確率変数の場合を示す。 X の確率密度関数を $f_X(x)$ とする。

$$\mathbb{E}[X] = \int_0^{\infty} x f_X(x) dx \geq \int_a^{\infty} x f_X(x) dx \geq a \int_a^{\infty} f_X(x) dx = a \mathbb{P}(X \geq a)$$

よりわかる。

□

定理 1.5 (チェビシエフの不等式) 確率変数 X が平均 μ , 分散 σ^2 ($0 < \sigma^2 < \infty$) をもつとき, 任意の正数 $a > 0$ に対して,

$$\mathbb{P}\{|X - \mu| \geq a\} \leq \frac{\sigma^2}{a^2}$$

が成り立つ.

証明 Markov の不等式から

$$\mathbb{P}\{|X - \mu| \geq a\} = \mathbb{P}\{|X - \mu|^2 \geq a^2\} \geq \frac{\mathbb{E}[(X - \mu)^2]}{a^2}$$

よりわかる.

□

例 1.13 X は平均 μ , 分散 σ^2 ($\sigma > 0$) の分布に従うとする. このとき, 任意の正の実数 t に対して,

$$\mathbb{P}\{|X - \mu| \geq t\sigma\} \leq \frac{1}{t^2} \mathbb{E}\left[\left(\frac{|X - \mu|}{\sigma}\right)^2\right] = \frac{1}{t^2} \frac{\mathbb{E}[(X - \mu)^2]}{\sigma^2} = \frac{1}{t^2}$$

となる. したがって, $t = 2$ の場合

$$\mathbb{P}\{|X - \mu| \geq 2\sigma\} \leq \frac{1}{2^2} = 0.25$$

となる.

例 1.14 Z は標準正規分布に従うとする. このとき, 任意の正の実数 t に対して,

$$\mathbb{P}\{|Z| \geq t\} \leq \sqrt{\frac{2}{\pi}} \frac{e^{-t^2/2}}{t} \quad (1.3)$$

となる.

$t = 2$ のとき, Chebychev の不等式から

$$\mathbb{P}\{|Z| \geq 2\} \leq \frac{1}{2^2} = 0.25$$

となる. しかし, (1.3) から

$$\mathbb{P}\{|Z| \geq 2\} \leq \sqrt{\frac{2}{\pi}} \frac{e^{-2}}{2} = 0.054$$

となる．また，

$$\mathbb{P}\{|Z| \geq 3\} \leq \sqrt{\frac{2}{\pi}} \frac{e^{-4.5}}{3} = 0.00295$$

となる．

(1.3) は以下からわかる．

$$\begin{aligned} \mathbb{P}\{Z \geq t\} &= \frac{1}{\sqrt{2\pi}} \int_t^{\infty} e^{-z^2/2} dz \\ &\leq \frac{1}{\sqrt{2\pi}} \int_t^{\infty} \frac{z}{t} e^{-z^2/2} dz \\ &= \frac{1}{\sqrt{2\pi}} \left[-\frac{1}{t} e^{-z^2/2} \right]_t^{\infty} = \frac{1}{\sqrt{2\pi}} \frac{e^{-t^2/2}}{t} \end{aligned}$$

と

$$\mathbb{P}\{|Z| \geq t\} = \mathbb{P}\{\{Z \geq t\} \cup \{Z \leq -t\}\} = \mathbb{P}\{Z \geq t\} + \mathbb{P}\{Z \leq -t\} = 2\mathbb{P}\{Z \geq t\}$$

からわかる．

ある区間 $I = (a, b)$ 上の実数値連続関数 $h(x)$ が凸 (convex) であるとは，任意の $c \in (0, 1)$ と任意の $x_1, x_2 \in I$ に対して，

$$h(cx_1 + (1-c)x_2) \leq ch(x_1) + (1-c)h(x_2)$$

が成り立つことである．もし， $x_1 \neq x_2$ に対して「 \leq 」のかわりに「 $<$ 」が常に成立するならば，狭義の凸関数という．

定理 1.6 (イェンセンの不等式) X を確率変数とし， $h(x)$ を X の値域を含む区間上で凸な関数とする． X と $h(X)$ の期待値が有限のとき，

$$\mathbb{E}[h(X)] \geq h(\mathbb{E}[X])$$

が成立する．もし， $h(x)$ が狭義の凸関数のとき，等号が成り立つのは 1 点分布の時に限られる．

証明 任意の固定した x_0 とある線形関数 $g(x) = ax + b$ が存在して， $h(x_0) = g(x_0)$ とすべての x に対して $h(x) \geq g(x)$ が成立する． $x_0 = \mathbb{E}[X]$ として，上

のことに利用すれば,

$$\begin{aligned}
 \mathbb{E}[h(X)] &\geq \mathbb{E}[g(X)] \\
 &= \mathbb{E}[aX + b] \\
 &= a\mathbb{E}[X] + b \\
 &= g(\mathbb{E}[X]) = h(\mathbb{E}[X])
 \end{aligned}$$

□

例 1.15 X を $\mathbb{E}[X^2] < \infty$ なる確率変数とする. Jensen の不等式を利用するために, まず, $\mathbb{E}[|X|] < \infty$ を確認する. たとえば, X が確率密度関数 $f_X(x)$ を持てば

$$\begin{aligned}
 \mathbb{E}[|X|] &= \int_{-\infty}^{\infty} |x| f_X(x) dx \\
 &= \int_{-\infty}^{-1} |x| f_X(x) dx + \int_{-1}^1 |x| f_X(x) dx + \int_1^{\infty} |x| f_X(x) dx \\
 &\leq \int_{-\infty}^{-1} x^2 f_X(x) dx + \int_{-1}^1 f_X(x) dx + \int_1^{\infty} x^2 f_X(x) dx \\
 &\leq \int_{-\infty}^{\infty} |x|^2 f_X(x) dx + \int_{-\infty}^{\infty} f_X(x) dx = \mathbb{E}[X^2] + 1 < \infty
 \end{aligned}$$

となる. X が離散型確率変数のときも同様にすれば, $\mathbb{E}[|X|] \leq \mathbb{E}[X^2] + 1$ であることがわかる. さらに, $\mathbb{E}[|X|]$ の評価が Jensen の不等式から得られる. $h(x) = x^2$ とすれば, $h(x)$ は狭義の凸関数なので, Jensen の不等式より

$$\{\mathbb{E}[X]\}^2 = h(\mathbb{E}[X]) \leq \mathbb{E}[h(X)] = \mathbb{E}[X^2]$$

となる. したがって,

$$|\mathbb{E}[X]| \leq \sqrt{\mathbb{E}[X^2]}$$

をえる.

例 1.16 a_1, a_2, \dots, a_n は正の数とし,

$$\begin{aligned} m_A &= \frac{1}{n}(a_1 + a_2 + \dots + a_n) && (\text{算術平均}), \\ m_G &= \{a_1 a_2 \dots a_n\}^{1/n} && (\text{幾何平均}), \\ m_H &= \frac{1}{\frac{1}{n} \left(\frac{1}{a_1} + \frac{1}{a_2} + \dots + \frac{1}{a_n} \right)} && (\text{調和平均}) \end{aligned}$$

とする. このとき,

$$m_H \leq m_G \leq m_A$$

が成立する.

Jensen の不等式を用いて示す: そのために, X を確率変数として

$$\mathbb{P}\{X = a_i\} = \frac{1}{n}, \quad i = 1, 2, \dots, n$$

とする. $-\log x$ は凸関数なので, $\mathbb{E}[\log X] \leq \log(\mathbb{E}[X])$ が成立する. したがって,

$$\log m_G = \frac{1}{n} \sum_{i=1}^n \log a_i = \mathbb{E}[\log X] \leq \log(\mathbb{E}[X]) = \log \left(\frac{1}{n} \sum_{i=1}^n a_i \right) = \log m_A$$

より $m_G \leq m_A$ がわかる. また,

$$\begin{aligned} \log \frac{1}{m_H} &= \log \left(\frac{1}{n} \sum_{i=1}^n \frac{1}{a_i} \right) = \log \left(\mathbb{E} \left[\frac{1}{X} \right] \right) \geq \mathbb{E} \left[\log \left(\frac{1}{X} \right) \right] \\ &= -\mathbb{E}[\log X] = -\log m_G \end{aligned}$$

から $m_G \geq m_H$ がわかる.

9 演習問題

問題 1.1 ジョーカーを除く 52 枚のトランプからカードを 1 枚無作為に抜き取る試行をし, 組み札に注目するとする: クラブ (C), ダイヤ (D), ハート (H), スペード (S). したがって, 標本空間は

$$\Omega = \{C, D, H, S\}$$

である. 事象 A, B を

$$A = \{C, D\}, \quad B = \{D, H, S\}$$

としたとき, つぎの事象を求めよ.

- (1) $A \cup B$
- (2) $A \cap B$
- (3) A^c

ヒント: $A \subset B \iff$ 任意の $\omega \in A$ に対し, $\omega \in B$
 $A = B \iff A \subset B$ かつ $B \subset A$

問題 1.2 つぎの等式を示せ. このとき,

- (1) $A \setminus B = A \setminus (A \cap B)$
- (2) $B = (B \cap A) \cup (B \cap A^c)$

ヒント: 事象 A に含まれる根元事象であって事象 B には含まれないものの全体のつくる事象を, A, B の差といい, $A \setminus B$ と記すことにする:

$$A \setminus B = \{\omega \in \Omega : \omega \in A \text{ かつ } \omega \notin B\}$$

問題 1.3 $a, b \in \mathbb{R} (a < b)$ とする. \mathbb{R} の部分集合に関するつぎの等式を確かめよ.

- (1) $\bigcap_{n=1}^{\infty} (a - 1/n, b] = [a, b]$
- (2) $\bigcup_{n=1}^{\infty} (a, b - 1/n] = (a, b)$
- (3) $\bigcap_{n=1}^{\infty} (a - 1/n, a] = \{a\}$
- (4) $\bigcup_{n=1}^{\infty} (-n, a] = (-\infty, a]$
- (5) $\bigcup_{n=-\infty}^{\infty} (n, n + 1] = \mathbb{R}$

ヒント: (1) の等式については, (左辺) \subset (右辺) は明らか. (左辺) \supset (右辺) を示すためには, 対偶をとり, $x \notin [a, b]$ ならば, $x \notin (a - 1/n, b]$ を言えばよい.

問題 1.4 (Ω, \mathcal{F}) を可測空間とする.

- (1) $A, B \in \mathcal{F}$ のとき, $A \cap B \in \mathcal{F}$ と $A \setminus B \in \mathcal{F}$ を示せ.
- (2) $A_1, A_2, \dots \in \mathcal{F}$ のとき, $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$ を示せ.

ヒント: $\{\bigcup_{i=1}^{\infty} A_i^c\}^c = \bigcap_{i=1}^{\infty} A_i$ を利用せよ.

問題 1.5 事象 A, B が互いに独立のとき, つぎの対も互いに独立であることを示せ.

- (1) A^c, B^c
- (2) A, B^c
- (3) A^c, B

ヒント: $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ を用いて, たとえば, 確率の性質 (iii) と (v) を利用して $\mathbb{P}(A^c \cap B^c) = \mathbb{P}(A^c)\mathbb{P}(B^c)$ を示せばよい. 計算は右辺から始めた方が見通しがいいようだ.

問題 1.6 大小 2 つのサイコロを投げる試行を考える. したがって, 標本空間は

$$\Omega = \{(1, 1), (1, 2), \dots, (1, 6), (2, 1), \dots, (2, 6), \dots, (6, 1), \dots, (6, 6)\}$$

である。つぎの事象を考えよう。

$$\begin{aligned} A &= \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\} \\ B &= \{\text{ふたつのサイコロの和が } 7 \text{ 以上 } 10 \text{ 以下}\} \\ C &= \{\text{ふたつのサイコロの和が } 2 \text{ または } 7 \text{ または } 8\} \end{aligned}$$

- (1) $\mathbb{P}(A) = 1/6$, $\mathbb{P}(B) = 1/2$, $\mathbb{P}(C) = 1/3$ を確認せよ。
- (2) $A \cap B \cap C$ を求め, $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$ を示せ。
- (3) $\mathbb{P}(B \cap C) \neq \mathbb{P}(B)\mathbb{P}(C)$ と $\mathbb{P}(A \cap B) \neq \mathbb{P}(A)\mathbb{P}(B)$ を確認せよ。

問題 1.7 標本空間 Ω を 3 つの字 a, b, c を並べたものとする。

$$\Omega = \{aaa, abc, acb, bbb, bca, bac, ccc, cba, cab\}$$

ここからどのならばが出現する確率も $1/9$ であるとし, 事象 A_i , $i = 1, 2, 3$, を

$$A_i = \{\text{字の並びの } i \text{ 番目の字が } a\}$$

とする。

- (1) 事象 A_1, A_2, A_3 を求め, $\mathbb{P}(A_i) = 1/3$ を確認せよ。
- (2) A_1, A_2, A_3 は対独立であることを示せ。
- (3) A_1, A_2, A_3 は独立でないことを確かめよ。

問題 1.8 ふたつの事象 A, B に対して, $\mathbb{P}(A) > 0$ ならば

$$\mathbb{P}(B|A) \geq 1 - \left\{ \frac{\mathbb{P}(B^c)}{\mathbb{P}(A)} \right\}$$

が成り立つことを示せ。

ヒント :

$$\begin{aligned} A \cup B &= (A \cap B^c) \cup (A^c \cap B) \cup (A \cap B) \\ A &= (A \cap B^c) \cup (A \cap B) \\ B &= (A^c \cap B) \cup (A \cap B) \end{aligned}$$

のような互いに排反な事象に分割できることを利用するとよい。

問題 1.9 A, B, C を事象としたとき, 以下を示せ。ただし, 条件付けられた事象の確率はすべて正とする。

- (1) $\mathbb{P}(B) = 1$ ならば, すべての事象 A に対して,

$$\mathbb{P}(A|B) = \mathbb{P}(A)$$

- (2) $A \subset B$ ならば,

$$\mathbb{P}(B|A) = 1$$

かつ

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A)}{\mathbb{P}(B)}$$

(3) A と B は互いに排反ならば

$$\mathbb{P}(A|A \cup B) = \frac{\mathbb{P}(A)}{\mathbb{P}(A) + \mathbb{P}(B)}$$

(4)

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A|B \cap C)\mathbb{P}(B|C)\mathbb{P}(C)$$

ヒント：(i) たとえば、 $B \subset A \cup B$ と $\mathbb{P}(B) = 1$ に注意し、命題 1.4(vi) を用いて、 $\mathbb{P}(A \cup B) = 1$ を示す。さらに、命題 1.4(v) を利用して、 $\mathbb{P}(A \cap B) = \mathbb{P}(A)$ を示せばよい。

問題 1.10 ふたつの事象 A, B に対して

$$\begin{aligned}\mathbb{P}(A|B) &= \frac{\mathbb{P}(A)\mathbb{P}(B|A)}{\mathbb{P}(A)\mathbb{P}(B|A) + \mathbb{P}(A^c)\mathbb{P}(B|A^c)} \\ \mathbb{P}(A^c|B) &= \frac{\mathbb{P}(A^c)\mathbb{P}(B|A^c)}{\mathbb{P}(A)\mathbb{P}(B|A) + \mathbb{P}(A^c)\mathbb{P}(B|A^c)}\end{aligned}$$

を示せ。ただし、条件付けられた事象の確率はすべて正とする。

問題 1.11 3つの箱がある：箱 1 には赤玉が 1つと白玉が 3つ、箱 2 には赤玉が 2つと白玉が 2つ、箱 3 には赤玉が 3つと白玉が 1つはっている。箱を 1つ無作為に選び、選ばれた箱から玉を 1つ取る。赤玉が取られたときに、箱 1 が選ばれた条件付き確率を求めよ。

ヒント：事象 $A_i, i = 1, 2, 3$ を「箱 i を選ぶ」とし、事象 B を「赤玉が取られる」とし、 B^c を「白玉が取られる」とおいてとき、 $\mathbb{P}(A_1|B)$ を求めればよいことになる。 $\mathbb{P}(A_i)$ および $\mathbb{P}(B|A_i)$ の確率は題意から簡単に計算できるので、ベイズの定理を使えばよい。

問題 1.12 正しい硬貨を 2 回投げる試行を考える：表を H 、裏を T と書けば、

$$\Omega = \{HH, HT, TH, TT\}$$

となる \mathcal{F} は Ω のすべての部分集合（ 2^4 個の事象から構成される）とする。確率変数 X を表の出た回数とする：

$$X(\omega) = \begin{cases} 0, & (\omega = TT) \\ 1, & (\omega = HT, TH) \\ 2, & (\omega = HH) \end{cases}$$

である。このとき、 X の分布関数を求め、そのグラフを掛け。

問題 1.13

$$F_X(x) = \frac{1}{1 + e^{-x}}$$

は命題 1.8 で与えられた分布関数の条件をみたしていることを確認し、そのグラフを掛け。

問題 1.14 確率変数 X は分布関数

$$F_X(x) = \begin{cases} 0 & (x < 0), \\ \frac{1}{2}x & (0 \leq x < 2), \\ 1 & (x \geq 2) \end{cases}$$

をもつとする。このとき、以下の問いに答えよ。

- (1) $F_X(x)$ のグラフを描き, 命題 1.8 で与えられた分布関数の条件をみたまことを確認せよ.
- (2) $P_X((1/2, 3/2]) = \mathbb{P}\left(\frac{1}{2} < X \leq \frac{3}{2}\right)$ を求めよ.

問題 1.15 つぎの関数を考える:

$$f_X(x) = \begin{cases} \frac{1}{x(x+1)}, & x = 1, 2, 3, \dots \\ 0, & (\text{その他}) \end{cases}$$

- (1) $f_X(\cdot)$ は命題 1.9 で与えられた確率関数の条件をみたまことを示せ.
- (2) 離散型確率変数 X が確率関数 $f_X(x)$ を持つとき, $\mathbb{P}(4 \leq X \leq 7)$ を求めよ.

問題 1.16

$$f_X(x) = \begin{cases} \frac{x^2(1-x)}{a}, & (0 < x < 1), \\ 0, & (\text{その他}) \end{cases}$$

とする. ただし, $a > 0$ とする.

- (1) $f_X(x)$ が確率密度関数となるように a を定めよ.
- (2) 連続型確率変数 X が (1) で求めた確率密度関数を持つとき, その分布関数を求めよ.
- (3) 連続型確率変数 X が (1) で求めた確率密度関数を持つとき, $\mathbb{P}(X \leq 1/2)$ を求めよ.

問題 1.17 $F_X(\cdot)$ を連続型確率変数 X の分布関数とし, $f_X(\cdot)$ をその確率密度関数とし, x_0 を $F_X(x_0) < 1$ なる固定された点とする. このとき,

$$g(x) = \begin{cases} f_X(x)/[1 - F_X(x_0)] & x \geq x_0 \\ 0 & x < x_0 \end{cases}$$

は命題 1.10 で与えられた確率密度関数の条件をみたまことを示せ.

問題 1.18 離散型確率変数 X は確率関数

$$f_X(x) = \begin{cases} \frac{1}{2} & (x = -1, 1) \\ 0 & (\text{その他}) \end{cases}$$

を持つとする.

- (i) X の分布関数 $F_X(x) = \mathbb{P}(X \leq x)$ を求めよ.
- (ii) X と X^2 の期待値 $\mathbb{E}[X]$ と $\mathbb{E}[X^2]$ を求めよ.

問題 1.19 連続型確率変数 X は確率密度関数

$$f_X(x) = \begin{cases} \frac{1}{2} & (-1 \leq x \leq 1) \\ 0 & (\text{その他}) \end{cases}$$

を持つとする.

- (i) X の分布関数 $F_X(x) = \mathbb{P}(X \leq x)$ を求めよ。
 (ii) X と X^2 の期待値 $\mathbb{E}[X]$ と $\mathbb{E}[X^2]$ を求めよ。

問題 1.20 連続型確率変数 X は確率密度関数

$$f_X(x) = \begin{cases} e^{-x} & (x > 0) \\ 0 & (\text{その他}) \end{cases}$$

を持つとする。

- (i) X の分布関数 $F_X(x) = \mathbb{P}(X \leq x)$ を求めよ。
 (ii) X と X^2 の期待値 $\mathbb{E}[X]$ と $\mathbb{E}[X^2]$ を求めよ。

問題 1.21 X を連続型確率変数とし、 a, b を定数とする。

- (i) $\mathbb{P}(X \geq a) = 1$ ならば、 $\mathbb{E}[X] \geq a$ を示せ。
 (ii) $\mathbb{P}(X \leq b) = 1$ ならば、 $\mathbb{E}[X] \leq b$ を示せ。

問題 1.22 X を離散型確率変数とし、確率関数

$$f_X(x) = \begin{cases} \frac{x}{6} & (x = 1, 2, 3), \\ 0 & (\text{その他}) \end{cases}$$

をもつとする。

- (i) $\mathbb{E}[X], \mathbb{E}[X^2]$ を求めよ。
 (ii) $\mathbb{E}[(X - \mathbb{E}[X])^2]$ を求めよ。
 (iii) $Y = (X - 1)^2$ としたとき、 $\mathbb{E}[Y]$ と $\mathbb{E}[(Y - \mathbb{E}[Y])^2]$ を求めよ。
 (iv) X の分布関数を求め、グラフに描け。

問題 1.23 X を連続型確率変数とし、確率密度関数

$$f_X(x) = \begin{cases} 6x(1-x) & (0 < x < 1), \\ 0 & (\text{その他}) \end{cases}$$

をもつとする。

- (i) $\mathbb{E}[X], \mathbb{E}[X^2]$ を求めよ。
 (ii) $\mathbb{E}[(X - \mathbb{E}[X])^2]$ を求めよ。
 (iii) $Y = -3X + 10$ としたとき、 $\mathbb{E}[Y]$ と $\mathbb{E}[(Y - \mathbb{E}[Y])^2]$ を求めよ。
 (iv) X の分布関数を求め、グラフに描け。

問題 1.24 X を連続型確率変数で $\mathbb{E}[X^2] < \infty$ とし、 a, b を定数とする。このとき、以下を示せ。

- (i) $\text{VAR}[a] = 0$ 。
 (ii) $\text{VAR}[aX + b] = a^2 \text{VAR}[X]$
 (iii) $\text{VAR}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2$

問題 1.25 X は連続型確率変数で $\mathbb{E}[X^2] < \infty$ とする. $g(t) = \mathbb{E}[(X - t)^2]$ は $t = \mathbb{E}[X]$ で最小となることを示せ. すなわち,

$$g(t) \geq \text{VAR}[X], \quad t \in \mathbb{R}$$

が成り立つ.

問題 1.26 連続型確率変数 X は確率密度関数

$$f_X(x) = \begin{cases} 1 & (0 < x < 1), \\ 0 & (\text{その他}), \end{cases}$$

をもつとする. このとき, 以下の問いに答えよ.

- (i) X の分布関数 $F_X(x) = \mathbb{P}(X \leq x)$ ($x \in \mathbb{R}$) と X の期待値 $\mathbb{E}[X]$ を求めよ. ただし, $F_X(x)$ は \mathbb{R} 上の関数であることがわかるように答えよ.
- (ii) $Y = -2 \log X$ としたとき, Y の分布関数 $F_Y(y) = \mathbb{P}(Y \leq y)$ ($y \in \mathbb{R}$) と確率密度関数 $f_Y(y)$ を求めよ. ただし, $F_Y(y)$ は \mathbb{R} 上の関数であることがわかるように答えよ. また, $f_Y(y) > 0$ となる y の範囲を明示すること.
- (iii) Y の期待値 $\mathbb{E}[Y]$ を求めよ.

問題 1.27 連続型確率変数 X は確率密度関数

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad (-\infty < x < \infty)$$

を持つとする. ただし, μ, σ は定数で $-\infty < \mu < \infty, 0 < \sigma < \infty$ とし, $\exp(x) = e^x$ である.

$$Z = \frac{X - \mu}{\sigma}$$

と定義したとき, 以下の問いに答えよ.

- (i) Z の確率密度関数 $f_Z(z)$ を求めよ.
- (ii) $10Z + 50$ の期待値 $\mathbb{E}[10Z + 50]$ と分散 $\text{VAR}[10Z + 50]$ を求めよ.

問題 1.28 確率変数 X は確率密度関数

$$f_X(x) = \begin{cases} \frac{1}{2}, & -1 < x < 1, \\ 0 & (\text{その他}) \end{cases}$$

を持つとする.

- (i) X の分布関数を求め, そのグラフを描け.
- (ii) $\mathbb{E}[X]$ と $\mathbb{E}[X^2]$ および $\text{VAR}[X]$ を求めよ.
- (iii) $Y = X^2$ としたとき, Y の確率密度関数 $f_Y(y)$ を求めよ. さらに,

$$\int_{-\infty}^{\infty} f_Y(y) dy = 1$$

を確認せよ.

(iv) $\mathbb{E}[Y]$ を Y の確率密度関数を用いて計算せよ.

問題 1.29 確率変数 X は確率密度関数

$$f_X(x) = \begin{cases} e^{-x}, & x > 0, \\ 0 & (\text{その他}) \end{cases}$$

を持つとする.

- (i) X の分布関数を求め、そのグラフを描け.
- (ii) $\mathbb{E}[X]$ と $\mathbb{E}[X^2]$ および $\text{VAR}[X]$ を求めよ.
- (iii) $Y = X^3$ としたとき、 Y の分布関数を求めよ.
- (iv) Y の確率密度関数を (iii) で求めた分布関数を微分することで求めよ.
- (v) Y の確率密度関数を命題 1.13 を用いて求めよ.
- (vi) $Z = \log X$ としたとき、 Z の分布関数を求めよ.
- (vii) Z の確率密度関数を (vi) で求めた分布関数を微分することで求めよ.
- (viii) Z の確率密度関数を命題 1.13 を用いて求めよ.
- (ix) $W = e^X$ としたとき、 W の分布関数を求めよ.
- (x) W の確率密度関数を (ix) で求めた分布関数を微分することで求めよ.
- (xi) W の確率密度関数を命題 1.13 を用いて求めよ.

問題 1.30 離散型確率変数 X は確率関数

$$f_X(x) = \begin{cases} \frac{1}{6}, & x = 1, 2, 3, 4, 5, 6, \\ 0 & (\text{その他}) \end{cases}$$

を持つとし、 $Y = (X - 3)^2$ とおく.

- (i) $\mathbb{E}[X]$ と $\mathbb{E}[X^2]$ を求めよ.
- (ii) Y の確率関数 $f_Y(y)$ を求めよ. さらに

$$\sum_y f_Y(y) = 1$$

を確認せよ.

- (iii) $\mathbb{E}[Y]$ を (ii) で求めた Y の確率関数を利用して求めよ.
- (iv) $\mathbb{E}[Y]$ を期待値の性質と (i) で求めた X の 2 次までの積率を利用して求めよ.

問題 1.31 離散型確率変数 X は確率関数

$$f_X(x) = \begin{cases} \frac{1}{3} & (x = -1, 0, 1), \\ 0 & \text{その他} \end{cases}$$

を持つとする.

- (i) X の分布関数 $F_X(x)$ のグラフを描け.

(ii) X の平均 $\mathbb{E}[X]$, X の 2 次と 3 次の原点まわり積率 $\mathbb{E}[X^2]$, $\mathbb{E}[X^3]$ を求めよ⁷ .

問題 1.32 連続型確率変数 X は確率密度関数

$$f_X(x) = \begin{cases} c(1-x^2) & (-1 < x < 1) \\ 0 & (\text{その他}) \end{cases}$$

をもつとする . ただし , c は正の定数とする . このとき , 以下の問いに答えよ .

- (i) $f_X(x)$ が確率密度関数になるように c を定めよ .
- (ii) X の期待値と分散 $\mathbb{E}(X)$, $\text{VAR}(X)$ を求めよ .
- (iii) X の分布関数 $F_X(x) = \mathbb{P}(X \leq x)$ を求めよ . ただし , $-\infty < x < \infty$ である .
- (iv) 確率 $\mathbb{P}(-0.9 < X < 0.9)$ を求めよ .
- (v) チェビシエフの不等式を用いて確率 $\mathbb{P}(-0.9 < X < 0.9)$ の下限を求めよ .

問題 1.33 連続型確率変数 X は確率密度関数

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \quad (-\infty < x < \infty)$$

を持つとする . ただし , μ, σ は定数で $-\infty < \mu < \infty, 0 < \sigma < \infty$ とし , $\exp(x) = e^x$ である .

$$Z = \frac{X - \mu}{\sigma}$$

と定義したとき , 以下の問いに答えよ .

- (i) Z の分布関数 $F_Z(z) = \mathbb{P}(Z \leq z) \quad (-\infty < z < \infty)$

を求めよ .

- (ii) Z の期待値と分散は

$$\mathbb{E}[Z] = 0, \quad \text{VAR}[Z] = 1$$

で与えられることを示せ . ただし , $\int_{-\infty}^{\infty} f_X(x) dx = 1$ および $\lim_{x \rightarrow \pm\infty} x \exp(-x^2/2) = 0$ は証明なしで用いてよい .

- (iii) X の期待値と分散 $\mathbb{E}[X]$ と $\text{VAR}[X]$ を求めよ .

問題 1.34 Ω を標本空間とし , \mathcal{F} を Ω 上の完全加法族⁸とする . 確率 \mathbb{P} は \mathcal{F} 上で定義された実数値関数でつぎの条件をみたすものであった .

(P1) 任意の $A \in \mathcal{F}$ に対して , $\mathbb{P}(A) \geq 0$

(P2) $\mathbb{P}(\Omega) = 1$

⁷ X の平均 は X の期待値と同じこと . また , X の 2 次の原点まわり積率は X^2 の期待値と同じこと .

⁸ Ω の部分集合のなす集まりで任意の可算回の集合演算に関して閉じている .

(P3) $i = 1, 2, \dots$ に対して $A_i \in \mathcal{F}$ かつ $A_i \cap A_j = \emptyset (i \neq j)$ ならば,

$$\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

(P1) から (P3) をどこでどのように使ったかを明示 (例に倣って) して以下の (i)–(iii) を証明せよ.

(i) (P1)–(P3) および (1.5) を用いてつぎのことを示せ: $B_1, B_2 \in \mathcal{F}$ とする.

$$B_1 \cap B_2 = \emptyset \quad \text{ならば,} \quad \mathbb{P}(B_1 \cup B_2) = \mathbb{P}(B_1) + \mathbb{P}(B_2) \quad (1.4)$$

(ii) (P1)–(P3) および (1.5) を用いてつぎのことを示せ: $B_3, B_4 \in \mathcal{F}$ に対して

$$\mathbb{P}(B_3 \cup B_4) \leq \mathbb{P}(B_3) + \mathbb{P}(B_4)$$

(iii) (P1)–(P3), (1.4) および (1.5) を用いてつぎのことを示せ: $C_1, C_2 \in \mathcal{F}$ とする.

$$C_1 \subset C_2 \quad \text{ならば,} \quad \mathbb{P}(C_1) \leq \mathbb{P}(C_2)$$

例: たとえば,

$$\mathbb{P}(\emptyset) = 0 \quad (1.5)$$

を示すには, $A_1 = \Omega, A_i = \emptyset (i \geq 2)$ とおくと

$$\Omega = \cup_{i=1}^{\infty} A_i \quad (1.6)$$

と

$$A_i \cap A_j = \emptyset (i \neq j) \quad (1.7)$$

となることに注意すれば,

$$\mathbb{P}(\Omega) = \mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i) = \mathbb{P}(\Omega) + \sum_{i=2}^{\infty} \mathbb{P}(\emptyset)$$

となる. ただし, 1 番目の等号は 1.6 からわかり, 2 番目の等号は (1.7) と (P3) からわかる. したがって,

$$\sum_{i=2}^{\infty} \mathbb{P}(\emptyset) = 0 \quad (1.8)$$

となる. しかし, (P1) から $\mathbb{P}(\emptyset) \geq 0$ なることと (1.8) から $\mathbb{P}(\emptyset) = 0$ がわかる.

第2章 1 次元の確率分布の代表的モデル

この章では、確率分布の具体的なモデルを学ぶ。

1 離散型確率変数のモデル

ベルヌーイ分布 確率変数 X は母数 p のベルヌーイ分布に従うとは、 X の確率関数が

$$f_X(x|p) = p^x(1-p)^{1-x}, \quad x = 0, 1$$

のときをいう。ただし、 $0 < p < 1$ である。この分布を $\text{Ber}(p)$ と記す。成功の確率が p ($0 < p < 1$)、失敗の確率が $1-p$ の試行をベルヌーイ試行とよび、この試行の成功を 1、失敗を 0 に対応させたものが X である。

定理 2.1 (ベルヌーイ分布の平均・分散)

$$\begin{aligned}\mathbb{E}[X] &= p \\ \text{VAR}[X] &= p(1-p)\end{aligned}$$

証明 $\mathbb{E}[X] = \sum_{x=0,1} x f_X(x|p) = 0 \times f_X(0|p) + 1 \times f_X(1|p) = p$ よりわかる。
分散も同様に $\text{VAR}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$ に注意すれば、

$$\begin{aligned}\text{VAR}[X] &= \sum_{x=0,1} (x-p)^2 f_X(x|p) \\ &= (0-p)^2 \times f_X(0|p) + (1-p)^2 \times f_X(1|p) \\ &= p(1-p)\end{aligned}$$

からわかる。

□

二項分布 確率変数 X は母数 n と p の二項分布に従うとは、 X の確率関数が

$$f_X(x|n, p) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \dots, n$$

のときをいう。ただし、 $n \geq 1$ は整数、 $0 < p < 1$ である。この分布を $\text{BN}(n, p)$ と記す。

定理 2.2 (二項分布の平均・分散・積率母関数)

$$\begin{aligned} \mathbb{E}[X] &= np \\ \text{VAR}[X] &= np(1-p) \\ M_X(t) &= (pe^t + (1-p))^n, \quad t \in \mathbb{R} \end{aligned}$$

証明 まず、積率母関数を求める。二項定理を用いると $t \in \mathbb{R}$ に対して、

$$\begin{aligned} M_X(t) &= \mathbb{E}[e^{tX}] \\ &= \sum_{x=0}^n e^{tx} f_X(x|n, p) \\ &= \sum_{x=0}^n \binom{n}{x} (pe^t)^x (1-p)^{n-x} \\ &= (pe^t + (1-p))^n \end{aligned}$$

となることがわかる。つぎに、定理 1.3 (積率母関数と積率の関係) を用いると

$$\begin{aligned} \mathbb{E}[X] &= \left. \frac{d}{dt} M_X(t) \right|_{t=0} = \left. n(pe^t + (1-p))^{n-1} pe^t \right|_{t=0} \\ &= np \\ \mathbb{E}[X^2] &= \left. \frac{d^2}{dt^2} M_X(t) \right|_{t=0} \\ &= \left. n(n-1)(pe^t + (1-p))^{n-2} (pe^t)^2 \right|_{t=0} + \left. n(pe^t + (1-p))^{n-1} pe^t \right|_{t=0} \\ &= n(n-1)p^2 + np \end{aligned}$$

となり、

$$\text{VAR}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2 = np(1-p)$$

がわかる .

□

幾何分布 確率変数 X はが母数 p の幾何分布に従うとは , X の確率関数が

$$f_X(x|p) = p(1-p)^{x-1}, \quad x = 1, 2, \dots$$

のときをいう . ただし , $0 < p < 1$ である . この分布を $G(p)$ と記す .

定理 2.3 (幾何分布の平均・分散・積率母関数)

$$\begin{aligned} \mathbb{E}[X] &= \frac{p}{1-p} \\ \text{VAR}[X] &= \frac{1-p}{p^2} \\ M_X(t) &= \frac{p}{1-t(1-p)}, \quad t < \frac{1}{1-p} \end{aligned}$$

証明 略 .

□

負の二項分布 確率変数 X は母数 n と p の負の二項分布に従うとは , X の確率関数が

$$f_X(x|n, p) = \binom{n+x-1}{n-1} p^n (1-p)^x, \quad x = 0, 1, 2, \dots$$

のときをいう . ただし , $n \geq 1$ は整数 , $0 < p < 1$ である . この分布を $\text{NBN}(n, p)$ と記す . 負の二項展開式 :

$$\frac{1}{p^n} = (p - (1-p))^n = \sum_{x=0}^{\infty} \binom{-n}{x} (-(1-p))^x = \sum_{x=0}^{\infty} \binom{n+x-1}{n-1} (1-p)^x$$

の両辺に p^n をかけた式の各項が確率関数である . ここで負の二項係数は

$$\binom{-n}{x} = \frac{1}{x!} (-n)(-n+1) \cdots (-n-x+1) = (-1)^x \binom{n+x-1}{n-1}$$

である .

定理 2.4 (負の二項分布の平均・分散)

$$\begin{aligned}\mathbb{E}[X] &= n \frac{p}{1-p} \\ \text{VAR}[X] &= n \frac{1-p}{p^2}\end{aligned}$$

証明 略 .

□

ポアソン分布 確率変数 X は母数 λ のポアソン分布に従うとは, X の確率関数が

$$f_X(x|\lambda) = \frac{\lambda^x}{x!} e^{-\lambda}, \quad x = 0, 1, 2, \dots$$

のときをいう. ただし, $\lambda > 0$ である. この分布を $\text{Po}(\lambda)$ と記す.

定理 2.5 (ポアソンの平均・分散・積率母関数)

$$\begin{aligned}\mathbb{E}[X] &= \lambda \\ \text{VAR}[X] &= \lambda \\ M_X(t) &= e^{\lambda(e^t-1)}\end{aligned}$$

証明 まず, 積率母関数を求める: $t \in \mathbb{R}$ に対して,

$$M_X(t) = \mathbb{E}[e^{tX}] = \sum_{x=0}^{\infty} e^{tx} f_X(x|\lambda) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t-1)}$$

となることがわかる. つぎに, 定理 1.3 (積率母関数と積率の関係) を用いると

$$\begin{aligned}\mathbb{E}[X] &= \left. \frac{d}{dt} M_X(t) \right|_{t=0} = \left. \lambda e^t e^{\lambda(e^t-1)} \right|_{t=0} = \lambda \\ \mathbb{E}[X^2] &= \left. \frac{d^2}{dt^2} M_X(t) \right|_{t=0} = \left. \lambda e^t e^{\lambda(e^t-1)} \right|_{t=0} + \left. (\lambda e^t)^2 e^{\lambda(e^t-1)} \right|_{t=0} \\ &= \lambda + \lambda^2\end{aligned}$$

となり,

$$\text{VAR}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2 = \lambda$$

がわかる.

□

分布	母数	確率関数 $P(X = x)$	平均	分散
ベルヌーイ分布	$0 < p < 1$	$p^x(1-p)^{1-x}, x = 0, 1$	p	$p(1-p)$
二項分布	$n \geq 1, 0 < p < 1$	$\binom{n}{x} p^x (1-p)^{n-x}$ ただし, $x = 0, 1, \dots, n$	np	$np(1-p)$
幾何分布	$0 < p < 1$	$p(1-p)^{x-1}$ ただし, $x = 1, 2, \dots$	$\frac{p}{1-p}$	$\frac{1-p}{p^2}$
負の二項分布	$n \geq 1, 0 < p < 1$	$\binom{n+x-1}{n-1} p^n (1-p)^x$ ただし, $x = 1, 2, \dots$	$n \frac{p}{1-p}$	$n \frac{1-p}{p^2}$
ポアソン分布	$0 < \lambda < \infty$	$\frac{\lambda^x}{x!} e^{-\lambda}, x = 0, 1, 2, \dots$	λ	λ

2 連続型確率変数のモデル

一様分布 確率変数 X は区間 (α, β) の一様分布に従うとは, X が確率密度関数

$$f_X(x) = \begin{cases} 1/(\beta - \alpha), & (\alpha < x < \beta), \\ 0 & \text{その他} \end{cases}$$

を持つときをいう. この分布を $U(\alpha, \beta)$ と記すことにする.

定理 2.6 (一様分布の平均・分散・積率母関数)

$$\begin{aligned} \mathbb{E}[X] &= \frac{\alpha + \beta}{2} \\ \text{VAR}[X] &= \frac{(\beta - \alpha)^2}{12} \end{aligned}$$

証明

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{\infty} x f_X(x) dx = \frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} x dx = \frac{\alpha + \beta}{2} \\ \mathbb{E}[X^2] &= \int_{-\infty}^{\infty} x^2 f_X(x) dx = \frac{1}{\beta - \alpha} \int_{\alpha}^{\beta} x^2 dx = \frac{\beta^3 - \alpha^3}{3(\beta - \alpha)} \end{aligned}$$

となることから

$$\text{VAR}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2 = \frac{\beta^3 - \alpha^3}{3(\beta - \alpha)} - \left\{ \frac{\alpha + \beta}{2} \right\}^2 = \frac{(\beta - \alpha)^2}{12}$$

となる.

□

正規分布 $(N(\mu, \sigma^2))$ 確率変数 X は母数 (μ, σ^2) の正規分布に従うとは、 X が確率密度関数

$$f_X(x|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad -\infty < x < \infty$$

を持つときをいう。ただし、 $-\infty < \mu < \infty, \sigma > 0$ である。特に、 $\mu = 0, \sigma^2 = 1$ のとき、正規分布 $N(0, 1)$ のことを標準正規分布という。

定理 2.7 (正規分布の平均・分散・積率母関数)

$$\begin{aligned} \mathbb{E}[X] &= \mu \\ \text{VAR}[X] &= \sigma^2 \\ M_X(t) &= \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right), \quad -\infty < t < \infty \end{aligned}$$

証明 まず、 Z が標準正規分布に従うとき、 $X = \sigma Z + \mu$ とおくと X は $N(\mu, \sigma^2)$ に従うことを示す。 $g(z) = \sigma z + \mu$ と $g^{-1}(x) = (1/\sigma)(x - \mu)$ として命題 1.13 を用いると

$$\begin{aligned} f_X(x) &= f_Z(g^{-1}(x)) \left| \frac{d}{dx} g^{-1}(x) \right| = \frac{1}{\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)} \left| \frac{1}{\sigma} \right| \\ &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \end{aligned}$$

となり、 X は $N(\mu, \sigma^2)$ に従うがわかった。 Z の確率密度関数を $f_Z(z) = (1/\sqrt{2\pi})e^{-z^2/2}$ とおくことにする。つぎに、期待値の線形性と分散の性質より

$$\begin{aligned} \mathbb{E}[X] &= \sigma\mathbb{E}[Z] + \mu \\ \text{VAR}[X] &= \sigma^2\text{VAR}[Z] \end{aligned}$$

となることに注意する。 $zf_Z(z)$ は奇関数であるので、

$$\mathbb{E}[Z] = \int_{-\infty}^{\infty} zf_Z(z) dz = 0$$

がわかる．また， $z^2 f_Z(z)$ は偶関数であることと部分積分の公式を用いると

$$\begin{aligned}\text{VAR}[Z] &= \int_{-\infty}^{\infty} z^2 f_Z(z) dz = 2 \int_0^{\infty} z^2 f_Z(z) dz \\ &= 2 \int_0^{\infty} z(-f_Z(z))' dz \\ &= 2 \left\{ [-zf_Z(z)]_0^{\infty} + \int_0^{\infty} f_Z(z) dz \right\} = 1\end{aligned}$$

がわかる．最後の等号は $\lim_{z \rightarrow \infty} z f_Z(z) = 0$ と $f_Z(z)$ は偶関数であることからわかる．

最後に， Z の積率母関数を求める：

$$\begin{aligned}M_Z(t) &= \mathbb{E}[e^{tZ}] \\ &= \int_{-\infty}^{\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(z-t)^2}{2} + \frac{t^2}{2}\right\} dx \\ &= e^{t^2/2} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(z-t)^2}{2}\right\} dz = e^{t^2/2}.\end{aligned}$$

したがって，

$$\begin{aligned}M_X(t) &= \mathbb{E}[e^{tX}] = \mathbb{E}[e^{t(\sigma Z + \mu)}] = e^{\mu t} \mathbb{E}[e^{\sigma t Z}] = e^{\mu t} M_Z(\sigma t) \\ &= \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}.\end{aligned}$$

□

指数分布 ($\text{Ex}(\lambda)$) 確率変数 X は母数 λ の指数分布に従うとは， X が確率密度関数

$$f_X(x|\lambda) = \begin{cases} \lambda e^{-\lambda x}, & (0 < x < \infty) \\ 0, & \text{その他} \end{cases}$$

を持つときをいう．ただし， $0 < \lambda < \infty$ である．この分布を $\text{Ex}(\lambda)$ と記すことにする．

定理 2.8 (指数分布の平均・分散・積率母関数)

$$\begin{aligned}\mathbb{E}[X] &= \frac{1}{\lambda} \\ \text{VAR}[X] &= \frac{1}{\lambda^2} \\ M_X(t) &= \frac{\lambda}{\lambda - t}, \quad -\infty < t < \lambda\end{aligned}$$

証明 まず、積率母関数を求める。 $t < \lambda$ に対して、

$$\begin{aligned}M_X(t) &= \mathbb{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x|\lambda) dx = \int_0^{\infty} e^{tx} f_X(x|\lambda) dx \\ &= \lambda \int_0^{\infty} e^{-(\lambda-t)x} dx = \frac{\lambda}{\lambda - t}\end{aligned}$$

となることがわかる。つぎに、定理 1.3 (積率母関数と積率の関係) を用いると

$$\begin{aligned}\mathbb{E}[X] &= \left. \frac{d}{dt} M_X(t) \right|_{t=0} = \left. \frac{\lambda}{(\lambda - t)^2} \right|_{t=0} = \frac{1}{\lambda} \\ \mathbb{E}[X^2] &= \left. \frac{d^2}{dt^2} M_X(t) \right|_{t=0} = \left. \frac{2\lambda}{(\lambda - t)^3} \right|_{t=0} = \frac{2}{\lambda^2}\end{aligned}$$

となり、

$$\text{VAR}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2 = \frac{1}{\lambda^2}$$

がわかる。

直接、積分を計算して積率をもとめてみよう：まず、

$$\int_{-\infty}^{\infty} f_X(x|\lambda) dx = \int_0^{\infty} \lambda e^{-\lambda x} dx = \left[-e^{-\lambda x} \right]_0^{\infty} = -\lim_{x \rightarrow \infty} e^{-\lambda x} + 1 = 1$$

に注意する。これと部分積分より

$$\begin{aligned}\mathbb{E}[X] &= \int_{-\infty}^{\infty} x f_X(x|\lambda) dx \\ &= \int_0^{\infty} x (-\lambda e^{-\lambda x})' dx \\ &= \left[-x e^{-\lambda x} \right]_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx \\ &= -\lim_{x \rightarrow \infty} x e^{-\lambda x} + \frac{1}{\lambda} \int_0^{\infty} \lambda e^{-\lambda x} dx \\ &= \frac{1}{\lambda}.\end{aligned}$$

再度，上の結果と部分積分を持ちいれば，

$$\begin{aligned}
 \mathbb{E}[X^2] &= \int_{-\infty}^{\infty} x^2 f_X(x|\lambda) dx \\
 &= \int_0^{\infty} x^2 (-e^{-\lambda x})' dx \\
 &= \left[-x^2 e^{-\lambda x} \right]_0^{\infty} + 2 \int_0^{\infty} x e^{-\lambda x} dx \\
 &= -\lim_{x \rightarrow \infty} x^2 e^{-\lambda x} + \frac{2}{\lambda} \int_0^{\infty} \lambda x e^{-\lambda x} dx \\
 &= \frac{2}{\lambda} \frac{1}{\lambda} = \frac{2}{\lambda^2}.
 \end{aligned}$$

□

ガンマ分布 確率変数 X は母数 α, β のガンマ分布に従うとは， X が確率密度関数

$$f_X(x|\alpha, \beta) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp(-x/\beta), & (0 < x < \infty) \\ 0, & \text{その他} \end{cases}$$

を持つときをいう．この分布を $GA(\alpha, \beta)$ と記すことにする．

特に， $\alpha = n/2$, ($n \in \mathbb{N}$) と $\beta = 2$ のとき，ガンマ分布 $GA(n/2, 2)$ を自由度 n のカイ自乗分布という．ただし， $0 < \alpha, \beta < \infty$ である．

注意 2.1 $\beta = 1/\lambda$ と $\alpha = 1$ のとき，指数分布になる．

定理 2.9 (ガンマ分布の平均・分散・積率母関数)

$$\begin{aligned}
 \mathbb{E}[X] &= \alpha\beta \\
 \text{VAR}[X] &= \alpha\beta^2 \\
 M_X(t) &= (1 - \beta t)^{-\alpha}, \quad t < 1/\beta
 \end{aligned}$$

証明 まず，積率母関数を求める．

$$\int_0^{\infty} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp(-x/\beta) dx = 1$$

に注意する． $t < 1/\beta$ に対して，

$$\begin{aligned} M_X(t) &= \mathbb{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x|\alpha, \beta) dx = \int_0^{\infty} e^{tx} f_X(x|\alpha, \beta) dx \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha} \lambda \int_0^{\infty} x^{\alpha-1} \exp(-x(1-\beta t)/\beta) dx = \left(\frac{1}{1-\beta t}\right)^\alpha \end{aligned}$$

となることがわかる．つぎに，定理 1.3 (積率母関数と積率の関係) を用いると

$$\begin{aligned} \mathbb{E}[X] &= \left. \frac{d}{dt} M_X(t) \right|_{t=0} = \left. \frac{\alpha\beta}{(1-\beta t)^{\alpha+1}} \right|_{t=0} = \alpha\beta \\ \mathbb{E}[X^2] &= \left. \frac{d^2}{dt^2} M_X(t) \right|_{t=0} = \left. \frac{\alpha(\alpha+1)\beta^2}{(1-\beta t)^{\alpha+2}} \right|_{t=0} = \alpha(\alpha+1)\beta^2 \end{aligned}$$

となり，

$$\text{VAR}[X] = \mathbb{E}[X^2] - \{\mathbb{E}[X]\}^2 = \alpha\beta^2$$

がわかる．

□

分布	母数	確率密度関数	平均	分散
一様分布	$\alpha, \beta (\alpha < \beta) \in \mathbb{R}$	$\frac{1}{(\beta - \alpha)}, \alpha < x < \beta$	$\frac{\alpha + \beta}{2}$	$\frac{(\beta - \alpha)^2}{12}$
標準正規分布	なし	$\frac{1}{\sqrt{2\pi}} e^{-x^2/2}, x \in \mathbb{R}$	0	1
正規分布	$\mu \in \mathbb{R}, \sigma^2 > 0$	$\frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}, x \in \mathbb{R}$	μ	σ^2
指数分布	$\lambda > 0$	$\lambda e^{-\lambda x}, x > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
ガンマ分布	$\alpha, \beta > 0$	$\frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \exp(-x/\beta), x > 0$	$\alpha\beta$	$\alpha\beta^2$
カイ自乗分布	$n \in \mathbb{N}$	$\frac{1}{\Gamma(n/2)2^{n/2}} e^{-x/2} x^{(n/2)-1}, x > 0$	n	$2n$

3 演習問題

問題 2.1 コインを 3 回投げたときの次の確率を求めよ．

- (1) 3 回表が出る確率．
- (2) 2 回表で 1 回裏の出る確率．
- (3) 少なくとも 1 回表の出る確率．

問題 2.2 サイコロを 5 回投げたときの次の確率を求めよ．

- (1) 3 の目が 2 回が出る確率 .
- (2) 3 の目が多くとも 1 回出る確率 .
- (3) 少なくとも 2 回出る確率 .

答えは二項分布の確率に式の数値を代入したままでよい .

問題 2.3 X は母数 n と p の二項分布に従うとする . ただし , n は自然数 , $0 < p < 1$ とする . $np = \lambda (\lambda > 0$ は固定した値) として , $n \rightarrow \infty$ としたときに ,

$$\mathbb{P}(X = x) \rightarrow \frac{\lambda^x e^{-\lambda}}{x!}$$

となることを示せ .

ちなみに , 母数 10 と 0.1 の二項分布に従う確率変数を X とすれば ,

$$\mathbb{P}(X = 2) = \binom{10}{2} \left(\frac{1}{10}\right)^2 \left(\frac{9}{10}\right)^{10-2} = 0.1937$$

$\lambda = np = (10)(0.1) = 1$ のポアソン分布でこの確率を近似すれば

$$\mathbb{P}(X = 2) \approx \frac{\lambda^2 e^{-\lambda}}{2!} = \frac{1^2 e^{-1}}{2!} = 0.1839$$

となる .

問題 2.4 確率変数 Z は標準正規分布 $N(0, 1)$ に従うとする . すなわち , 標準正規分布の確率密度関数 $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ を用いて , 以下の確率を表現し , 数表を用いて確率を求めよ .

- (1) $\mathbb{P}(0 < Z \leq 1.2)$
- (2) $\mathbb{P}(-0.68 < Z \leq 0)$
- (3) $\mathbb{P}(-0.46 < Z \leq 2.21)$
- (4) $\mathbb{P}(0.81 < Z \leq 1.94)$
- (5) $\mathbb{P}(Z > -1.28)$

ヒント : $\mathbb{P}(Z > 1.2) = 0.115$, $\mathbb{P}(Z > 0.68) = 0.248$, $\mathbb{P}(Z > 0.46) = 0.323$, $\mathbb{P}(Z > 2.21) = 0.014$, $\mathbb{P}(Z > 0.81) = 0.209$, $\mathbb{P}(Z > 1.94) = 0.026$, $\mathbb{P}(Z > 1.28) = 0.100$

問題 2.5 確率変数 X は正規分布 $N(151, (15)^2)$ に従うとする . 標準正規分布の確率密度関数 $f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$ を用いて , 以下の確率を表現し , 数表を用いて確率を求めよ .

- (1) $\mathbb{P}(119.5 < X \leq 155.5)$
- (2) $\mathbb{P}(X > 185.5)$

ヒント : $\mathbb{P}(Z > 2.10) = 0.018$, $\mathbb{P}(Z > 0.30) = 0.382$, $\mathbb{P}(Z > 2.30) = 0.011$

第3章 多次元の確率変数

この章では、複数の確率変数の分布を学ぶ。

1 同時分布と周辺分布

$(\Omega, \mathcal{F}, \mathbb{P})$ を確率空間とし、 X, Y をこの確率空間上の確率変数とする。これらふたつの確率変数を組として考えた (X, Y) を 2 次元確率ベクトルという。さらに、 (X, Y) の分布を同時分布とよび、任意の $A, B \in \mathcal{B}(\mathbb{R})$ に対して、

$$P_{X,Y}(A \times B) = \mathbb{P}\{\omega \in \Omega : X(\omega) \in A, Y(\omega) \in B\} = \mathbb{P}(X \in A, Y \in B)$$

で定める。すなわち、 $P_{X,Y}$ は確率ベクトル (X, Y) によって \mathbb{P} より誘導された $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ 上¹の確率測度である。

X, Y それぞれの分布 P_X, P_Y をそれぞれの周辺分布という。

定義 3.1 2 次元確率ベクトル (X, Y) の同時分布関数を

$$\begin{aligned} F_{X,Y}(x, y) &= P_{X,Y}((-\infty, x] \times (-\infty, y]) \\ &= \mathbb{P}(X \leq x, Y \leq y) \\ &= \mathbb{P}(\{\omega \in \Omega : X(\omega) \leq x, Y(\omega) \leq y\}), \quad x, y \in \mathbb{R} \end{aligned}$$

で定める。各成分だけに注目した分布関数

$$F_X(x) = P_X((-\infty, x]) = \mathbb{P}(X \leq x), \quad F_Y(y) = P_Y((-\infty, y]) = \mathbb{P}(Y \leq y)$$

をそれぞれの周辺分布関数とよぶ。

命題 3.1 (同時分布関数の性質) (i) すべての $(x, y) \in \mathbb{R}^2$ に対して、 $0 \leq$

$$F_{X,Y}(x, y) \leq 1.$$

¹ $\mathcal{B}(\mathbb{R}^2) = \sigma(\{A \times B : A \in \mathcal{B}(\mathbb{R}), B \in \mathcal{B}(\mathbb{R})\})$ である。

(ii) $x_1 < x_2, y_1 < y_2$ に対して, $F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_2)$.

(iii) $\lim_{x \rightarrow -\infty} F_{X,Y}(x, y) = 0, \lim_{y \rightarrow -\infty} F_{X,Y}(x, y) = 0,$
 $\lim_{x \rightarrow \infty, y \rightarrow \infty} F_{X,Y}(x, y) = 1.$

証明 (i) 同時分布関数の定義と確率の定義からわかる.

(ii) $\{(X, Y) \in (-\infty, x_1] \times (-\infty, y_1]\} \subset \{(X, Y) \in (-\infty, x_2] \times (-\infty, y_2]\}$ に注意して, 命題 1.4(vi) を用いればよい.

(iii) $\cap_{n=1}^{\infty} (-\infty, -n] = \emptyset$ に注意をして, 命題 1.4(i) と (ix) を用いると

$$\begin{aligned} \lim_{n \rightarrow \infty} F_{X,Y}(x, -n) &= \lim_{n \rightarrow \infty} \mathbb{P}((X, Y) \in (-\infty, x] \times (-\infty, -n]) \\ &= \mathbb{P}((X, Y) \in \cap_{n=1}^{\infty} (-\infty, x] \times (-\infty, -n]) = \mathbb{P}(\emptyset) = 0 \end{aligned}$$

よりわかる. のこりも同様である. □

命題 3.2 (同時分布関数と周辺分布関数の関係)

$$F_X(x) = \lim_{y \rightarrow \infty} F_{X,Y}(x, y), \quad F_Y(y) = \lim_{x \rightarrow \infty} F_{X,Y}(x, y)$$

証明 命題 1.4(viii) に注意して

$$\begin{aligned} \lim_{n \rightarrow \infty} F_{X,Y}(x, n) &= \lim_{n \rightarrow \infty} \mathbb{P}((X, Y) \in (-\infty, x] \times (-\infty, n]) \\ &= \mathbb{P}((X, Y) \in \cup_{n=1}^{\infty} (-\infty, x] \times (-\infty, n]) \\ &= \mathbb{P}((X, Y) \in (-\infty, x] \times \mathbb{R}) = \mathbb{P}(X \leq x) \end{aligned}$$

からわかる. □

定義 3.2 2つの確率変数 X, Y が独立であるとは, その同時分布 $P_{X,Y}$ が周辺分布 P_X, P_Y の積で表されることである: すなわち, 任意の $A, B \in \mathcal{B}(\mathbb{R})$ に対して

$$P_{X,Y}(A \times B) = P_X(A)P_Y(B)$$

が成り立つこと²である. 独立でないときを従属という.

²これは $\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$ である.

注意 3.1 つぎは同値である .

(1) X, Y は独立である .

(2) $F_{X, Y}(x, y) = F_X(x)F_Y(y)$. ただし , $x, y \in \mathbb{R}$ である .

証明 (1) \Rightarrow (2) は独立性の定義において , $A = (-\infty, x], B = (-\infty, y]$ とすればわかる . 逆については , 略 . \square

1.1 同時確率関数

確率変数 X, Y はともに離散型であって , それぞれは高々可算個の点で値をとるとする .

定義 3.3 離散型確率変数 (X, Y) の同時確率関数とは , \mathbb{R}^2 上の実数値関数 $f_{X, Y}(x, y)$ で

$$f_{X, Y}(x, y) = \mathbb{P}(X = x, Y = y)$$

をみたすものをいう .

$S = \{(x, y) \in \mathbb{R}^2 : f_{X, Y}(x, y) > 0\}$ とおけば , S は可算集合となる . さらに , $S_X = \{x \in \mathbb{R} : f_{X, Y}(x, y) > 0(\text{ある } y \in \mathbb{R})\}$ と $S_Y = \{y \in \mathbb{R} : f_{X, Y}(x, y) > 0(\text{ある } x \in \mathbb{R})\}$ とする . このとき , 同時確率関数は

$$(i) f_{X, Y}(x, y) \geq 0$$

$$(ii) \sum_{(x, y) \in S} f_{X, Y}(x, y) = 1$$

(iii) \mathbb{R}^2 の任意の部分集合³ A に対して ,

$$\mathbb{P}((X, Y) \in A) = \sum_{(x, y) \in A \cap S} f_{X, Y}(x, y)$$

定義 3.4 離散型確率変数 (X, Y) の同時分布関数とは , \mathbb{R}^2 上の実数値関数 $F_{X, Y}(x, y)$ で

$$\begin{aligned} F_{X, Y}(x, y) &= \mathbb{P}\{(X, Y) \in (-\infty, x] \times (-\infty, y]\} \\ &= \sum_{(s, t) : s \leq x, t \leq y, (s, t) \in S} f_{X, Y}(s, t) \end{aligned}$$

で定義されるものをいう .

³正確には , 任意のボレル集合 $A \in \mathcal{B}(\mathbb{R}^2)$

X と Y のそれぞれの確率関数を

$$f_X(x) = P(X = x), \quad f_Y(y) = P(Y = y)$$

で定めることにする．同時確率関数に対して， f_X と f_Y を X と Y の周辺確率関数ということにする．

命題 3.3 離散型確率変数 (X, Y) は同時確率関数 $f_{X,Y}(x, y)$ を持つとする．このとき，

$$\begin{aligned} f_X(x) &= \sum_{y \in S_Y} f_{X,Y}(x, y), \\ f_Y(y) &= \sum_{x \in S_X} f_{X,Y}(x, y) \end{aligned}$$

が成立する．

証明 f_X について示す．固定した x に対して， $A_x = \{(x, y) \in \mathbb{R}^2 : -\infty < y < \infty\}$ とおく．このとき， $x \in S_X$ に対して，

$$\begin{aligned} f_X(x) &= P(X = x) \\ &= \mathbb{P}(X = x, -\infty < Y < \infty) \\ &= \mathbb{P}((X, Y) \in A_x) \\ &= \sum_{(x, y) \in A_x \cap S} f_{X,Y}(x, y) \\ &= \sum_{y \in S_Y} f_{X,Y}(x, y) \end{aligned}$$

よりわかる． f_Y についても同様に示される． □

1.2 同時確率密度関数

定義 3.5 連続型確率ベクトル (X, Y) とし， $F_{X,Y}(x, y)$ をその同時分布関数とする． \mathbb{R}^2 上の実数値関数 $f_{X,Y}(x, y)$ ですべての $A \subset \mathbb{R}^2$ に対して，

$$\mathbb{P}((X, Y) \in A) = \int \int_A f_{X,Y}(x, y) dx dy$$

をみたすものが存在するとき， $f_{X,Y}(x, y)$ を (X, Y) の同時確率密度関数という．

命題 3.4 (同時確率密度関数の性質) (i) すべての $x \in \mathbb{R}, y \in \mathbb{R}$ に対して,
 $f_{X,Y}(x, y) \geq 0$.

(ii) すべての $(x, y) \in \mathbb{R}^2$ に対して,

$$F_{X,Y}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}(s, t) ds dt.$$

(iii) $F_{X,Y}(x, y)$ が同時確率密度関数を持つならば, $x \in \mathbb{R}, y \in \mathbb{R}$ に対して,

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y)$$

となる.

証明 証明は明らか. □

注意 3.2 確率ベクトル (X, Y) が同時確率密度関数 $f_{X,Y}(x, y)$ を持つとき,
 X と Y の周辺確率密度関数は

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy, \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$$

と表現できること⁴に注意せよ.

1.3 独立性

定義 3.6 確率ベクトル (X, Y) は同時確率関数または同時確率密度関数 $f_{X,Y}(x, y)$ をもつとする. このとき, X と Y が独立であるとは, すべての $x \in \mathbb{R}, y \in \mathbb{R}$ に対して

$$f_{X,Y}(x, y) = f_X(x)f_Y(y)$$

が成立することである.

補題 3.1 確率ベクトル (X, Y) は同時確率関数または同時確率密度関数 $f_{X,Y}(x, y)$ をもつとする. このとき, X と Y が独立であるとはための必要十分条件は, \mathbb{R} 上で定義されたある関数 $g(x)$ と $h(y)$ が存在し, すべての $x \in \mathbb{R}, y \in \mathbb{R}$ に対して

$$f_{X,Y}(x, y) = g(x)h(y)$$

⁴証明は定理 3.3 と同様に証明される.

とかけることである。

証明 \Rightarrow (必要条件) は $g(x) = f_X(x)$, $h(y) = f_Y(y)$ とおけばよい。

\Leftarrow (十分条件) は連続型についてのみ示すことにする。同時確率密度関数が $f_{X,Y}(x, y) = g(x)h(y)$ と表現されたとする。さらに,

$$\int_{-\infty}^{\infty} g(x) dx = c, \quad \int_{-\infty}^{\infty} h(y) dy = d$$

とおくと定数 c と d は関係式

$$\begin{aligned} cd &= \left(\int_{-\infty}^{\infty} g(x) dx \right) \left(\int_{-\infty}^{\infty} h(y) dy \right) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)h(y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1 \end{aligned} \quad (3.1)$$

をみたく。さらに,

$$f_X(x) = \int_{-\infty}^{\infty} g(x)h(y) dy = g(x)d, \quad f_Y(y) = \int_{-\infty}^{\infty} g(x)h(y) dx = h(y)c \quad (3.2)$$

となる。(3.1) と (3.2) から

$$f_{X,Y}(x, y) = g(x)h(y) = g(x)h(y)cd = f_X(x)f_Y(y)$$

となり⁵, X と Y が独立であることが示せた。□

例 3.1 離散型確率ベクトル (X, Y) の同時確率関数が以下のように与えられているとする:

$$\begin{aligned} f_{X,Y}(0, 10) &= f_{X,Y}(0, 20) = \frac{2}{18}, & f_{X,Y}(1, 10) &= f_{X,Y}(1, 30) = \frac{3}{18}, \\ f_{X,Y}(2, 20) &= \frac{4}{18}, & f_{X,Y}(2, 30) &= \frac{4}{18}. \end{aligned}$$

ただし, その他の (x, y) では $f_{X,Y}(x, y) = 0$ である。 X の周辺確率関数は

$$f_X(0) = \frac{4}{18}, f_X(1) = \frac{6}{18}, f_X(2) = \frac{8}{18}$$

⁵二番目の等号は $cd = 1$, 三番目の等号は (3.2) よりわかる。

となり, Y の周辺確率関数は

$$f_Y(10) = \frac{5}{18}, \quad f_Y(20) = \frac{6}{18}, \quad f_Y(30) = \frac{7}{18}$$

となる. よって, X と Y は独立でない. たとえば,

$$f_{X,Y}(0, 10) = \frac{2}{18} \neq \frac{4}{18} \times \frac{5}{18} = f_X(0)f_Y(10)$$

からわかる.

1.4 同時分布に関する期待値

定義 3.7 確率ベクトル (X, Y) は同時確率関数または同時確率密度関数 $f_{X,Y}(x, y)$ を持つとし, $g(x, y)$ を \mathbb{R}^2 上の実数値関数とする. このとき, $g(X, Y)$ の期待値を

$$\mathbb{E}[g(X, Y)] = \begin{cases} \sum_{(x,y) \in S} g(x, y) f_{X,Y}(x, y), & (\text{離散型}) \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f_{X,Y}(x, y) dx dy, & (\text{連続型}) \end{cases}$$

で定義する. ただし, 離散型の場合は $\sum_{(x,y) \in S} |g(x, y)| f_{X,Y}(x, y) < \infty$ のとき,

連続型の場合は $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g(x, y)| f_{X,Y}(x, y) dx dy < \infty$ のとき, $g(X, Y)$ の期待値を定義することにする. 期待値が定義されるとき, $g(X, Y)$ の期待値が存在するという.

記法について

確率変数のベクトルや行列に対する期待値の作用を以下のように書くことにする. たとえば, 確率ベクトル (X, Y) に対して,

$$\mathbb{E}(X, Y) = (\mathbb{E}(X), \mathbb{E}(Y))$$

などと書き, 行列の成分が確率変数である確率行列に対しては,

$$\mathbb{E} \begin{bmatrix} X^2 & XY \\ XY & Y^2 \end{bmatrix} = \begin{bmatrix} \mathbb{E}[X^2] & \mathbb{E}[XY] \\ \mathbb{E}[XY] & \mathbb{E}[Y^2] \end{bmatrix}$$

である.

定理 3.1 X と Y は独立な確率変数とし、実数上で定義された実数値関数 $h_1(x)$ と $h_2(y)$ は x と y にのみにそれぞれ依存するものとする。このとき、

$$\mathbb{E}[h_1(X)h_2(Y)] = \mathbb{E}[h_1(X)]\mathbb{E}[h_2(Y)]$$

が成立する。ただし、それぞれの期待値は存在するものと仮定する。

証明 (X, Y) がともに連続型確率変数とし、同時確率密度関数 $f_{X, Y}(x, y)$ を持つ場合について証明する。独立性の定義を利用すれば、

$$\begin{aligned} \mathbb{E}[h_1(X)h_2(Y)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_1(x)h_2(y)f_{X, Y}(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_1(x)h_2(y)f_X(x)f_Y(y) dx dy \\ &= \int_{-\infty}^{\infty} h_1(x)f_X(x) \left(\int_{-\infty}^{\infty} h_2(y)f_Y(y) dy \right) dx \\ &= \int_{-\infty}^{\infty} h_1(x)f_X(x) dx \int_{-\infty}^{\infty} h_2(y)f_Y(y) dy \\ &= \mathbb{E}[h_1(X)]\mathbb{E}[h_2(Y)] \end{aligned}$$

より示せた。離散型の場合は積分記号を和の記号に直せばよい。 □

定理 3.2 X と Y は独立な確率変数とし、それぞれは積率母関数 $M_X(t)$ と $M_Y(t)$ を持つとする。このとき、 $Z = X + Y$ の積率母関数は

$$M_Z(t) = M_X(t)M_Y(t)$$

で与えられる。

証明 定理 3.1 から

$$M_Z(t) = \mathbb{E}[e^{tZ}] = \mathbb{E}[e^{tX}e^{tY}] = \mathbb{E}[e^{tX}]\mathbb{E}[e^{tY}] = M_X(t)M_Y(t)$$

がわかる。 □

注意 3.3 X と Y は独立な確率変数とし、それぞれは $N(\mu_1, \sigma_1^2)$ と $N(\mu_2, \sigma_2^2)$ に従うとする。このとき、それぞれの積率母関数は

$$M_X(t) = \exp(\mu_1 t + \frac{1}{2}\sigma_1^2 t^2), \quad M_Y(t) = \exp(\mu_2 t + \frac{1}{2}\sigma_2^2 t^2), \quad t \in \mathbb{R}$$

となった． $Z = X + Y$ の積率母関数は定理 3.2 から

$$\begin{aligned} M_Z(t) &= M_X(t)M_Y(t) = \exp(\mu_1 t + \frac{1}{2}\sigma_1^2 t^2) \exp(\mu_2 t + \frac{1}{2}\sigma_2^2 t^2) \\ &= \exp\{(\mu_1 + \mu_2)t + \frac{1}{2}(\sigma_1^2 + \sigma_2^2)t^2\} \end{aligned}$$

となる．したがって， Z は $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$ に従うことがわかる．

2 共分散と相関係数

X と Y の 2 次までの積率としてそれぞれの平均と分散：

$$\mathbb{E}[X], \quad \mathbb{E}[Y], \quad \text{VAR}[X], \quad \text{VAR}[Y]$$

に加えて，平均まわりの相互積率を X と Y の共分散といい，

$$\text{COV}[X, Y] = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))]$$

と記すこと⁶にする．

定理 3.3 (共分散の公式)

$$\text{COV}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

証明

$$\begin{aligned} \text{COV}[X, Y] &= \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] \\ &= \mathbb{E}[XY - X\mathbb{E}[Y] - Y\mathbb{E}[X] + \mathbb{E}[X]\mathbb{E}[Y]] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \end{aligned}$$

がわかる．

□

定理 3.4 X と Y が独立ならば，

$$\text{COV}[X, Y] = 0$$

である．

⁶ $\mathbb{E}[X^2]$ と $\mathbb{E}[Y^2]$ の期待値の存在を仮定すれば，ここで考えている期待値はすべて存在する．

証明

$$\begin{aligned}\text{COV}[X, Y] &= \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \\ &= \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[X]\mathbb{E}[Y]\end{aligned}$$

がわかる .

□

注意 3.4 定理 3.4 の逆は必ずしも成立しないことに注意せよ .

定理 3.5 a, b を定数として , 期待値 , 分散 , および共分散はつぎのようになる :

$$\begin{aligned}\mathbb{E}[aX + bY] &= a\mathbb{E}[X] + b\mathbb{E}[Y], \\ \text{VAR}[aX + bY] &= a^2\text{VAR}[X] + b^2\text{VAR}[Y] + 2ab\text{COV}[X, Y] \\ \text{COV}[X, Y] &= \text{COV}[Y, X] \\ \text{COV}[aX + bY, Z] &= a\text{COV}[X, Z] + b\text{COV}[Y, Z]\end{aligned}$$

となる . 特に , X と Y が独立ならば ,

$$\text{VAR}[aX + bY] = a^2\text{VAR}[X] + b^2\text{VAR}[Y]$$

である .

証明

□

例 3.2 (X, Y) は同時確率密度関数

$$f_{X, Y}(x, y) = \begin{cases} 1, & 0 < x < 1, x < y < x + 1, \\ 0, & (\text{その他}), \end{cases}$$

を持つとする . このときの X と Y の共分散を求めよう . まず , X と Y の周辺確率密度関数を求めよう . $0 < x < 1$ のとき , $f_{X, Y}(x, y) > 0$ であることに注意すれば , $0 < x < 1$ のとき ,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X, Y}(x, y) dy = \int_x^{x+1} dy = 1$$

となる．それ以外では $f_{X,Y}(x,y) = 0$ なので， $f_X(x) = 0$ となる．したがって

$$f_X(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & (\text{その他}), \end{cases}$$

となる．また， $0 < y < 2$ のとき， $f_{X,Y}(x,y) > 0$ となることに注意すれば，

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx \\ &= \begin{cases} \int_0^y dx, & 0 < y < 1, \\ \int_{y-1}^1 dx, & 1 \leq y < 2, \end{cases} \\ &= \begin{cases} y, & 0 < y < 1, \\ 2 - y, & 1 \leq y < 2, \end{cases} \end{aligned}$$

となる．したがって

$$f_Y(y) = \begin{cases} y, & 0 < y < 1, \\ 2 - y, & 1 \leq y < 2, \\ 0, & (\text{その他}) \end{cases}$$

となる．これらに注意すれば，

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x dx = \frac{1}{2}, \\ \mathbb{E}[Y] &= \int_{-\infty}^{\infty} y f_Y(y) dy = \int_0^1 y^2 dy + \int_1^2 y(2-y) dy = \frac{1}{3} + \frac{2}{3} = 1, \\ \mathbb{E}[XY] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_{X,Y}(x,y) dx dy = \int_0^1 \int_x^{x+1} xy dy dx \\ &= \int_0^1 x \left[\frac{y^2}{2} \right]_x^{x+1} dx = \int_0^1 \frac{x(2x+1)}{2} dx = \left[\frac{x^3}{3} + \frac{x^2}{4} \right]_0^1 = \frac{7}{12} \end{aligned}$$

となる．したがって，

$$\text{COV}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{7}{12} - \frac{1}{2} = \frac{1}{12}$$

となる．

相関係数 X と Y の標準化

$$\frac{X - \mathbb{E}[X]}{\sqrt{\text{VAR}[X]}}, \quad \frac{Y - \mathbb{E}[Y]}{\sqrt{\text{VAR}[Y]}}$$

の積の平均を X と Y の相関係数といい、

$$\rho[X, Y] = \mathbb{E} \left[\frac{X - \mathbb{E}[X]}{\sqrt{\text{VAR}[X]}} \times \frac{Y - \mathbb{E}[Y]}{\sqrt{\text{VAR}[Y]}} \right] = \frac{\text{COV}[X, Y]}{\sqrt{\text{VAR}[X]}\sqrt{\text{VAR}[Y]}}$$

で記すことにする。

$\rho[X, Y] > 0$ のときには X と Y に正の相関, $\rho[X, Y] < 0$ のときには X と Y に負の相関があるといい, $\rho[X, Y] = 0$ のときには X と Y は無相関であるという. 定義と定理 3.4 から, X と Y が独立ならば, X と Y は無相関である. また, 注意 3.4 から X と Y は無相関であっても X と Y が必ずしも独立ではない.

定理 3.6 X と Y は 2 次の積率をもつ任意の確率変数とする. このとき,

- (i) $-1 \leq \rho[X, Y] \leq 1$.
- (ii) $|\rho[X, Y]| = 1$ となるための必要十分条件は定数 $a, b (a \neq 0)$ が存在して,

$$P(Y = aX + b) = 1$$

が成立することである. 特に, $\rho[X, Y] = 1$ ならば, $a > 0$ となり, $\rho[X, Y] = -1$ ならば, $a < 0$ となる.

証明 (i) を示すために,

$$h(t) = \mathbb{E}\{(X - \mu_X)t + (Y - \mu_Y)\}^2$$

とする. ただし, $\mu_X = \mathbb{E}[X]$, $\mu_Y = \mathbb{E}[Y]$ である. 期待値の中を展開すれば,

$$\begin{aligned} h(t) &= t^2 \mathbb{E}[(X - \mu_X)^2] + 2t \mathbb{E}[(X - \mu_X)(Y - \mu_Y)] + \mathbb{E}[(Y - \mu_Y)^2] \\ &= t^2 \text{VAR}[X] + 2t \text{COV}[X, Y] + \text{VAR}[Y] \\ &= \text{VAR}[X] \left(t + \frac{\text{COV}[X, Y]}{\text{VAR}[X]} \right)^2 + \frac{\text{VAR}[X]\text{VAR}[Y] - \{\text{COV}[X, Y]\}^2}{\text{VAR}[X]} \end{aligned}$$

となる．しかし， $h(t) \geq 0$ から

$$\text{VAR}[X]\text{VAR}[Y] - \{\text{COV}[X, Y]\}^2 \geq 0$$

を得る．したがって，

$$\frac{\{\text{COV}[X, Y]\}^2}{\text{VAR}[X]\text{VAR}[Y]} \leq 1 \iff \left| \frac{\text{COV}[X, Y]}{\sqrt{\text{VAR}[X]}\sqrt{\text{VAR}[Y]}} \right| \leq 1$$

となり，(i) は示せた．

(ii) 上の議論から $|\rho[X, Y]| = 1$ となるための必要十分条件は

$$\text{VAR}[X]\text{VAR}[Y] - \{\text{COV}[X, Y]\}^2 = 0 \iff t \text{ の二次方程式 } h(t) = 0 \text{ は単根である．}$$

である．この解を t_0 とする．しかし， $\mathbb{E}[(X - \mu_X)t_0 + (Y - \mu_Y)] = 0$ に注意すれば，

$$0 = h(t_0) = \mathbb{E}\{(X - \mu_X)t_0 + (Y - \mu_Y)\}^2 = \text{VAR}[(X - \mu_X)t_0 + (Y - \mu_Y)]$$

となり，分散の性質 (iii) から

$$P((X - \mu_X)t_0 + (Y - \mu_Y) = 0) = 1$$

となる．したがって， $a = -t_0$ ， $b = \mu_X t_0 + \mu_Y$ とおけば， $P(aX + b = Y) = 1$ が成立することがわかる．また，

$$h(t) = \text{VAR}[X] \left(t + \frac{\text{COV}[X, Y]}{\text{VAR}[X]} \right)^2 = 0$$

に注意すれば，方程式

$$h(t) = \text{VAR}[X] \left(t + \frac{\text{COV}[X, Y]}{\text{VAR}[X]} \right)^2 = 0$$

の解は

$$t = t_0 = -\frac{\text{COV}[X, Y]}{\text{VAR}[X]}$$

なので，

$$a = \frac{\text{COV}[X, Y]}{\text{VAR}[X]}$$

となる．したがって， a の符号と $\rho[X, Y]$ の符号は同じになる．よって，(ii) は示された． \square

3 条件付き分布と独立性

3.1 離散型確率変数の場合

定義 3.8 (X, Y) は離散型確率ベクトルとし, 同時確率関数 $f_{X, Y}(x, y)$ および周辺確率関数 $f_X(x)$ と $f_Y(y)$ を持つとする.

(i) $\mathbb{P}(X = x) = f_X(x) > 0$ なる任意の x に対して, $X = x$ が与えられたときの Y の条件付確率関数を $f_{Y|X}(y|x)$ で記し,

$$f_{Y|X}(y|x) = \mathbb{P}(Y = y|X = x) = \frac{f_{X, Y}(x, y)}{f_X(x)}$$

で定める.

(ii) $\mathbb{P}(Y = y) = f_Y(y) > 0$ なる任意の y に対して, $Y = y$ が与えられたときの X の条件付確率関数を $f_{X|Y}(x|y)$ で記し,

$$f_{X|Y}(x|y) = \mathbb{P}(X = x|Y = y) = \frac{f_{X, Y}(x, y)}{f_Y(y)}$$

で定める.

注意 3.5 $f_{Y|X}(y|x)$ は確率関数であることに注意せよ. すなわち, 各 x に対して,

- $f_{Y|X}(y|x) \geq 0, y \in \mathbb{R},$
- $\sum_{y \in S_Y} f_{Y|X}(y|x) = 1$

となっている.

例 3.3 離散型確率ベクトル (X, Y) の同時確率関数が以下のように与えられているとする:

$$\begin{aligned} f_{X, Y}(0, 10) &= f_{X, Y}(0, 20) = \frac{2}{18}, & f_{X, Y}(1, 10) &= f_{X, Y}(1, 30) = \frac{3}{18}, \\ f_{X, Y}(1, 20) &= \frac{4}{18}, & f_{X, Y}(2, 30) &= \frac{4}{18}. \end{aligned}$$

ただし, その他の (x, y) では $f_{X, Y}(x, y) = 0$ である. $X = x, x = 0, 1, 2$ が与えられたときの Y の条件付確率関数を求めよう. そのために, X の周辺確

率関数を求める：

$$f_X(0) = f_{X,Y}(0, 10) + f_{X,Y}(0, 20) = \frac{4}{18},$$

$$f_X(1) = f_{X,Y}(1, 10) + f_{X,Y}(1, 20) + f_{X,Y}(1, 30) = \frac{10}{18},$$

$$f_X(2) = f_{X,Y}(2, 30) = \frac{4}{18}.$$

$x = 0$ のとき, $y = 10, 20$ のとき $f_{X,Y}(0, y) > 0$ であるので, $y = 10, 20$ のとき $f_{Y|X}(y|0) > 0$ となり,

$$f_{Y|X}(10|0) = \frac{f_{X,Y}(0, 10)}{f_X(0)} = \frac{\frac{2}{18}}{\frac{4}{18}} = \frac{1}{2},$$

$$f_{Y|X}(20|0) = \frac{f_{X,Y}(0, 20)}{f_X(0)} = \frac{\frac{2}{18}}{\frac{4}{18}} = \frac{1}{2}$$

となる. したがって, $X = 0$ という情報から Y の条件付確率は $y = 10, 20$ にそれぞれ $1/2$ の確率を与える.

$x = 1$ のとき, $y = 10, 20, 30$ のとき $f_{X,Y}(1, y) > 0$ となり,

$$f_{Y|X}(10|1) = \frac{f_{X,Y}(1, 10)}{f_X(1)} = \frac{\frac{3}{18}}{\frac{10}{18}} = \frac{3}{10},$$

$$f_{Y|X}(20|1) = \frac{f_{X,Y}(1, 20)}{f_X(1)} = \frac{\frac{4}{18}}{\frac{10}{18}} = \frac{4}{10},$$

$$f_{Y|X}(30|1) = \frac{f_{X,Y}(1, 30)}{f_X(1)} = \frac{\frac{3}{18}}{\frac{10}{18}} = \frac{3}{10},$$

となる. したがって, $X = 1$ という情報から Y の条件付確率は $y = 10, 20, 30$ にそれぞれ $3/10, 4/10, 3/10$ の確率を与える.

$x = 2$ のとき, $y = 30$ のとき $f_{X,Y}(2, y) > 0$ となり,

$$f_{Y|X}(30|2) = \frac{f_{X,Y}(2, 30)}{f_X(2)} = \frac{\frac{4}{18}}{\frac{4}{18}} = 1$$

となる. したがって, $X = 2$ という情報から $Y = 30$ がわかる.

たとえば,

$$\mathbb{P}(Y > 10 | X = 0) = f_{Y|X}(20|0) = \frac{1}{2},$$

$$\mathbb{P}(Y > 10 | X = 1) = f_{Y|X}(20|1) + f_{Y|X}(30|1) = \frac{7}{10},$$

となる。

3.2 連続型確率変数の場合

定義 3.9 (X, Y) は連続型確率ベクトルとし、同時確率密度関数 $f_{X, Y}(x, y)$ および周辺確率密度関数 $f_X(x)$ と $f_Y(y)$ を持つとする。

(i) $f_X(x) > 0$ なる任意の x に対して、 $X = x$ が与えられたときの Y の条件付確率密度関数を $f_{Y|X}(y|x)$ で記し、

$$f_{Y|X}(y|x) = \frac{f_{X, Y}(x, y)}{f_X(x)}$$

で定める。

(ii) $f_Y(y) > 0$ なる任意の y に対して、 $Y = y$ が与えられたときの X の条件付確率密度関数を $f_{X|Y}(x|y)$ で記し、

$$f_{X|Y}(x|y) = \frac{f_{X, Y}(x, y)}{f_Y(y)}$$

で定める。

例 3.4 連続型確率ベクトル (X, Y) は同時確率密度関数

$$f_{X, Y}(x, y) = \begin{cases} e^{-y}, & 0 < x < y < \infty, \\ 0, & \text{その他} \end{cases}$$

を持つとする。 $X = x$ が与えられたときの Y の条件付確率密度関数を求めるために、 X の周辺確率密度関数を求めよう。 $x \leq 0$ の場合、すべての y に対して $f_{X, Y}(x, y) = 0$ なので、 $f_X(x) = 0$ となる。 $x > 0$ の場合、 $y > x$ ならば、 $f_{X, Y}(x, y) > 0$ なので、

$$f_X(x) = \int_{-\infty}^{\infty} f_{X, Y}(x, y) dy = \int_x^{\infty} f_{X, Y}(x, y) dy = e^{-x}$$

となる。したがって

$$f_X(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & \text{その他} \end{cases}$$

となる．これより， $X = x$ が与えられたときの Y の条件付確率密度関数は $x > 0$ の場合のみに定義される．各 $x > 0$ に対して

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}, \quad y > x,$$

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{0}{e^{-x}} = 0, \quad y \leq x$$

となる．

3.3 独立性との関係

注意 3.6 もし， X と Y が独立ならば， x の値に関わらず

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{f_X(x)f_Y(y)}{f_X(x)} = f_Y(y)$$

となる．

4 条件付き期待値

定義 3.10 $X = x$ が与えられたときの Y の条件付確率関数または条件付確率密度関数を $f_{Y|X}(y|x)$ とする． $g: \mathbb{R} \rightarrow \mathbb{R}$ としたとき， $X = x$ が与えられたときの $g(Y)$ の条件付期待値を

$$\mathbb{E}[g(Y)|x] = \begin{cases} \sum_y g(y)f_{Y|X}(y|x), & (\text{離散型}), \\ \int_{-\infty}^{\infty} g(y)f_{Y|X}(y|x) dy, & (\text{連続型}) \end{cases}$$

で定める．ただし，条件付期待値は $\mathbb{E}[|g(Y)||x] < \infty$ のときに存在するものとする．

命題 3.5 (条件付期待値の性質) a_1, a_2, b を定数， $g_1: \mathbb{R} \rightarrow \mathbb{R}$ ， $g_2: \mathbb{R} \rightarrow \mathbb{R}$ とする．

$$(i) \quad \mathbb{E}[a_1g_1(Y) + a_2g_2(Y) + b|x] = a_1\mathbb{E}[g_1(Y)|x] + a_2\mathbb{E}[g_2(Y)|x] + b.$$

$$(ii) \quad g_1(y) \geq 0 \text{ ならば, } \mathbb{E}[g_1(Y)|x] \geq 0.$$

(iii) $a_1 \leq g_1(y) \leq a_2$ ならば, $a_1 \leq \mathbb{E}[g_1(Y)|x] \leq a_2$.

(iv) $\mathbb{E}[g_1(X)g_2(Y)|x] = g_1(x)\mathbb{E}[g_2(Y)|x]$.

が成立する. ただし, 上の条件付期待値はすべて存在するものとする.

証明 積分の性質からわかる. □

定義 3.11 X と Y を確率変数とし, $\mathbb{E}[Y^2] < \infty$ とする. $X = x$ が与えられたときの Y の条件付分散を

$$\text{VAR}[Y|x] = \mathbb{E}[Y^2|x] - \{\mathbb{E}[Y|x]\}^2$$

で定義する. $v(x) := \text{VAR}[Y|x]$ においてとき, $\text{VAR}[Y|X] := v(X)$ で定める. 以上の定義から $\text{VAR}[Y|X] = \mathbb{E}[Y^2|X] - \{\mathbb{E}[Y|X]\}^2$ となることに注意する.

例 3.5 連続型確率ベクトル (X, Y) は同時確率密度関数

$$f_{X,Y}(x, y) = \begin{cases} e^{-y}, & 0 < x < y < \infty, \\ 0, & \text{その他} \end{cases}$$

を持つとき, $X = x$ が与えられたときの Y の条件付確率密度関数は $x > 0$ の場合のみに定義され, 各 $x > 0$ に対して

$$\begin{aligned} f_{Y|X}(y|x) &= \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{e^{-y}}{e^{-x}} = e^{-(y-x)}, & y > x, \\ f_{Y|X}(y|x) &= \frac{f_{X,Y}(x, y)}{f_X(x)} = \frac{0}{e^{-x}} = 0, & y \leq x \end{aligned}$$

であった. $X = x$ ($x > 0$) が与えられたときの Y の条件付期待値は

$$\mathbb{E}[Y|x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy = \int_x^{\infty} y e^{-(y-x)} dy = 1 + x$$

となる. 同様に, $X = x$ ($x > 0$) が与えられたときの Y^2 の条件付期待値は

$$\begin{aligned} \mathbb{E}[Y^2|x] &= \int_{-\infty}^{\infty} y^2 f_{Y|X}(y|x) dy = \int_x^{\infty} y^2 e^{-(y-x)} dy \\ &= \int_0^{\infty} (t+x)^2 e^{-t} dt = x^2 + 2x + 2 \end{aligned}$$

となる．したがって， $X = x (x > 0)$ が与えられたときの Y の条件付分散は

$$\text{VAR}[Y|x] = \mathbb{E}[Y^2|x] - \{\mathbb{E}[Y|x]\}^2 = 1$$

となる．

$X = x$ が与えられたときの $g(Y)$ の条件付期待値 (存在するならば) $\mathbb{E}[g(Y)|x]$ は x の関数であるので， $h(x) = \mathbb{E}[g(Y)|x]$ とおくと，確率変数 $\mathbb{E}[g(Y)|X]$ を $\mathbb{E}[g(Y)|X] = h(X)$ で定めることにする．すなわち， $X = x$ のとき，確率変数 $\mathbb{E}[g(Y)|X]$ の値は $\mathbb{E}[g(Y)|x]$ である．記号の読み方であるが， $\mathbb{E}[g(Y)|X]$ に期待値の記号 \mathbb{E} が使われているが， $\mathbb{E}[g(Y)|X]$ は X に依存する確率変数である．

定理 3.7 X と Y を確率変数とし， Y の期待値は存在するとする．このとき，

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]]$$

が成立する．

証明 X と Y が連続型確率変数で，すべての $x \in \mathbb{R}$ に対して， $f_X(x) > 0$ の場合の証明を与える． (X, Y) の同時確率密度関数を $f_{X,Y}(x, y)$ とする． $X = x$ が与えられたときの Y の条件付確率密度関数 $f_{Y|X}(y|x)$ および条件付期待値の定義から

$$\begin{aligned} \mathbb{E}[Y] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f_{X,Y}(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y \frac{f_{X,Y}(x, y)}{f_X(x)} f_X(x) dx dy \\ &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy \right] f_X(x) dx \\ &= \int_{-\infty}^{\infty} \mathbb{E}[Y|x] f_X(x) dx \\ &= \mathbb{E}[\mathbb{E}[Y|X]] \end{aligned}$$

となる．

□

例 3.6 確率変数 X と Y は同時確率密度関数

$$f_{X,Y}(x,y) = \begin{cases} 2 & x+y < 1, x > 0, y > 0, \\ 0 & (\text{その他}), \end{cases}$$

を持つとする．このとき， X の周辺確率密度関数は， $0 < x < 1$ のとき，

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy = \int_0^{-x+1} 2 dy = 2(1-x)$$

となる．したがって， X の周辺確率密度関数は

$$f_X(x) = \begin{cases} 2(1-x) & 0 < x < 1, \\ 0 & (\text{その他}), \end{cases}$$

である．また， $0 < y < 1$ に対して

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) dx = \int_0^{-y+1} 2 dx = 2(1-y)$$

となることから Y の周辺確率密度関数は

$$f_Y(y) = \begin{cases} 2(1-y) & 0 < y < 1, \\ 0 & (\text{その他}), \end{cases}$$

となる．これから X の二次までの積率を求めると

$$\mathbb{E}[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy = \int_0^1 y 2(1-y) dy = \frac{1}{3},$$

$$\mathbb{E}[Y^2] = \int_{-\infty}^{\infty} y^2 f_Y(y) dy = \int_0^1 y^2 2(1-y) dy = \frac{1}{6},$$

$$\text{VAR}[Y] = \mathbb{E}[Y^2] - \{\mathbb{E}[Y]\}^2 = \frac{1}{18}$$

となる． $0 < x < 1$ としたとき， $X = x$ が与えられたときの Y の条件付確率密度関数は

$$f_{Y|X}(y|x) = \begin{cases} \frac{1}{1-x} & 0 < y < 1-x, \\ 0 & (\text{その他}) \end{cases}$$

となる．したがって， $X = x$ ($0 < x < 1$) が与えられたときの Y の条件付き期待値は

$$\mathbb{E}[Y|x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy = \int_0^{1-x} y \frac{1}{1-x} dy = \frac{1}{1-x} \left[\frac{y^2}{2} \right]_0^{1-x} = \frac{1-x}{2}$$

となる．さらに，

$$\mathbb{E}[\mathbb{E}[Y|X]] = \int_{-\infty}^{\infty} \mathbb{E}[Y|x]f_X(x) dx = \int_0^1 \frac{1-x}{2} 2(1-x) dx = \frac{1}{3}$$

となる．

次に， $X = x$ ($0 < x < 1$) が与えられたときの Y の条件付き分散を求める．
そのために，

$$\begin{aligned} \mathbb{E}[Y^2|x] &= \int_{-\infty}^{\infty} y^2 f_{Y|X}(y|x) dy = \int_0^{1-x} y^2 \frac{1}{1-x} dy = \frac{1}{1-x} \left[\frac{y^3}{3} \right]_0^{1-x} \\ &= \frac{(1-x)^2}{3} \end{aligned}$$

となる．よって， $X = x$ ($0 < x < 1$) が与えられたときの Y の条件付き分散は

$$\text{VAR}[Y|x] = \mathbb{E}[Y^2|x] - \{\mathbb{E}[Y|x]\}^2 = \frac{(1-x)^2}{3} - \frac{(1-x)^2}{4} = \frac{(1-x)^2}{12}$$

となる．

定理 3.8 X, Y を確率変数とし， $\mathbb{E}[Y^2] < \infty$ とする．このとき，

$$\text{VAR}[Y] = \mathbb{E}[\text{VAR}[Y|X]] + \text{VAR}[\mathbb{E}[Y|X]]$$

である．ただし， $h(x) = \text{VAR}[Y|x]$ とおいたとき， $\text{VAR}[Y|X] = h(X)$ で定めた．

証明 まず，

$$\begin{aligned} \text{VAR}[Y] &= \mathbb{E}\{\{Y - \mathbb{E}[Y]\}^2\} = \mathbb{E}\{\{Y - \mathbb{E}[Y|X] + \mathbb{E}[Y|X] - \mathbb{E}[Y]\}^2\} \\ &= \mathbb{E}\{\{Y - \mathbb{E}[Y|X]\}^2\} + \mathbb{E}\{\{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}^2\} \\ &\quad + 2\mathbb{E}\{\{Y - \mathbb{E}[Y|X]\}\{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}\} \end{aligned}$$

に注意する．上の式の最右辺の 3 項目は

$$\mathbb{E}\{\{Y - \mathbb{E}[Y|X]\}\{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}\} = \mathbb{E}[\mathbb{E}\{\{Y - \mathbb{E}[Y|X]\}\{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}|X]]$$

となる．ここで， X が与えられたとき， $\mathbb{E}[Y|X]$ と $\mathbb{E}[Y]$ は定数であることに注意して，上の式の右辺の期待値の中を評価する：

$$\begin{aligned} & \mathbb{E}[\{Y - \mathbb{E}[Y|X]\}\{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}|X] \\ &= \{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}\mathbb{E}[Y - \mathbb{E}[Y|X]|X] \\ &= \{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}\{\mathbb{E}[Y|X] - \mathbb{E}[Y]\} \\ &= 0 \end{aligned}$$

となる．したがって，

$$\mathbb{E}[\{Y - \mathbb{E}[Y|X]\}\{\mathbb{E}[Y|X] - \mathbb{E}[Y]\}] = 0$$

となり，定理は証明された． □

5 2次元の確率変数の変換

(X, Y) を2次元の確率ベクトルとし，実数値関数 $g_1(x, y)$, $g_2(x, y)$ によって定められる新たな確率ベクトル $(U, V) = (g_1(X, Y), g_2(X, Y))$ の分布を求めることを考える． \mathbb{R}^2 の任意の部分集合 B に対して， \mathbb{R}^2 の部分集合 A を

$$A = \{(x, y) \in \mathbb{R}^2 : (g_1(x, y), g_2(x, y)) \in B\}$$

で定めると

$$\mathbb{P}((U, V) \in B) = \mathbb{P}((X, Y) \in A)$$

となる．したがって，固定された g_1, g_2 に対して， (U, V) の分布は (X, Y) の分布のみに依存することがわかる．

5.1 離散型確率ベクトルの場合

(X, Y) が離散型確率ベクトルの場合をまず考えよう． (X, Y) の同時確率関数を $f_{X, Y}(x, y)$ とし， (U, V) の同時確率密度関数 $f_{U, V}(u, v)$ を求めよう．そのために， $S = \{(x, y) \in \mathbb{R}^2 : f_{X, Y}(x, y) > 0\}$ とし，

$$T = \{(u, v) \in \mathbb{R}^2 : \text{ある } (x, y) \in S \text{ に対して, } u = g_1(x, y), v = g_2(x, y)\}$$

とする．すると $(u, v) \in T$ ならば, $f_{U, V}(u, v) > 0$ となり, T は可算集合となることに注意する．任意の $(u, v) \in T$ に対して,

$$S_{uv} = \{(x, y) \in S : g_1(x, y) = u, g_2(x, y) = v\}$$

とおけば,

$$f_{U, V}(u, v) = \mathbb{P}(U = u, V = v) = \mathbb{P}((X, Y) \in S_{uv}) = \sum_{(x, y) \in S_{uv}} f_{X, Y}(x, y)$$

となる．

定理 3.9 X と Y は独立に平均 θ ($\theta > 0$) と λ ($\lambda > 0$) のポアソン分布に従うとする．このとき, $X + Y$ は平均 $\theta + \lambda$ のポアソン分布に従う．

証明 仮定より (X, Y) の同時確率関数は

$$f_{X, Y}(x, y) = \frac{\theta^x e^{-\theta}}{x!} \frac{\lambda^y e^{-\lambda}}{y!}, \quad x = 0, 1, 2, \dots, y = 0, 1, 2, \dots$$

となる．したがって, $S = \{0, 1, 2, \dots\} \times \{0, 1, 2, \dots\}$ である．

いま, $U = X + Y, V = Y$ とおく．すなわち, $g_1(x, y) = x + y, g_2(x, y) = y$ である． $y = v, x = u - v$ から $v = 0, 1, 2, \dots$ かつ $u - v = 0, 1, \dots$ から $u = v, v + 1, v + 2, \dots$ を得る．したがって,

$$\begin{aligned} T &= \{v = 0, 1, 2, \dots; u = v, v + 1, v + 2, \dots\} \\ &= \{u = 0, 1, 2, \dots; v = 0, 1, \dots, u\} \end{aligned}$$

となる．任意の $(u, v) \in T$ に対して, $S_{uv} = \{(u - v, v)\}$ となるので,

$$f_{U, V}(u, v) = f_{X, Y}(u - v, v) = \frac{\theta^{u-v} e^{-\theta}}{(u-v)!} \frac{\lambda^v e^{-\lambda}}{v!},$$

をえる． U の周辺確率関数を求めるために, 固定した非負整数 u を考える．

$$\{U = u\} = \cup_{v: f_{U, V}(u, v) > 0} \{U = u, V = v\}$$

と

$$\{v : f_{U, V}(u, v) > 0\} = \{0, 1, \dots, u\}$$

となる．よって，

$$\begin{aligned} f_U(u) &= \mathbb{P}(U = u) = \mathbb{P}(\cup_{v=0}^u \{U = u, V = v\}) = \sum_{v=0}^u \mathbb{P}(U = u, V = v) \\ &= \sum_{v=0}^u f_{U,V}(u, v) = \sum_{v=0}^u \frac{\theta^{u-v} e^{-\theta}}{(u-v)!} \frac{\lambda^v e^{-\lambda}}{v!} = e^{-(\theta+\lambda)} \sum_{v=0}^u \frac{\theta^{u-v}}{(u-v)!} \frac{\lambda^v}{v!} \\ &= \frac{e^{-(\theta+\lambda)}}{u!} \sum_{v=0}^u \binom{u}{v} \theta^{u-v} \lambda^v = \frac{e^{-(\theta+\lambda)} (\theta + \lambda)^u}{u!} \end{aligned}$$

となり，定理は証明された． \square

5.2 連続型の場合

(X, Y) を連続型確率変数とし，同時確率密度関数 $f_{X,Y}(x, y)$ を持つとする．前節と同様，確率ベクトル (U, V) は $U = g_1(X, Y)$, $V = g_2(X, Y)$ で定義され，

$$S = \{(x, y) \in \mathbb{R}^2 : f_{X,Y}(x, y) > 0\}$$

$$T = \{(u, v) \in \mathbb{R}^2 : \text{ある } (x, y) \in S \text{ に対して, } u = g_1(x, y), v = g_2(x, y)\}$$

であったことを思い出そう．議論を簡単にするために， g_1, g_2 は一対一対応の変換と仮定し，逆変換を

$$x = h_1(u, v), \quad y = h_2(u, v)$$

と書くことにする．さらに，

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u} \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \frac{\partial y}{\partial u}$$

と⁷おく．ただし，

$$\frac{\partial x}{\partial u} = \frac{\partial h_1(u, v)}{\partial u}, \quad \frac{\partial x}{\partial v} = \frac{\partial h_1(u, v)}{\partial v}, \quad \frac{\partial y}{\partial u} = \frac{\partial h_2(u, v)}{\partial u}, \quad \frac{\partial y}{\partial v} = \frac{\partial h_2(u, v)}{\partial v},$$

である．ヤコビアン J は恒等的にゼロではないとする．このとき，

$$f_{U,V}(u, v) = \begin{cases} f_{X,Y}(h_1(u, v), h_2(u, v)) |J|, & (u, v) \in T \\ 0, & (u, v) \notin T \end{cases}$$

⁷これをヤコビアンという．

となる．これは， $(u, v) \in T$ に対して，重積分の変数変換の公式を使えば，

$$\begin{aligned} \frac{\partial^2}{\partial u \partial v} \mathbb{P}(U \leq u, V \leq v) &= \frac{\partial^2}{\partial u \partial v} \int \int_{g_1(x, y) \leq u, g_2(x, y) \leq v} f_{X, Y}(x, y) dx dy \\ &= \frac{\partial^2}{\partial u \partial v} \int_{-\infty}^u \int_{-\infty}^v f_{X, Y}(h_1(s, t), h_2(s, t)) |J| ds dt \end{aligned}$$

となることからわかる．

例 3.7 X と Y は独立に標準正規分布に従っていると仮定し，変換

$$U = X + Y, \quad V = X - Y$$

を考える．したがって， $g_1(x, y) = x + y$ ， $g_2(x, y) = x - y$ である． (X, Y) の同時確率密度関数は

$$f_{X, Y}(x, y) = \frac{1}{2\pi} e^{-(x^2 + y^2)/2}$$

であるので， $S = \mathbb{R}^2$ となる．また， g_1, g_2 は一対一対応で逆変換は

$$x = h_1(u, v) = \frac{u + v}{2}, \quad y = h_2(u, v) = \frac{u - v}{2}$$

となる．したがって， $T = \mathbb{R}^2$ となる．さらに，

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

となるので， (U, V) の同時確率密度関数は

$$\begin{aligned} f_{U, V}(u, v) &= f_{X, Y}(h_1(u, v), h_2(u, v)) |J| \\ &= \frac{1}{2\pi} \exp\left(-\frac{1}{2} \times \frac{(u + v)^2}{4}\right) \exp\left(-\frac{1}{2} \times \frac{(u - v)^2}{4}\right) \frac{1}{2} \\ &= \left(\frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-u^2/4}\right) \left(\frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-v^2/4}\right) \end{aligned}$$

となる．これと補題 3.1 から U と V は独立となる．

定理 3.10 X, Y は独立な確率変数とし， $g_1(x)$ は x のみの関数で $g_2(y)$ は y のみの関数とする．このとき， $U = g_1(X)$ と $V = g_2(Y)$ も独立である．

証明 U, V は連続型の場合を示すことにする．任意の $u \in \mathbb{R}, v \in \mathbb{R}$ に対して，

$$S_u = \{x \in \mathbb{R} : g_1(x) \leq u\}, \quad S_v = \{y \in \mathbb{R} : g_2(y) \leq v\},$$

とおく．このとき， (U, V) の同時分布関数は

$$\begin{aligned} F_{U,V}(u, v) &= \mathbb{P}(U \leq u, V \leq v) \\ &= \mathbb{P}(X \in S_u, Y \in S_v) = \mathbb{P}(X \in S_u)\mathbb{P}(Y \in S_v) \end{aligned}$$

となる．これより (U, V) の同時確率密度関数は

$$f_{U,V}(u, v) = \frac{\partial^2}{\partial u \partial v} F_{U,V}(u, v) = \left(\frac{d}{du} \mathbb{P}(X \in S_u) \right) \left(\frac{d}{dv} \mathbb{P}(Y \in S_v) \right)$$

となる．したがって，補題 3.1 から定理は証明された． \square

6 多次元分布の代表的なモデル

6.1 二変量正規分布

定義 3.12 確率変数 (X, Y) がつぎの同時確率密度関数をもつとき， (X, Y) は母数 $-\infty < \mu_X < \infty, -\infty < \mu_Y < \infty, 0 < \sigma_X < \infty, 0 < \sigma_Y < \infty, -1 < \rho < 1$ の二変量正規分布に従うとする．

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}Q(x, y)\right\}$$

ただし， $-\infty < x < \infty, -\infty < y < \infty$ で

$$Q(x, y) = \frac{(x - \mu_X)^2}{\sigma_X^2} - 2\rho \frac{(x - \mu_X)}{\sigma_X} \frac{(y - \mu_Y)}{\sigma_Y} + \frac{(y - \mu_Y)^2}{\sigma_Y^2}$$

である．

命題 3.6 (二変量正規分布の性質) (i) すべての x, y で $f_{X,Y}(x, y) > 0$ である．
(ii) $\int \int f_{X,Y}(x, y) dx dy = 1$ である．

(iii) X の周辺確率密度関数は

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left\{-\frac{(x - \mu_X)^2}{2\sigma_X^2}\right\}$$

である .

(iv) (X, Y) の積率母関数は , $-\infty < s, t < \infty$ に対して ,

$$M_{X,Y}(s, t) = \exp\left[s\mu_X + t\mu_Y + \frac{1}{2}(s^2\sigma_X^2 + 2st\rho\sigma_X\sigma_Y + t^2\sigma_Y^2)\right]$$

となる .

(v) X と Y の一次と二次の積率は以下ようになる :

$$\mathbb{E}[X] = \mu_X, \quad \mathbb{E}[X^2] = \sigma_X^2 + \mu_X^2, \quad \mathbb{E}[XY] = \rho\sigma_X\sigma_Y + \mu_X\mu_Y$$

となる . これらより

$$\begin{aligned} \text{VAR}[X] &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \sigma_X^2, \\ \text{COV}[X, Y] &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \rho\sigma_X\sigma_Y \end{aligned}$$

を得る .

証明 (i) は明らか . (ii) を示すために ,

$$u = \frac{x - \mu_X}{\sigma_X}, \quad v = \frac{y - \mu_Y}{\sigma_Y}$$

とおく . すると

$$Q(x, y) = u^2 - 2\rho uv + v^2 = (u - \rho v)^2 + (1 - \rho^2)v^2$$

となる . さらに ,

$$w = \frac{u - \rho v}{\sqrt{1 - \rho^2}}$$

とおけば

$$\begin{aligned} &\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{1 - \rho^2}} \exp\left\{-\frac{(u - \rho v)^2}{2(1 - \rho^2)} - \frac{v^2}{2}\right\} du dv \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{w^2}{2}\right) dw \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2}\right) dv = 1 \end{aligned}$$

となる .

(iii) X の周辺確率密度関数は

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f(x, y) dy \\ &= \frac{1}{2\pi\sigma_X\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \frac{1}{2(1-\rho^2)}\left(v - \rho\frac{x-\mu_X}{\sigma_X}\right)^2\right] dv \end{aligned}$$

となり, さらに

$$w = \frac{1}{\sqrt{1-\rho^2}} \left(v - \rho\frac{x-\mu_X}{\sigma_X}\right)$$

とおけば,

$$\begin{aligned} f_X(x) &= \frac{1}{2\pi\sigma_X} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right] \int_{-\infty}^{\infty} e^{-w^2/2} dw \\ &= \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right] \end{aligned}$$

を得る .

(iv) (X, Y) の積率母関数は

$$\begin{aligned} M_{X, Y}(s, t) &= \mathbb{E}[\exp(sX + tY)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(sx + ty) f_{X, Y}(x, y) dx dy \\ &= \exp\left[s\mu_X + t\mu_Y + \frac{1}{2}(s^2\sigma_X^2 + 2st\rho\sigma_X\sigma_Y + t^2\sigma_Y^2)\right] \end{aligned}$$

であること示す . ただし, $-\infty < s, t < \infty$ である .

まず,

$$u = \frac{x - \mu_X}{\sigma_X}, \quad v = \frac{y - \mu_Y}{\sigma_Y}$$

とおく . すると

$$\begin{aligned} M_{X, Y}(s, t) &= \iint_{\mathbb{R}^2} \exp(sx + ty) f_{X, Y}(x, y) dx dy \\ &= \exp\{\mu_X s + \mu_Y t\} \\ &\times \iint_{\mathbb{R}^2} \exp\{s\sigma_X u + t\sigma_Y v\} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{(u-\rho v)^2}{2(1-\rho^2)} - \frac{v^2}{2}\right\} du dv \end{aligned}$$

となる．ここで

$$w = \frac{u - \rho v}{\sqrt{1 - \rho^2}}$$

とおけば，

$$\begin{aligned} & \exp\{-(\mu_X s + \mu_Y t)\} M_{X, Y}(s, t) \\ &= \iint_{\mathbb{R}^2} \exp\{s\sigma_X(\sqrt{1 - \rho^2}w + \rho v) + t\sigma_Y v\} \frac{1}{2\pi} \exp\left\{-\frac{w^2}{2} - \frac{v^2}{2}\right\} dw dv \\ &= \frac{1}{2\pi} \iint_{\mathbb{R}^2} \exp\left\{-\frac{1}{2}\left(w - s\sigma_X\sqrt{1 - \rho^2}\right)^2 - \frac{1}{2}\left(v - (s\rho\sigma_X + t\sigma_Y)\right)^2\right\} dw dv \\ &\quad \times \exp\left\{\frac{1}{2}\left(s^2\sigma_X^2(1 - \rho^2) + (s\rho\sigma_X + t\sigma_Y)^2\right)\right\} \\ &= \exp\left\{\frac{1}{2}\left(s^2\sigma_X^2 + 2st\rho\sigma_X\sigma_Y + t^2\sigma_Y^2\right)\right\} \end{aligned}$$

よりわかる．最後の等号は

$$\frac{1}{2\pi} \iint_{\mathbb{R}^2} \exp\left\{-\frac{1}{2}\left(w - s\sigma_X\sqrt{1 - \rho^2}\right)^2 - \frac{1}{2}\left(v - (s\rho\sigma_X + t\sigma_Y)\right)^2\right\} dw dv = 1$$

は左辺の被積分関数は平均 $N(s\sigma_X\sqrt{1 - \rho^2}, 1)$ と $N(s\rho\sigma_X + t\sigma_Y, 1)$ に独立に従うふたつの確率変数の同時確率密度関数であることに注意すればよい．

(v) X と Y の一次と二次の積率を求める：

$$\begin{aligned} \mathbb{E}[X] &= \left. \frac{\partial}{\partial s} M_{X, Y}(s, t) \right|_{s=0, t=0} = \mu_X, \\ \mathbb{E}[X^2] &= \left. \frac{\partial^2}{\partial s^2} M_{X, Y}(s, t) \right|_{s=0, t=0} = \sigma_X^2 + \mu_X^2, \\ \mathbb{E}[XY] &= \left. \frac{\partial^2}{\partial s \partial t} M_{X, Y}(s, t) \right|_{s=0, t=0} = \rho\sigma_X\sigma_Y + \mu_X\mu_Y \end{aligned}$$

となる．これらより

$$\begin{aligned} \text{VAR}[X] &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \sigma_X^2, \\ \text{COV}[X, Y] &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \rho\sigma_X\sigma_Y \end{aligned}$$

を得る．

□

注意 3.7 ここで

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}$$

とおけば,

$$\frac{1}{1-\rho^2} Q(x, y) = (x - \mu_X, y - \mu_Y)\Sigma^{-1} \begin{pmatrix} x - \mu_X \\ y - \mu_Y \end{pmatrix},$$

$$|\Sigma| = \sigma_X^2\sigma_Y^2(1-\rho^2)$$

となり,

$$f_{X,Y}(x, y) = \frac{1}{2\pi|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(x - \mu_X, y - \mu_Y)\Sigma^{-1} \begin{pmatrix} x - \mu_X \\ y - \mu_Y \end{pmatrix}\right]$$

7 演習問題

問題 3.1 確率変数 X と Y の同時確率関数は下のよな表で与えられるとする.

$x \setminus y$	-1	0	1	4
1	0.12	0.1	0.1	0.08
2	0.1	0.1	0.05	0.04
3	0.08	0.1	0.1	0.03

(i) つぎの確率を求めよ.

- (a) $\mathbb{P}(X > 1, Y \leq 3)$, (b) $\mathbb{P}(X > Y)$, (c) $\mathbb{P}(X = 1, Y < 1)$,
 (d) $\mathbb{P}(X = 1, Y = 0)$, (e) $\mathbb{P}(X = 3)$,

(ii) X と Y の周辺確率関数 $f_X(x)$ と $f_Y(y)$ を求めよ. さらに

$$\sum_x f_X(x) = 1, \quad \sum_y f_Y(y) = 1$$

を確認せよ.

(iii) X と Y は独立でないことを示せ.

問題 3.2 (X, Y) の同時確率密度関数がつぎのように与えられるとする.

$$f_{X,Y}(x, y) = \begin{cases} 4xy, & 0 < x < 1, 0 < y < 1, \\ 0, & \text{その他} \end{cases}$$

(i) つぎの確率を求めよ.

- (a) $\mathbb{P}(0 < X \leq 0.2, 0.5 < Y < 0.8)$, (b) $\mathbb{P}(X \leq Y)$, (c) $\mathbb{P}(X = Y)$,

- (ii)
- X
- と
- Y
- の周辺確率密度関数
- $f_X(x)$
- と
- $f_Y(y)$
- を求めよ。さらに

$$\int_{-\infty}^{\infty} f_X(x) dx = 1, \quad \int_{-\infty}^{\infty} f_Y(y) dy = 1$$

を確認せよ。

問題 3.3 (X, Y) の同時分布関数がつぎのように与えられているとする。

$$F_{X, Y}(x, y) = \begin{cases} \frac{xy(x+y)}{6}, & 0 \leq x \leq 1, 0 \leq y \leq 2, \\ \frac{x(x+2)}{6}, & 0 \leq x \leq 1, y > 2, \\ \frac{y(y+1)}{6}, & x > 1, 0 \leq y \leq 2, \\ 1, & x > 1, y > 2, \\ 0, & \text{その他} \end{cases}$$

- (i) X と Y の周辺分布関数 $F_X(x)$ と $F_Y(y)$ を求めよ。
 (ii) (X, Y) の同時確率密度関数 $f_{X, Y}(x, y)$ を求めよ。さらに

$$\int \int_{\mathbb{R}^2} f_{X, Y}(x, y) dx dy = 1$$

を確認せよ。

- (iii)
- X
- と
- Y
- の周辺確率密度関数
- $f_X(x)$
- と
- $f_Y(y)$
- を求めよ。さらに、

$$\int_{-\infty}^{\infty} f_X(x) dx = 1, \quad \int_{-\infty}^{\infty} f_Y(y) dy = 1$$

を確認せよ。

問題 3.4 (X, Y) の同時分布関数がつぎのように与えられているとする。

$$F_{X, Y}(x, y) = \begin{cases} \frac{cxy}{(1+x)(1+2y)}, & x \geq 0, y \geq 0, \\ 0, & \text{その他} \end{cases}$$

ただし、 c は非負の定数とする。

- (i) c を求めよ。
 (ii) X と Y の周辺分布関数 $F_X(x)$ と $F_Y(y)$ を求めよ。
 (iii) (X, Y) の同時確率密度関数 $f_{X, Y}(x, y)$ を求めよ。さらに

$$\int \int_{\mathbb{R}^2} f_{X, Y}(x, y) dx dy = 1$$

を確認せよ。

- (iv)
- X
- と
- Y
- の周辺確率密度関数
- $f_X(x)$
- と
- $f_Y(y)$
- を求めよ。さらに、

$$\int_{-\infty}^{\infty} f_X(x) dx = 1, \quad \int_{-\infty}^{\infty} f_Y(y) dy = 1$$

を確認せよ。

問題 3.5

離散型確率変数 X, Y の同時確率関数 $f_{X, Y}(x, y)$ は下の表のように与えられているとする。

$x \backslash y$	-1	0	1
-1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$
0	$\frac{3}{16}$	0	$\frac{3}{16}$
1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$

- (i) X と Y の周辺確率関数 $f_X(x)$ と $f_Y(y)$ を求めよ。
- (ii) X の周辺分布関数 $F_X(x) = \mathbb{P}(X \leq x)$ ($y \in \mathbb{R}$) を求め、 $F_X(x)$ のグラフを作図せよ。
- (iii) X, Y, XY の期待値 $\mathbb{E}(X), \mathbb{E}(Y), \mathbb{E}(XY)$ を計算せよ。
- (iv) X と Y の共分散 $\text{COV}[X, Y]$ を求めよ。
- (v) X と Y は独立か従属かを調べよ。さらに、その理由を述べること。

問題 3.6

確率変数 X と Y の同時確率関数は下のような表で与えられるとする。

4	1 / 16	1 / 16	1 / 16	4 / 16
3	1 / 16	1 / 16	3 / 16	
2	1 / 16	2 / 16		
1	1 / 16			
y / x	1	2	3	4

- (i) X と Y の周辺確率関数 $f_X(x)$ と $f_Y(y)$ を求めよ。さらに

$$\sum_x f_X(x) = 1, \quad \sum_y f_Y(y) = 1$$

を確認せよ。

- (ii) $T := XY$ と $S := X + Y$ のそれぞれの周辺確率関数 $f_T(t)$ と $f_S(s)$ を求めよ。さらに

$$\sum_t f_T(t) = 1, \quad \sum_s f_S(s) = 1$$

を確認せよ。

- (iii) $\mathbb{E}[XY], \mathbb{E}[X + Y]$ を計算せよ。
- (iv) $\mathbb{E}[X], \mathbb{E}[Y]$ を計算せよ。
- (v) $X = x$ が与えられたときの Y の条件付き確率関数 $f_{Y|X}(y|x)$ と求めよ。 $X = x$ が与えられたときの Y の条件付き確率関数 $f_{Y|X}(y|x)$ が定義される各 x に対してさらに

$$\sum_y f_{Y|X}(y|x) = 1$$

を確認せよ。

(vi) $g(X) = \mathbb{E}[Y|X]$ を求めよ．更に，

$$\mathbb{E}[g(X)] = \mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}[Y]$$

を確認せよ．

問題 3.7 つぎの関数が同時確率関数となるように定数 c を定め， $\mathbb{P}(X = 1, Y > 1)$ ， $\mathbb{P}(X = Y)$ を求めよ．

$$(a) \quad f_{X,Y}(x, y) = \begin{cases} cx, & x = 1, 2, 3; y = 1, 2, \\ 0, & \text{その他} \end{cases}$$

$$(a) \quad f_{X,Y}(x, y) = \begin{cases} c|x + y|, & x = -1, 0, 1, 3; y = -1, 0, 2, \\ 0, & \text{その他} \end{cases}$$

問題 3.8 連続型確率変数 X, Y は同時確率密度関数

$$f_{X,Y}(x, y) = \begin{cases} x + y & (0 < x < 1, 0 < y < 1), \\ 0 & (\text{その他}), \end{cases}$$

を持つとする．

- (i) X の周辺確率密度関数 $f_X(x)$ を求めよ．なお，解答には $f_X(x) > 0$ となる x の範囲を明示すること．
- (ii) X の期待値 $\mathbb{E}[X]$ ，分散 $\text{VAR}[X]$ ， XY の期待値 $\mathbb{E}[XY]$ および X と Y の共分散 $\text{COV}[X, Y]$ を求めよ．ただし，最終的な数値を求めるために必要な分数の足し算や引き算の計算（通分）はしなくともよい．
- (iii) $X = x$ が与えられたときの Y の条件付確率密度関数 $f_{Y|X}(y|x)$ を求めよ．ただし，解答には $f_{Y|X}(y|x)$ が定義される x の範囲と $f_{Y|X}(y|x) > 0$ となる y の範囲を明示すること．
- (iv) $X = x$ が与えられたときの Y の条件付期待値 $\mathbb{E}[Y|x]$ を求めよ．
- (v) $U = X + Y$ と $V = X - Y$ とおいたとき， U と V の同時確率密度関数 $f_{U,V}(u, v)$ を求めよ．ただし， $f_{U,V}(u, v) > 0$ となる u と v の範囲を明示すること．
- (vi) V の周辺確率密度関数 $f_V(v)$ を求めよ．ただし， $f_V(v) > 0$ となる v の範囲を明示せよ．

問題 3.9 $0 \leq a \leq 1/2$ とし，確率変数 X と Y は -1 と 1 の値を取り，

$$\mathbb{P}(X = -1, Y = -1) = \mathbb{P}(X = 1, Y = 1) = a$$

と

$$\mathbb{P}(X = -1, Y = 1) = \mathbb{P}(X = 1, Y = -1) = \frac{1}{2} - a$$

をする．このとき，以下の問いに答えよ．

- (i) $\mathbb{E}[XY]$ を計算せよ．
- (ii) X の周辺確率、 X の期待値 $\mathbb{E}[X]$ と分散 $\text{VAR}[X]$ を求めよ．
- (iii) $X = 1$ が与えられたときの Y の条件付き確率と条件付き期待値 $\mathbb{E}[Y|X = 1]$ を求めよ．
- (iv) X と Y の相関係数 $\rho(X, Y)$ を求め， X と Y が無相関になるよう a を定めよ．
- (v) 上の問いで求めた a のとき， X と Y が独立かを調べよ．

問題 3.10 確率変数 (X, Y) の確率密度関数

$$f_{X,Y}(x, y) = \begin{cases} c(x+y) & (x, y) \in (0, 1) \times (0, 1), \\ 0 & (\text{その他}), \end{cases}$$

で与えられているとする。このとき、以下の問いに答えよ。

- (i) 確率密度関数の性質を満たすよう c を定めよ。

以下は (i) で求めた c の値を使うことにする。

- (ii) X の周辺確率密度関数 $f_X(x)$ を求めよ。さらに

$$\int_{-\infty}^{\infty} f_X(x) dx = 1$$

を確認せよ。

- (iii) $\mathbb{E}[XY]$, $\mathbb{E}[X+Y]$ を計算せよ。

- (iv) $\mathbb{E}[X]$, $\mathbb{E}[Y]$ を計算せよ。

- (v) $X = x$ が与えられたときの Y の条件付き確率密度関数 $f_{Y|X}(y|x)$ と求めよ。さらに、 $X = x$ が与えられたときの Y の条件付き確率密度関数 $f_{Y|X}(y|x)$ される各 x に対して

$$\int_{-\infty}^{\infty} f_{Y|X}(y|x) dy = 1$$

を確認せよ。

- (vi) $g(X) = \mathbb{E}[Y|X]$ を求めよ。更に、求めた $g(X)$ について等式

$$\mathbb{E}[g(X)] = \mathbb{E}[\mathbb{E}[Y|X]] = \mathbb{E}[Y]$$

が成立することを確認せよ。(一般に正しいことがわかるが、具体的にこの問題の同時確率密度関数について成立することを確認すること)

問題 3.11 X と Y の同時確率密度関数を

$$f_{X,Y}(x, y) = \begin{cases} 2 & 0 < x < y, 0 < y < 1, \\ 0 & (\text{その他}), \end{cases}$$

とする。

- (i) $\text{COV}[X, Y]$ を計算せよ。

- (ii) X の周辺確率密度関数 $f_X(x)$ を求めよ。さらに

$$\int_{-\infty}^{\infty} f_X(x) dx = 1$$

を確認せよ。

- (iii) $X = x$ が与えられたときの Y の条件付き確率密度関数 $f_{Y|X}(y|x)$ と求めよ。さらに、 $X = x$ が与えられたときの Y の条件付き確率密度関数 $f_{Y|X}(y|x)$ される各 x に対して

$$\int_{-\infty}^{\infty} f_{Y|X}(y|x) dy = 1$$

を確認せよ。

問題 3.12 X を離散型確率変数とし、その確率関数が

$$f_X(x) = \frac{x}{3} \quad x = 1, 2$$

で与えられ、 $X = x$ が与えられたときの Y の条件付き確率関数が

$$f_{Y|X}(y|x) = P[Y = y|X = x] = \binom{x}{y} \left(\frac{1}{2}\right)^x \quad y = 0, \dots, x$$

で与えられるとする。ただし、

$$\binom{x}{y} = \frac{x!}{y!(x-y)!}$$

である。

- (i) $\mathbb{E}[X]$ と $\text{VAR}[X]$ を計算せよ。
- (ii) $\mathbb{E}[Y]$ を計算せよ。
- (iii) X と Y の同時確率関数 $f_{X,Y}(x, y)$ を求めよ。さらに

$$\sum_x \sum_y f_{X,Y}(x, y) = 1$$

を確認せよ。

問題 3.13 X と Y の同時確率密度関数が

$$f_{X,Y}(x, y) = \begin{cases} 4xy - 2x - 2y + 2 & (0 < x < 1, 0 < y < 1), \\ 0 & \text{その他} \end{cases}$$

で与えられとする。

- (i) Y の平均と分散を求めよ。
- (ii) X の周辺確率密度関数 $f_X(x)$ を求めよ。さらに

$$\int_{-\infty}^{\infty} f_X(x) dx = 1$$

を確認せよ。

- (iii) X と Y の共分散と相関係数を求めよ。
- (iv) $\hat{Y} = \mathbb{E}[Y|X]$ を求めよ (ヒント: 答えは X の関数となる。また、 $\mathbb{E}[Y|X]$ は $g(x) = \mathbb{E}[Y|x]$ としたとき、 $\mathbb{E}[Y|X] = g(X)$ で定める。)
- (v) $\tilde{Y} = a + bX$ (a と b は定数) としてとき、 $\mathbb{E}\{\tilde{Y} - Y\}^2$ を最小にする a と b を $\mathbb{E}[X]$ 、 $\mathbb{E}[Y]$ 、 $\mathbb{E}[XY]$ 、および $\mathbb{E}[X^2]$ で表現した後、それを求めよ。

問題 3.14 X と Y は互いに独立な確率変数でそれぞれは正規分布 $N(0, 1)$ と $N(0, 4)$ に従うとする。

$$U = X + Y, \quad V = X - Y$$

とおく。

- (i) U と V の同時確率密度関数 $f_{U,V}(u, v)$ を求めよ。さらに

$$\int \int_{\mathbb{R}^2} f_{U,V}(u, v) du dv = 1$$

を確認せよ。

(ii) U と V の分散共分散行列

$$\begin{pmatrix} \text{VAR}[U] & \text{COV}[U, V] \\ \text{COV}[U, V] & \text{VAR}[V] \end{pmatrix}$$

の逆行列を求めよ.

問題 3.15 X と Y 互いに独立な確率変数でそれぞれは標準正規分布に従うとし,

$$U = \frac{1}{\sqrt{2}}(X + Y), \quad V = \frac{1}{\sqrt{2}}(X - Y)$$

とおく.

(i) U と V の期待値と U と V の分散共分散行列

$$\begin{pmatrix} \text{VAR}[U] & \text{COV}[U, V] \\ \text{COV}[U, V] & \text{VAR}[V] \end{pmatrix}$$

を求めよ.

(ii) U と V の同時確率密度関数 $f_{U, V}(u, v)$ を求めよ. さらに

$$\int \int_{\mathbb{R}^2} f_{U, V}(u, v) du dv = 1$$

を確認せよ.

問題 3.16 確率変数 X と Y は同時確率密度関数

$$f_{X, Y}(x, y) = \begin{cases} x + y, & 0 < x < 1, 0 < y < 1, \\ 0, & \text{その他} \end{cases}$$

を持つとする.

(i) $U = X + Y, V = X$ とおく. U と V の同時確率密度関数 $f_{U, V}(u, v)$ を求めよ. さらに

$$\int \int_{\mathbb{R}^2} f_{U, V}(u, v) du dv = 1$$

を確認せよ.

(ii) U の周辺確率密度関数 $f_U(u)$ を求めよ. さらに

$$\int_{-\infty}^{\infty} f_U(u) du = 1$$

を確認せよ.

(iii) $Z = XY$ とおく. Z と X の同時確率密度関数 $f_{Z, X}(z, x)$ を求めよ. さらに

$$\int \int_{\mathbb{R}^2} f_{Z, X}(z, x) dz dx = 1$$

を確認せよ.

(iv) Z の周辺確率密度関数 $f_Z(z)$ を求めよ. さらに

$$\int_{-\infty}^{\infty} f_Z(z) dz = 1$$

を確認せよ.

- (v)
- $T = X/Y$
- とおく.
- T
- と
- Y
- の同時確率密度関数
- $f_{T, Y}(t, y)$
- を求めよ. さらに

$$\int \int_{\mathbb{R}^2} f_{T, Y}(t, y) dt dy = 1$$

を確認せよ.

- (vi)
- T
- の周辺確率密度関数
- $f_T(t)$
- を求めよ. さらに

$$\int_{-\infty}^{\infty} f_T(t) dt = 1$$

を確認せよ.

問題 3.17 X と Y は独立同一の分布に従い, それぞれは確率密度関数

$$f(x) = \begin{cases} e^{-x}, & x > 0, \\ 0, & x \leq 0 \end{cases}$$

を持つとする.

- (i)
- $V = X + Y, U = Y$
- としたとき,
- U
- と
- V
- の同時確率密度関数
- $f_{U, V}(u, v)$
- を求めよ. さらに

$$\int \int_{\mathbb{R}^2} f_{U, V}(u, v) du dv = 1$$

を確認せよ.

- (ii)
- V
- の周辺確率密度関数
- $f_V(v)$
- を求めよ. さらに

$$\int_{-\infty}^{\infty} f_V(v) dv = 1$$

を確認せよ.

- (iii)
- $W = X/Y$
- とおく.
- W
- と
- Y
- の同時確率密度関数
- $f_{W, Y}(w, y)$
- を求めよ. さらに

$$\int \int_{\mathbb{R}^2} f_{W, Y}(w, y) dw dy = 1$$

を確認せよ.

- (iv)
- W
- の周辺確率密度関数
- $f_W(w)$
- を求めよ. さらに

$$\int_{-\infty}^{\infty} f_W(w) dw = 1$$

を確認せよ.

問題 3.18 連続型確率変数 X, Y は同時確率密度関数

$$f_{X, Y}(x, y) = \begin{cases} \frac{1}{8}(2 + x + y + \beta xy) & (-1 < x < 1, -1 < y < 1) \\ 0 & (\text{その他}) \end{cases} \quad (3.3)$$

をもつとする. ただし, β は $0 \leq \beta \leq 1/2$ なる定数とする. また, $0 \leq \beta \leq 1/2$ かつ $-1 < x < 1, -1 < y < 1$ ならば, $f_{X, Y}(x, y) \geq 0$ となっていることに注意せよ. この事実の証明は不要とする. このとき, 以下の問いに答えよ.

- (i)
- X
- と
- Y
- の周辺確率密度関数
- $f_X(x)$
- と
- $f_Y(y)$
- をそれぞれ求めよ. さらに

$$\int_{-\infty}^{\infty} f_X(x) dx = 1, \quad \int_{-\infty}^{\infty} f_Y(y) dy = 1,$$

を確認せよ.

- (ii) XY の期待値 $\mathbb{E}[XY]$ を求めよ.
- (iii) X と Y が無相関になるときの β の値を求めよ.
- (iv) X と Y が無相関になるとき (すなわち, X と Y の同時確率密度関数が問い (iii) で求めた β の値をもつ (3.3) で与えられるとき), X と Y は独立であるかどうかを調べよ.

問題 3.19 取りうる値の集合が $\{1, 2, 3\}$ である確率変数 X と Y の同時確率関数 $f_{X,Y}(x, y)$ が以下のように与えられているとする. 以下の問いに答えよ.

$Y \setminus X$	1	2	3
1	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{3}{36}$
2	$\frac{1}{36}$	$\frac{10}{36}$	$\frac{3}{36}$
3	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$

- (i)
- X
- と
- Y
- の周辺確率関数
- $f_X(x)$
- と
- $f_Y(y)$
- を求めよ. さらに

$$\sum_x f_X(x) = 1, \quad \sum_y f_Y(y) = 1$$

を確認せよ.

- (ii)
- $Y = y (y = 1, 2, 3)$
- が与えられたときの
- X
- の条件付確率関数
- $f_{X|Y}(x|y)$
- を求めよ. さらに,
- $Y = y (y = 1, 2, 3)$
- が与えられたときの
- X
- の条件付確率関数
- $f_{X|Y}(x|y)$
- が定義される各
- y
- に対して

$$\sum_x f_{X|Y}(x|y) = 1$$

を確認せよ.

- (iii) $Y = y (y = 1, 2, 3)$ が与えられたときの X の条件付期待値 $\mathbb{E}[X|y]$ を求めよ.
- (iv) $g(y) = \mathbb{E}[X|y]$ とおいたとき, $g(Y)$ の期待値 $\mathbb{E}[g(Y)]$ を求めよ.

問題 3.20 離散型確率変数 X, Y の同時確率関数 $f_{X,Y}(x, y)$ は下の表のように与えられているとする.

$x \setminus y$	-1	0	1
-1	α	β	α
0	β	0	β
1	α	β	α

ただし, α, β は定数で $\alpha > 0, \beta > 0$ かつ $\alpha + \beta = \frac{1}{4}$ である.

- (i) X と Y の周辺確率関数 $f_X(x)$ と $f_Y(y)$ を求めよ.
- (ii) X の周辺分布関数 $F_X(x) = \mathbb{P}(X \leq x)$ を求め, $F_X(x)$ のグラフを作図せよ.
- (iii) X, Y, XY の期待値 $\mathbb{E}(X), \mathbb{E}(Y), \mathbb{E}(XY)$ を計算せよ.
- (iv) X と Y の共分散 $\text{COV}[X, Y]$ を求めよ.
- (v) X と Y は独立か従属かを調べよ. さらに, その理由を述べること.

第4章 標本分布論と漸近分布論

ある母数をもつ母集団分布から無作為標本を抽出したとき，無作為標本の関数である統計量の分布を学ぶ．

1 標本分布論の枠組み

1.1 ランダム標本

定義 4.1 X_1, X_2, \dots, X_n が母集団分布 $f(x)$ からの標本の大きさが n のランダム標本であるとは， X_1, X_2, \dots, X_n は互いに独立な確率変数列であって，各 $X_i, i = 1, 2, \dots, n$, は確率関数または確率密度関数 $f(x)$ に従うときをいう．また， X_1, X_2, \dots, X_n は独立同一に確率関数または確率密度関数 $f(x)$ に従うともいう．

注意 4.1 多くの場合は $n > 1$ である．また， $X = (X_1, X_2, \dots, X_n)$ の同時確率関数または同時確率密度関数は

$$f_{\mathbf{X}}(x_1, x_2, \dots, x_n) = \prod_{k=1}^n f(x_k)$$

である．

1.2 統計量と標本分布

定義 4.2 X_1, X_2, \dots, X_n をある母集団分布からの標本の大きさが n のランダム標本とし， $T(x_1, x_2, \dots, x_n)$ をランダム標本 $X = (X_1, X_2, \dots, X_n)$ の値域上で定義された実数値または実ベクトル値関数とする．このとき，確率変数または確率ベクトル $Y = T(X_1, X_2, \dots, X_n)$ を統計量という．さらに，統計量 Y の確率分布を Y の標本分布とよぶ．

例 4.1 ランダム標本の算術平均は統計量であり，標本平均という．通常

$$\bar{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i$$

と記す．

また，

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

で定義される統計量を標本分散という．

補題 4.1 x_1, x_2, \dots, x_n を実数列とし， $\bar{x}_n = (1/n)(x_1 + x_2 + \cdots + x_n)$ と $s_n^2 = (1/(n-1)) \sum_{i=1}^n (x_i - \bar{x}_n)^2$ とおく．このとき，

$$(1) \quad \min_a \sum_{i=1}^n (x_i - a)^2 = \sum_{i=1}^n (x_i - \bar{x}_n)^2.$$

$$(2) \quad (n-1)s_n^2 = \sum_{i=1}^n x_i^2 - n\bar{x}_n^2.$$

となる．

証明 (1) を証明するために， \bar{x}_n を加えて引けば，

$$\begin{aligned} \sum_{i=1}^n (x_i - a)^2 &= \sum_{i=1}^n (x_i - \bar{x}_n + \bar{x}_n - a)^2 \\ &= \sum_{i=1}^n (x_i - \bar{x}_n)^2 + 2 \sum_{i=1}^n (x_i - \bar{x}_n)(\bar{x}_n - a) + \sum_{i=1}^n (\bar{x}_n - a)^2 \\ &= \sum_{i=1}^n (x_i - \bar{x}_n)^2 + n(\bar{x}_n - a)^2 \end{aligned} \quad (4.1)$$

となる．最後の等式は

$$\sum_{i=1}^n (x_i - \bar{x}_n)(\bar{x}_n - a) = (\bar{x}_n - a) \sum_{i=1}^n (x_i - \bar{x}_n) = 0$$

よりわかる．(4.1) は $a = \bar{x}_n$ の時に最小になることがわかる．

(2) を示すためには，(4.1) において， $a = 0$ とすればよい． □

命題 4.1 X_1, X_2, \dots, X_n をある母集団分布からの標本の大きさが n のランダム標本とし, $g_i (i = 1, 2, \dots, n)$ を X_i の値域上で定義された実数値関数とする. $g_i^2(X_i)$ の期待値が存在するとき,

$$\mathbb{E}\left[\sum_{i=1}^n g_i(X_i)\right] = \sum_{i=1}^n \mathbb{E}[g_i(X_i)], \quad (4.2)$$

$$\text{VAR}\left[\sum_{i=1}^n g(X_i)\right] = \sum_{i=1}^n \text{VAR}[g_i(X_i)] \quad (4.3)$$

が成立する.

証明 (4.2) は期待値の線形性と

$$\mathbb{E}\left[\sum_{i=1}^n g_i(X_i)\right] = \sum_{i=1}^n \mathbb{E}[g_i(X_i)]$$

からわかる¹.

¹連続型の場合を考える. X_1, X_2, \dots, X_n の同時確率密度関数を $f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n)$ としたとき,

$$\begin{aligned} \mathbb{E}\left[\sum_{i=1}^n g(X_i)\right] &= \int \cdots \int \sum_{i=1}^n g(x_i) f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) \Pi_{i=1}^n dx_i \\ &= \sum_{i=1}^n \int \cdots \int g(x_i) f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) \Pi_{i=1}^n dx_i \\ &= \sum_{i=1}^n \int g(x_i) f_{X_i}(x_i) dx_i \\ &= \mathbb{E}[g(X_i)] \end{aligned}$$

よりわかる. ただし, $f_{X_i}(x_i)$ は X_i の周辺確率密度関数である.

(4.3) を示すために、分散の定義と期待値の線形性から

$$\begin{aligned}
 \text{VAR}\left[\sum_{i=1}^n g_i(X_i)\right] &= \mathbb{E}\left[\left\{\sum_{i=1}^n g_i(X_i) - \mathbb{E}\left[\sum_{i=1}^n g_i(X_i)\right]\right\}^2\right] \\
 &= \mathbb{E}\left[\left\{\sum_{i=1}^n (g_i(X_i) - \mathbb{E}[g_i(X_i)])\right\}^2\right] \\
 &= \mathbb{E}\left[\sum_{i=1}^n (g_i(X_i) - \mathbb{E}[g_i(X_i)])^2 + \sum_{i \neq j} (g_i(X_i) - \mathbb{E}[g_i(X_i)])(g_j(X_j) - \mathbb{E}[g_j(X_j)])\right] \\
 &= \sum_{i=1}^n \mathbb{E}[(g_i(X_i) - \mathbb{E}[g_i(X_i)])^2] \\
 &\quad + \sum_{i \neq j} \mathbb{E}\{[g_i(X_i) - \mathbb{E}[g_i(X_i)]]\}[g_j(X_j) - \mathbb{E}[g_j(X_j)]]\} \quad (4.4)
 \end{aligned}$$

となる。しかし、

$$\sum_{i=1}^n \mathbb{E}[(g_i(X_i) - \mathbb{E}[g_i(X_i)])^2] = \sum_{i=1}^n \text{VAR}[g_i(X_i)]$$

と $i \neq j$ に対して、

$$\begin{aligned}
 &\mathbb{E}[(g_i(X_i) - \mathbb{E}[g_i(X_i)])(g_j(X_j) - \mathbb{E}[g_j(X_j)])] \\
 &= \mathbb{E}\{g_i(X_i) - \mathbb{E}[g_i(X_i)]\}\mathbb{E}\{g_j(X_j) - \mathbb{E}[g_j(X_j)]\} = 0
 \end{aligned}$$

となる。ただし、最後の等号は定理 3.1 からわかる。上のふたつの式を (4.4) に代入すれば、(2) は示される。□

系 4.1 X_1, X_2, \dots, X_n をある母集団分布からの標本の大きさが n のランダム標本とする。 X_i^2 の期待値が存在するとき、

$$\begin{aligned}
 \mathbb{E}\left[\sum_{i=1}^n a_i X_i\right] &= \sum_{i=1}^n a_i \mathbb{E}[X_i], \\
 \text{VAR}\left[\sum_{i=1}^n a_i X_i\right] &= \sum_{i=1}^n a_i^2 \text{VAR}[X_i]
 \end{aligned}$$

が成立する．ただし， a_1, a_2, \dots, a_n は定数である．

命題 4.2 $n \geq 2$ とする． X_1, X_2, \dots, X_n を平均 μ と分散 $\sigma^2 < \infty$ の母集団分布からの標本の大きさが n のランダム標本とする．このとき，

$$(1) \quad \mathbb{E}[\bar{X}_n] = \mu,$$

$$(2) \quad \text{VAR}[\bar{X}_n] = \frac{\sigma^2}{n},$$

$$(3) \quad \mathbb{E}[S_n^2] = \sigma^2.$$

となる．

証明 (1) を示すために，命題 4.1 において， $g_i(x) = x/n$ ($i = 1, 2, \dots, n$) とすれば，

$$\mathbb{E}[\bar{X}_n] = \frac{1}{n} \mathbb{E}\left[\sum_{i=1}^n X_i\right] = \frac{1}{n} n \mathbb{E}[X_1] = \mathbb{E}[X_1] = \mu$$

となることがわかる．

(2) は分散の性質と定理 を同様に利用すれば，

$$\text{VAR}[\bar{X}_n] = \frac{1}{n^2} \text{VAR}\left[\sum_{i=1}^n X_i\right] = \frac{1}{n^2} n \text{VAR}[X_1] = \frac{1}{n} \text{VAR}[X_1] = \frac{\sigma^2}{n}$$

となることがわかる．

(3) を示すために補題 4.1 を使えば，

$$\begin{aligned} \mathbb{E}[S_n^2] &= \mathbb{E}\left[\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2\right] = \frac{1}{n-1} \mathbb{E}\left[\sum_{i=1}^n X_i^2 - n\bar{X}_n^2\right] \\ &= \frac{1}{n-1} (n\mathbb{E}[X_1^2] - n\mathbb{E}[\bar{X}_n^2]) \end{aligned} \quad (4.5)$$

となる．しかし，

$$\mathbb{E}[X_1^2] = \text{VAR}[X_1] + (\mathbb{E}[X_1])^2 = \sigma^2 + \mu^2,$$

$$\mathbb{E}[\bar{X}_n^2] = \text{VAR}[\bar{X}_n] + (\mathbb{E}[\bar{X}_n])^2 = \frac{\sigma^2}{n} + \mu^2$$

となる．これらと (4.5) をあわせれば，

$$\mathbb{E}[S_n^2] = \frac{1}{n-1} \left(n(\sigma^2 + \mu^2) - n \left(\frac{\sigma^2}{n} + \mu^2 \right) \right) = \sigma^2$$

となることがわかる． □

命題 4.3 X_1, X_2, \dots, X_n を積率母関数 $M_X(t)$ を持つ母集団分布からの標本の大きさが n のランダム標本とする．このとき，

$$M_{\bar{X}_n}(t) = (M_X(t/n))^n$$

が成立する．

証明

$$M_{\bar{X}_n}(t) = \mathbb{E}[e^{t\bar{X}_n}] = \mathbb{E} \left[\prod_{i=1}^n e^{tX_i/n} \right] = \prod_{i=1}^n \mathbb{E} [e^{tX_i/n}] = [M_X(t/n)]^n$$

からわかる． □

例 4.2 X_1, X_2, \dots, X_n を正規分布 $N(\mu, \sigma^2)$ からの標本の大きさが n のランダム標本とする．このとき，標本平均 \bar{X}_n は正規分布 $N(\mu, \sigma^2/n)$ に従うことがわかる．なぜならば，

$$M_{\bar{X}_n}(t) = \exp \left[n \left(\frac{t}{n} \mu + \frac{1}{2} \left(\frac{t}{n} \right)^2 \sigma^2 \right) \right] = \exp \left[t\mu + \frac{(\sigma^2/n)t^2}{2} \right]$$

からわかる．

例 4.3 X_1, X_2, \dots, X_n を母数 p のベルヌーイ分布からの標本の大きさが n のランダム標本とする．ただし， $0 < p < 1$ である．このとき，標本平均 $n\bar{X}_n$ は母数 n と p の二項分布に従うことがわかる．なぜならば，

$$M_{n\bar{X}_n}(t) = \prod_{i=1}^n \mathbb{E}[e^{tX_i}] = [pe^t + (1-p)]^n$$

からわかる．

2 正規分布からのランダム標本

命題 4.4 X_1, X_2, \dots, X_n を正規分布 $N(\mu, \sigma^2)$ からの標本の大きさが n のランダム標本とし, $\bar{X}_n = (1/n) \sum_{i=1}^n X_i$ と $S_n^2 = (1/(n-1)) \sum_{i=1}^n (X_i - \bar{X}_n)^2$ とおく. このとき, 以下が成立する:

- (1) \bar{X}_n と S_n^2 は独立である.
- (2) \bar{X}_n は正規分布 $N(\mu, \sigma^2/n)$ に従う.
- (3) $(n-1)S_n^2/\sigma^2$ は自由度 $n-1$ のカイ自乗分布に従う.

証明 (2) は例 4.2 からわかる. 次に, (1) を示す. 各 i に対して, $Y_i = (X_i - \mu)/\sigma$ とすれば, Y_i は正規分布 $N(0, 1)$ に従い, $(\bar{X}_n - \mu)/\sigma = \bar{Y}_n = (1/n) \sum_{i=1}^n Y_i$ と $S_n^2/\sigma^2 = (1/(n-1)) \sum_{i=1}^n (Y_i - \bar{Y}_n)^2$ となるので, 一般性を失わず, X_i は正規分布 $N(0, 1)$ に従うとして, \bar{X}_n と S_n^2 の独立性を示せばよいことがわかる.

$X_i - \bar{X}_n$ は正規分布 $N(0, (1-1/n))$ に従うことがわかる. なぜならば,

$$\begin{aligned}
 M_{X_i - \bar{X}_n}(t) &= \mathbb{E} \left[\exp \left(t(1-1/n)X_i - \sum_{j \neq i} (t/n)X_j \right) \right] \\
 &= \mathbb{E}[e^{t(1-1/n)X_i}] \prod_{j \neq i} \mathbb{E}[e^{-(t/n)X_j}] \\
 &= \exp \left(\frac{t^2(1-(1/n))^2}{2} \right) \prod_{j \neq i} \exp \left(\frac{(t/n)^2}{2} \right) \\
 &= \exp \left(\frac{t^2(1-(1/n))^2}{2} \right) \times \exp \left((n-1) \frac{(t/n)^2}{2} \right) \\
 &= \exp \left(\frac{t^2(1-(1/n))^2}{2} + (n-1) \frac{(t/n)^2}{2} \right) \\
 &= \exp \left\{ \frac{t^2}{2} \left(1 - \frac{2}{n} + \frac{1}{n^2} + \frac{n-1}{n^2} \right) \right\} \\
 &= \exp \left\{ \left(1 - \frac{1}{n} \right) \frac{t^2}{2} \right\}
 \end{aligned}$$

からわかる. \bar{X}_n と $X_i - \bar{X}_n$ はともに正規分布に従うので, \bar{X}_n と $X_i - \bar{X}_n$ が独立であることをいうためには, $\text{COV}[\bar{X}_n, X_i - \bar{X}_n] = 0$ を示せばよい. し

かし,

$$\begin{aligned}
 \text{COV}[\bar{X}_n, X_i - \bar{X}_n] &= \mathbb{E}[\bar{X}_n(X_i - \bar{X}_n)] \\
 &= \mathbb{E}[(1/n) \sum_{j=1}^n X_j X_i] - \mathbb{E}[\bar{X}_n^2] \\
 &= \frac{1}{n} \sum_{j \neq i} \mathbb{E}[X_i X_j] + \frac{1}{n} \mathbb{E}[X_i^2] - \mathbb{E}[\bar{X}_n^2] \\
 &= \frac{1}{n} \text{VAR}[X_i] - \text{VAR}[\bar{X}_n] = \frac{1}{n} - \frac{1}{n} = 0
 \end{aligned}$$

となる。よって、 \bar{X}_n と $X_i - \bar{X}_n$ が独立である。これから \bar{X}_n と $(X_1 - \bar{X}_n, X_2 - \bar{X}_n, \dots, X_n - \bar{X}_n)$ は独立²となり、 \bar{X}_n と S_n^2 は独立であることがわかる。

(3) を示すために,

$$\sum_{i=1}^n (X_i - \mu)^2 = \sum_{i=1}^n (X_i - \bar{X}_n)^2 + n(\bar{X}_n - \mu)^2$$

に注意する。 $Y_i = (X_i - \mu)/\sigma$, $i = 1, 2, \dots, n$, とおけば、 $\bar{Y}_n = (1/n) \sum_{i=1}^n Y_i = (\bar{X}_n - \mu)/\sigma$ となり、 Y_i と \bar{Y}_n は $N(0, 1)$ と $N(0, 1/n)$ に従う。したがって、 $U = \sum_{i=1}^n (Y_i - \bar{Y}_n)^2$ が自由度 $n - 1$ のカイ自乗分布に従うことを示せばよい。いま、 $W = \sum_{i=1}^n Y_i^2$, $V = n\bar{Y}_n^2$ とおけば、 $t < 1/2$ に対して、 W, V の積率母関数は

$$\begin{aligned}
 M_W(t) &= \mathbb{E}[\exp(tW)] = \left(\frac{1}{1 - 2t} \right)^{n/2}, \\
 M_V(t) &= \mathbb{E}[\exp(tV)] = \left(\frac{1}{1 - 2t} \right)^{1/2}
 \end{aligned}$$

となる。さらに、 \bar{Y}_n と $(Y_1 - \bar{Y}_n, Y_2 - \bar{Y}_n, \dots, Y_n - \bar{Y}_n)$ は独立であることに注意して、 $U = \sum_{i=1}^n (Y_i - \bar{Y}_n)^2$ の積率母関数を求める： $t < 1/2$ のとき、

$$\begin{aligned}
 M_W(t) &= \mathbb{E}[\exp(tW)] = \mathbb{E}[\exp\{t \sum_{i=1}^n (Y_i - \bar{Y}_n)^2\} + tn\bar{Y}_n^2] \\
 &= \mathbb{E}[\exp\{tU + tV\}] = \mathbb{E}[\exp\{tU\}]\mathbb{E}[\exp\{tV\}] = M_U(t)M_V(t)
 \end{aligned}$$

²この議論は正規分布に従っていることから保障される。

より

$$M_U(t) = \frac{M_W(t)}{M_V(t)} = \left(\frac{1}{1-2t} \right)^{(n-1)/2}$$

がわかる . □

2.1 t 分布と F 分布

X_1, X_2, \dots, X_n が正規分布 $N(\mu, \sigma^2)$ からの標本の大きさが n のランダム標本としたとき ,

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \quad (4.6)$$

は標準正規分布 $N(0, 1)$ に従うことが定理 4.4 (2) からわかる . σ が既知であれば , \bar{X}_n を観測したときに , (4.6) は μ の推測に利用できる . しかし , σ が未知のときは , (4.6) の代わりに

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \quad (4.7)$$

を μ の推測に用いる . ただし , $S_n^2 = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ で S_n は S_n^2 の正の平方根である . (4.7) の標本分布を求めるために

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} = \frac{(\bar{X}_n - \mu)/(\sigma/\sqrt{n})}{\sqrt{S_n^2/\sigma^2}} \quad (4.8)$$

と書き換える . (4.8) の分子は標準正規分布 $N(0, 1)$ に従い , 分母は $\sqrt{\chi_{n-1}^2/(n-1)}$ と同じ分布で , 分母と分子は独立である . ただし , χ_{n-1}^2 は自由度 $(n-1)$ のカイ自乗分布である . したがって , (4.8) の分布は $V/\sqrt{U/(n-1)}$ と同じ分布である . ただし , U と V は独立に自由度 $(n-1)$ のカイ自乗分布と標準正規分布に従うものとする .

定義 4.3 p を自然数としたとき , 確率変数 T が自由度 p の t 分布に従うとは , T が確率密度関数

$$f_T(t) = \frac{\Gamma((p+1)/2)}{\Gamma(p/2)} \frac{1}{\sqrt{p\pi}} \frac{1}{(1+t^2/p)^{(p+1)/2}}, \quad -\infty < t < \infty$$

を持つときをいう .

注意 4.2 $f_T(t)$ は密度関数であるためには,

$$\int_{-\infty}^{\infty} f_T(t) dt = 1$$

でなければならない. $f_T(t)$ は偶関数なので,

$$2 \int_0^{\infty} \frac{\Gamma((p+1)/2)}{\Gamma(p/2)} \frac{1}{\sqrt{p\pi}} \frac{1}{(1+t^2/p)^{(p+1)/2}} dt = 1$$

を示せばよい. これを確認する. いま,

$$u \Big|_1^0 = \frac{1}{1 + \frac{t^2}{p}} \Big|_0^{\infty} \iff t^2 = \frac{p}{u} - p = \frac{p(1-u)}{u}$$

とおく. すると

$$2t dt = -\frac{p}{u^2} du$$

となるので,

$$dt = -\frac{1}{2t} \frac{p}{u^2} du = -\frac{1}{2} \frac{p}{u^2} \sqrt{\frac{u}{p(1-u)}} du = -\frac{\sqrt{p}}{2} u^{-3/2} (1-u)^{-1/2} du$$

となる. これらから

$$\begin{aligned} 2 \int_0^{\infty} f_T(t) dt &= 2 \frac{1}{\sqrt{p\pi}} \frac{\Gamma(\frac{p+1}{2})}{\Gamma(\frac{p}{2})} \int_0^{\infty} \frac{dt}{\left(1 + \frac{t^2}{p}\right)^{\frac{p+1}{2}}} \\ &= \frac{\Gamma(\frac{p+1}{2})}{\sqrt{\pi}\Gamma(\frac{p}{2})} \int_0^1 u^{(p+1)/2} u^{-3/2} (1-u)^{-1/2} du \\ &= \frac{\Gamma(\frac{p+1}{2})}{\sqrt{\pi}\Gamma(\frac{p}{2})} \int_0^1 u^{p/2-1} (1-u)^{1/2-1} du \\ &= \frac{\Gamma(\frac{p+1}{2})}{\sqrt{\pi}\Gamma(\frac{p}{2})} B\left(\frac{p}{2}, \frac{1}{2}\right) \\ &= \frac{\Gamma(\frac{p+1}{2})}{\sqrt{\pi}\Gamma(\frac{p}{2})} \frac{\Gamma(\frac{p}{2}) \Gamma(\frac{1}{2})}{\Gamma(\frac{p+1}{2})} = 1 \end{aligned}$$

がわかる.

命題 4.5 (t 分布の確率密度関数の導出について) $X_1, X_2, \dots, X_n, n \geq 2$, は正規分布 $N(\mu, \sigma^2)$ からの標本の大きさが n のランダム標本とするととき,

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n}}$$

は自由度 $n - 1$ の t 分布に従う.

証明

$$U = (n-1) \frac{S_n^2}{\sigma^2}, \quad V = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}, \quad p = n-1$$

とおくと U と V は独立で, それぞれは自由度 p のカイ自乗分布 χ_p^2 と標準正規分布 $N(0, 1)$ に従い,

$$T = \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} = \frac{V}{\sqrt{U/p}}$$

となる. したがって, U と V から出発して, $\sqrt{p}V/\sqrt{U}$ の確率密度関数を求める. まず,

$$f_{U,V}(u, v) = \frac{1}{\sqrt{2\pi}} e^{-v^2/2} \frac{1}{\Gamma(p/2)2^{p/2}} u^{(p/2)-1} e^{-u/2}, \quad -\infty < v < \infty; 0 < u < \infty$$

に注意する. いま

$$t = \frac{v}{\sqrt{u/p}}, \quad w = u$$

とおくと

$$J = \begin{vmatrix} (\partial v/\partial t) & (\partial v/\partial w) \\ (\partial u/\partial t) & (\partial u/\partial w) \end{vmatrix} = \begin{vmatrix} \sqrt{\frac{w}{p}} & * \\ 0 & 1 \end{vmatrix} = \sqrt{\frac{w}{p}}$$

から

$$\begin{aligned} f_T(t) &= \int_0^\infty f_{U,V}\left(w, t\sqrt{\frac{w}{p}}\right) \sqrt{\frac{w}{p}} dw \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma(p/2)2^{p/2}} \int_0^\infty e^{-(1/2)t^2 w/p} w^{(p/2)-1} e^{-w/2} \sqrt{\frac{w}{p}} dw \\ &= \frac{1}{\sqrt{2\pi}} \frac{1}{\Gamma(p/2)2^{p/2} p^{1/2}} \int_0^\infty e^{-(1/2)(1+t^2/p)w} w^{(p+1)/2-1} dw \end{aligned}$$

となる．さらに，

$$z = \left(1 + \frac{t^2}{p}\right)w$$

とおけば，

$$\begin{aligned} f_T(t) &= \frac{1}{\Gamma(p/2)\sqrt{p\pi}} \frac{1}{(1+t^2/p)^{(p+1)/2}} \frac{1}{2^{(p+1)/2}} \int_0^\infty z^{(p+1)/2-1} e^{-z/2} dz \\ &= \frac{\Gamma((p+1)/2)}{\Gamma(p/2)} \frac{1}{\sqrt{p\pi}} \frac{1}{(1+t^2/p)^{(p+1)/2}} \end{aligned}$$

を得る． □

定義 4.4 p, q を自然数としたとき，確率変数 F が自由度 p と q の F 分布に従うとは， F が確率密度関数

$$f_F(x) = \frac{\Gamma((p+q)/2)}{\Gamma(p/2)\Gamma(q/2)} \left(\frac{p}{q}\right)^{p/2} \frac{x^{(p/2)-1}}{(1+(p/q)x)^{(p+q)/2}}, \quad 0 < x < \infty$$

を持つときをいう．

注意 4.3 $f_F(x)$ は密度関数であるためには，

$$\int_0^\infty f_F(x) dx = 1$$

でなければならない．この積分を計算するために，

$$y = \frac{p}{q}x$$

とおく．すると

$$\begin{aligned} \int_0^\infty f_F(x) dx &= \frac{\Gamma((p+q)/2)}{\Gamma(p/2)\Gamma(q/2)} \left(\frac{p}{q}\right)^{p/2} \int_0^\infty \frac{\left(\frac{q}{p}y\right)^{(p/2)-1}}{(1+y)^{(p+q)/2}} \frac{q}{p} dy \\ &= \frac{\Gamma((p+q)/2)}{\Gamma(p/2)\Gamma(q/2)} \int_0^\infty \frac{y^{(p/2)-1}}{(1+y)^{(p+q)/2}} dy \end{aligned}$$

となる．つぎに，

$$z = \frac{y}{1+y} \iff y = \frac{z}{1-z}, \quad dy = \frac{1}{(1-z)^2} dz$$

とおく.

$$\begin{aligned}
 \int_0^\infty f_F(x) dx &= \frac{\Gamma((p+q)/2)}{\Gamma(p/2)\Gamma(q/2)} \int_0^1 \frac{\left(\frac{z}{1-z}\right)^{(p/2)-1}}{\left(1+\frac{z}{1-z}\right)^{(p+q)/2}} \frac{1}{(1-z)^2} dz \\
 &= \frac{\Gamma((p+q)/2)}{\Gamma(p/2)\Gamma(q/2)} \int_0^1 \left(\frac{z}{1-z}\right)^{(p/2)-1} \frac{1}{\left(\frac{1}{1-z}\right)^{(p+q)/2}} \frac{1}{(1-z)^2} dz \\
 &= \frac{\Gamma((p+q)/2)}{\Gamma(p/2)\Gamma(q/2)} \int_0^1 z^{(p/2)-1} (1-z)^{(q/2)-1} dz \\
 &= \frac{\Gamma((p+q)/2)}{\Gamma(p/2)\Gamma(q/2)} B\left(\frac{p}{2}, \frac{q}{2}\right) \\
 &= \frac{\Gamma((p+q)/2)}{\Gamma(p/2)\Gamma(q/2)} \frac{\Gamma(p/2)\Gamma(q/2)}{\Gamma((p+q)/2)} = 1
 \end{aligned}$$

がわかる.

命題 4.6 $n \geq 2$ と $m \geq 2$ とする. X_1, X_2, \dots, X_n は正規分布 $N(\mu_X, \sigma_X^2)$ からの標本の大きさが n のランダム標本とし, Y_1, Y_2, \dots, Y_m は $X_i, i = 1, 2, \dots, n$, とは独立な正規分布 $N(\mu_Y, \sigma_Y^2)$ からの標本の大きさが m のランダム標本とし,

$$\begin{aligned}
 S_X^2 &= \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2, \quad \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i, \quad S_Y^2 = \frac{1}{m-1} \sum_{i=1}^m (Y_i - \bar{Y}_m)^2, \\
 \bar{Y}_m &= \frac{1}{m} \sum_{i=1}^m Y_i,
 \end{aligned}$$

とする. このとき,

$$F = \frac{(S_X^2/\sigma_X^2)}{(S_Y^2/\sigma_Y^2)}$$

は自由度 $n-1$ と $m-1$ の F 分布に従う.

証明

$$U = (n-1) \frac{S_X^2}{\sigma_X^2}, \quad V = (m-1) \frac{S_Y^2}{\sigma_Y^2}, \quad p = n-1, \quad q = m-1$$

とおく．命題 4.4 より U と V は独立に自由度 p と q のカイ自乗分布に従う．
 $u > 0, v > 0$ に対して，

$$f_{U,V}(u, v) = \frac{1}{\Gamma(p/2)2^{p/2}} u^{p/2-1} e^{-u/2} \frac{1}{\Gamma(q/2)2^{q/2}} v^{q/2-1} e^{-v/2}$$

である．いま，

$$W = \frac{U/p}{V/q}, \quad Z = V$$

なる変換を考える．

$$\begin{cases} w = \frac{u/p}{v/q} \\ z = v \end{cases} \implies \begin{cases} u = \frac{p}{q} w z \\ v = z \end{cases}$$

となり，

$$J = \begin{vmatrix} \frac{\partial u}{\partial w} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial w} & \frac{\partial v}{\partial z} \end{vmatrix} = \begin{vmatrix} \frac{p}{q} z & \frac{p}{q} w \\ 0 & 1 \end{vmatrix} = \frac{p}{q} z$$

となる． $w > 0, z > 0$ に対して，

$$\begin{aligned} f_{W,Z}(w, z) &= \frac{1}{\Gamma(p/2)\Gamma(q/2)2^{\frac{1}{2}(p+q)}} \left(\frac{p}{q}\right)^{\frac{p}{2}-1} w^{\frac{p}{2}-1} z^{\frac{p}{2}-1} z^{\frac{q}{2}-1} \\ &\quad \times \exp\left(-\frac{p}{2q} w z\right) e^{-z/2} \frac{p}{q} z \\ &= \frac{(p/q)^{\frac{p}{2}} w^{\frac{p}{2}-1}}{\Gamma(p/2)\Gamma(q/2)2^{(1/2)(p+q)}} z^{(1/2)(p+q)-1} \exp\left[-\frac{z}{2}\left(1 + \frac{p}{q} w\right)\right] \end{aligned}$$

となる．したがって，

$$\begin{aligned} f_W(w) &= \int_0^\infty f_{W,Z}(w, z) dz \\ &= \frac{(p/q)^{\frac{p}{2}} w^{\frac{p}{2}-1}}{\Gamma(p/2)\Gamma(q/2)2^{\frac{1}{2}(p+q)}} \int_0^\infty z^{\frac{1}{2}(p+q)-1} \exp\left[-\frac{z}{2}\left(1 + \frac{p}{q} w\right)\right] dz \end{aligned}$$

となる．ここで，

$$\frac{z}{2}\left(1 + \frac{p}{q} w\right) = t \implies z = 2t\left(1 + \frac{p}{q} w\right)^{-1}, \quad dz = 2\left(1 + \frac{p}{q} w\right)^{-2} dt$$

とおく．これより

$$\begin{aligned} f_W(w) &= \frac{(p/q)^{\frac{p}{2}} w^{\frac{p}{2}-1}}{\Gamma(p/2)\Gamma(q/2)2^{\frac{1}{2}(p+q)}} 2^{\frac{1}{2}(p+q)-1} \left(1 + \frac{p}{q}w\right)^{-\frac{1}{2}(p+q)+1} \\ &\quad \times 2 \left(1 + \frac{p}{q}w\right)^{-1} \int_0^\infty t^{\frac{1}{2}(p+q)-1} e^{-t} dt \\ &= \frac{\Gamma(\frac{1}{2}(p+q))(p/q)^{\frac{p}{2}}}{\Gamma(p/2)\Gamma(q/2)} \times \frac{w^{\frac{p}{2}-1}}{\left(1 + \frac{p}{q}w\right)^{\frac{1}{2}(p+q)}} \end{aligned}$$

よりわかる．

□

3 順序統計量

定義 4.5 X_1, X_2, \dots, X_n をランダム標本としたとき，これを小さい順に並べかえたものを

$$X_{(1)}, X_{(2)}, \dots, X_{(n)}$$

を記し，これらを順序統計量という．すなわち，

$$\begin{aligned} X_{(1)} &= \min\{X_1, X_2, \dots, X_n\}, \\ X_{(2)} &= X_1, X_2, \dots, X_n \text{ の中で 2 番目に小さいもの}, \\ &\vdots \\ X_{(n)} &= \max\{X_1, X_2, \dots, X_n\}, \end{aligned}$$

である．

命題 4.7 X_1, X_2, \dots, X_n を離散型分布 $f_X(x_i) = p_i, i = 1, 2, \dots$ からの標本の大きさが n のランダム標本とする．ただし， $x_1 < x_2 < \dots$ は X の台³とする．さらに， $P_i = \sum_{j=1}^i p_j, i = 1, 2, \dots, P_0 = 0$ とおく． $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ を標本の順序統計量としたとき，

$$P(X_{(j)} \leq x_i) = \sum_{k=j}^n \binom{n}{k} P_i^k (1 - P_i)^{n-k}, \quad j = 1, 2, \dots, n, i = 1, 2, \dots, \quad (4.9)$$

³ $f_X(x) > 0$ なる点．

と

$$P(X_{(j)} = x_i) = \sum_{k=j}^n \binom{n}{k} [P_i^k (1 - P_i)^{n-k} - P_{i-1}^k (1 - P_{i-1})^{n-k}], \quad (4.10)$$

となる.

証明 $i \in \{1, 2, \dots\}$ を固定し, Y を確率変数とし,

$$Y = \#\{X_j, j = 1, 2, \dots, n \mid X_j \leq x_i\}$$

とする. ここで, 確率変数 Z_j を $\{X_j \leq x_i\}$ が起こったとき, $Z_j = 1$, さもなければ, $Z_j = 0$ と定める. X_1, X_2, \dots, X_n は同一分布に従うので, 各 j について,

$$\mathbb{P}(Z_j = 1) = P_i = \mathbb{P}(X_j \leq x_i)$$

である. また, Z_1, Z_2, \dots, Z_n は独立である. さらに, $Y = \sum_{j=1}^n Z_j$ に注意すれば, Y は母数 n, P_i の二項分布 $Bin(n, P_i)$ に従うことがわかる.

事象 $\{X_{(j)} \leq x_i\}$ は事象 $\{Y \geq j\}$ と同じなので,

$$\mathbb{P}(X_{(j)} \leq x_i) = \mathbb{P}(Y \geq j)$$

となり, (4.9) は示された. (4.10) を示すためには,

$$\mathbb{P}(X_{(j)} = x_i) = \mathbb{P}(X_{(j)} \leq x_i) - \mathbb{P}(X_{(j)} \leq x_{i-1})$$

を考えればよい. $i = 1$ の場合は $\mathbb{P}(X_{(j)} = x_i) = \mathbb{P}(X_{(j)} \leq x_i)$ よりわかる. 以上から (4.10) を示された. \square

命題 4.8 X_1, X_2, \dots, X_n を分布関数 $F_X(x)$ と確率密度関数 $f_X(x)$ をもつ連続型分布からの標本の大きさが n のランダム標本とする. $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ を標本の順序統計量としたとき, $X_{(j)}, j = 1, 2, \dots, n$, の確率密度関数は

$$f_{X_{(j)}}(x) = \frac{n!}{(j-1)!(n-j)!} f_X(x) [F_X(x)]^{j-1} [1 - F_X(x)]^{n-j}, \quad (4.11)$$

となる.

証明 $X_{(j)}$ の分布関数を求め、その導関数を計算することにより、 $X_{(j)}$ の確率密度関数を求めることにする。

実数 x を固定する。 Y を確率変数とし、

$$Y = \#\{X_j, j = 1, 2, \dots, n \mid X_j \leq x\}$$

とする。ここで、確率変数 Z_j を $\{X_j \leq x\}$ が起こったとき、 $Z_j = 1$ 、さもなければ、 $Z_j = 0$ と定める。 X_1, X_2, \dots, X_n は同一分布に従うので、各 j について、 $\mathbb{P}(Z_j = 1) = F_X(x)$ である。また、 Z_1, Z_2, \dots, Z_n は独立である。さらに、 $Y = \sum_{j=1}^n Z_j$ に注意すれば、 Y は母数 n 、 P_i の二項分布 $Bin(n, F_X(x))$ に従うことがわかる。これらより

$$F_{X_{(j)}}(x) = \mathbb{P}(Y \geq j) = \sum_{k=j}^n \binom{n}{k} [F_X(x)]^k [1 - F_X(x)]^{n-k}$$

となる。したがって、 $X_{(j)}$ の確率密度関数は

$$\begin{aligned} f_{X_{(j)}}(x) &= \frac{d}{dx} F_{X_{(j)}}(x) \\ &= \sum_{k=j}^n \binom{n}{k} \left(k [F_X(x)]^{k-1} [1 - F_X(x)]^{n-k} f_X(x) \right. \\ &\quad \left. - (n-k) [F_X(x)]^k [1 - F_X(x)]^{n-k-1} f_X(x) \right) \\ &= \binom{n}{j} j f_X(x) [F_X(x)]^{j-1} [1 - F_X(x)]^{n-j} \\ &\quad + \sum_{k=j+1}^n \binom{n}{k} k [F_X(x)]^{k-1} [1 - F_X(x)]^{n-k} f_X(x) \\ &\quad - \sum_{k=j}^{n-1} \binom{n}{k} (n-k) [F_X(x)]^k [1 - F_X(x)]^{n-k-1} f_X(x) \\ &= \frac{n!}{(j-1)!(n-j)!} f_X(x) [F_X(x)]^{j-1} [1 - F_X(x)]^{n-j} \\ &\quad + \sum_{k=j}^{n-1} \binom{n}{k+1} (k+1) [F_X(x)]^k [1 - F_X(x)]^{n-k-1} f_X(x) \\ &\quad - \sum_{k=j}^{n-1} \binom{n}{k} (n-k) [F_X(x)]^k [1 - F_X(x)]^{n-k-1} f_X(x) \end{aligned}$$

となる．最後に，

$$\binom{n}{k+1}(k+1) = \frac{n!}{k!(n-k-1)!} = \binom{n}{k}(n-k)$$

から (4.11) は示される． \square

例 4.4 X_1, X_2, \dots, X_n を $(0, 1)$ 上の一様分布からの標本の大きさが n のランダム標本とする．すなわち，

$$F_X(x) = \begin{cases} x, & 0 < x < 1, \\ 0, & x \leq 0, \\ 1, & x \geq 1 \end{cases},$$

$$f_X(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{その他,} \end{cases}$$

である．このとき，(4.11) から

$$f_{X_{(j)}} = \begin{cases} \frac{n!}{(j-1)!(n-j)!} x^{j-1} (1-x)^{n-j}, & 0 < x < 1, \\ 0, & \text{その他} \end{cases}, \quad j = 1, 2, \dots, n$$

となる．また，補遺の代表的な広義積分 (iii) から

$$\begin{aligned} \mathbb{E}[X_{(j)}] &= \frac{n!}{(j-1)!(n-j)!} \int_0^1 x^j (1-x)^{n-j} dx \\ &= \frac{n!}{(j-1)!(n-j)!} \frac{\Gamma(j+1)\Gamma(n-j+1)}{\Gamma(n+2)} \\ &= \frac{n!}{(j-1)!(n-j)!} \frac{j!(n-j)!}{(n+1)!} \\ &= \frac{j}{n+1} \end{aligned}$$

となる．

命題 4.9 X_1, X_2, \dots, X_n を分布関数 $F_X(x)$ と確率密度関数 $f_X(x)$ をもつ連続型分布からの標本の大きさが n のランダム標本とする． $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ を標本の順序統計量としたとき， $X_{(i)}$ と $X_{(j)}$ ， $1 \leq i < j \leq n$ ，の同時確率密度

関数は

$$f_{X_{(i)}, X_{(j)}}(u, v) = \begin{cases} \frac{n!}{(i-1)!(j-1-i)!(n-j)!} f_X(u) f_X(v) [F_X(u)]^{i-1} \\ \quad \times [F_X(v) - F_X(u)]^{j-i-1} [1 - F_X(v)]^{n-j}, & -\infty < u < v < \infty, \\ 0, & (\text{その他}) \end{cases}$$

となる .

証明 証明は省略する .

□

系 4.2 $X_{(1)}$ と $X_{(n)}$ の同時確率密度関数は

$$f_{X_{(1)}, X_{(n)}}(x_1, x_n) = \begin{cases} n(n-1)[F_X(x_n) - F_X(x_1)]^{n-2} f_X(x_1) f(x_n) & x_1 < x_n, \\ 0 & x_1 \geq x_n \end{cases}$$

で与えられる .

例 4.5 $X_1, X_2, \dots, X_n, n \geq 2$ は独立同一に $(0, 1)$ 上の一様分布に従うとする . すなわち , 各 $X_i, i = 1, 2, \dots, n$ は確率密度関数と分布関数

$$f_X(x) = \begin{cases} 1 & 0 < x < 1, \\ 0 & (\text{その他}) \end{cases}, \quad F_X(x) = \begin{cases} 0 & x \leq 0, \\ x & 0 < x < 1, \\ 1 & x \geq 1 \end{cases}$$

を持つ . 命題 4.9 から $X_{(1)}$ と $X_{(n)}$ の同時確率密度関数は

$$f_{X_{(1)}, X_{(n)}}(u, v) = \begin{cases} n(n-1)(v-u)^{n-2} & 0 < u < v < 1, \\ 0 & \text{その他} \end{cases}$$

となる .

これより, $0 < u < 1$ と $0 < v < 1$ に対して,

$$\begin{aligned} f_{X_{(1)}}(u) &= \int_u^1 n(n-1)(v-u)^{n-2} dv \\ &= n(n-1) \left[\frac{(v-u)^{n-1}}{n-1} \right]_u^1 = n(1-u)^{n-1}, \\ f_{X_{(n)}}(v) &= n(n-1) \int_0^v (v-u)^{n-2} du = n(n-1) \left[-\frac{(v-u)^{n-1}}{n-1} \right]_0^v \\ &= nv^{n-1}, \end{aligned}$$

を得る. さらに,

$$\begin{aligned} &\int_{0 < u < v < 1} n(n-1)(v-u)^{n-2} du dv \\ &= n(n-1) \int_0^1 \left\{ \int_u^1 (v-u)^{n-2} dv \right\} du \\ &= n(n-1) \int_0^1 \left\{ \int_0^v (v-u)^{n-2} du \right\} dv = 1 \end{aligned}$$

となる.

4 確率変数の列の収束について

以下では, 特に断りがない限り $\{X_n\}_{n=1}^{\infty}$ を確率変数列とし, X を確率変数とし, これらは同一の確率空間上で定義されているとする.

定義 4.6 (確率収束) $n \rightarrow \infty$ のとき, $\{X_n\}$ が X に確率収束するとは, 任意の正数 ϵ に対して

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \epsilon) = 0$$

をみたすときをいい, $X_n \xrightarrow{P} X$ と記す.

定義 4.7 (分布収束) $\{X_n\}_{n=1}^{\infty}$, X は確率変数とし, $F_X(x)$ を X の分布関数とする. $n \rightarrow \infty$ のとき, $\{X_n\}$ が X に分布収束するとは, $F_X(x)$ の任意の連続点において,

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n \leq x) = F_X(x)$$

をみたすときをいい, $X_n \xrightarrow{d} X$ と記す.

注意 4.4 $X_n \xrightarrow{d} F_X(x)$ のように記すこともある．また， X が正規分布 $N(0, \sigma^2)$ に従うとき， $X_n \xrightarrow{d} N(0, \sigma^2)$ と記すこともある．また， $F_X(x)$ のことを X_n の極限分布という．

例 4.6 X_1, X_2, \dots, X_n は独立同一に $[0, 1]$ 上の一様分布に従うとし，

$$M_n = \max\{X_1, X_2, \dots, X_n\}$$

とする．直感的には， $n \rightarrow \infty$ のとき， M_n は 1 に近づくことがわかるであろう．これはつぎのことから保障される．まず， M_n の分布関数は

$$F_{M_n}(x) = \begin{cases} 0 & (x < 0), \\ x^n & (0 \leq x \leq 1), \\ 1 & (x > 1), \end{cases}$$

と⁴なる．したがって，十分小さな正数 ϵ に対して，

$$\begin{aligned} \mathbb{P}(|M_n - 1| > \epsilon) &= \mathbb{P}(\{M_n < 1 - \epsilon\} \cup \{M_n > 1 + \epsilon\}) \\ &\leq \mathbb{P}(M_n < 1 - \epsilon) + \mathbb{P}(M_n > 1 + \epsilon) \\ &= \mathbb{P}(M_n < 1 - \epsilon) = (1 - \epsilon)^n \rightarrow 0 \quad (n \rightarrow \infty), \end{aligned}$$

がわかる⁵．

次に， $n(1 - M_n)$ の極限分布を求めよう： $x \geq 0$ と十分大きな正の整数 n に対して，

$$\begin{aligned} \mathbb{P}(n(1 - M_n) \leq x) &= \mathbb{P}(M_n \geq 1 - x/n) = 1 - \mathbb{P}(M_n < 1 - x/n) \\ &= 1 - \left(1 - \frac{x}{n}\right)^n \\ &\rightarrow 1 - e^{-x} \quad (n \rightarrow \infty), \end{aligned}$$

⁴ $0 \leq x \leq 1$ に対して， $\mathbb{P}(X_1 \leq x) = x$ になることに注意すれば，

$$\mathbb{P}(M_n \leq x) = \mathbb{P}(\cap_{i=1}^n \{X_i \leq x\}) = \prod_{i=1}^n \mathbb{P}(X_i \leq x) = x^n$$

となることからわかる．

⁵下から 2 番目の等号は

$$\mathbb{P}(M_n > 1 + \epsilon) = 1 - \mathbb{P}(M_n \leq 1 + \epsilon) = 1 - F_{M_n}(1 + \epsilon) = 1 - 1 = 0$$

よりわかる．また， $|1 - \epsilon| < 1$ より最後の極限は求まる．

となる．また， $x < 0$ のときは $P(n(1 - M_n) \leq x) = 0$ となる．したがって，

$$n(1 - M_n) \xrightarrow{d} F_X(x)$$

を得る．ただし，

$$F_X(x) = \begin{cases} 1 - e^{-x} & (x \geq 0), \\ 0 & (x < 0), \end{cases}$$

である．すなわち，母数 1 の指数分布に分布収束することがわかる．

以下では確率変数列の収束に関する重要な命題を証明するための補題である．

補題 4.2 $\{A_n\}_{n=1}^{\infty}$ と $\{B_n\}_{n=1}^{\infty}$ を事象の列とする． $n \uparrow \infty$ のとき，

$$\mathbb{P}(A_n) \rightarrow 1, \quad \mathbb{P}(B_n) \rightarrow 1$$

ならば，

$$\mathbb{P}(A_n \cap B_n) \rightarrow 1$$

が成立する．

証明 $\mathbb{P}(A_n^c) = 1 - \mathbb{P}(A_n) \rightarrow 0 (n \rightarrow \infty)$ に注意すれば，

$$\mathbb{P}\{(A_n \cap B_n)^c\} = \mathbb{P}(A_n^c \cup B_n^c) \leq \mathbb{P}(A_n^c) + \mathbb{P}(B_n^c) \rightarrow 0, \quad (n \rightarrow \infty)$$

よりわかる． □

命題 4.10 $\{X_n\}_{n=1}^{\infty}$ と $\{Y_n\}_{n=1}^{\infty}$ を確率変数列で

$$X_n \xrightarrow{P} c, \quad Y_n \xrightarrow{P} d, \quad (n \rightarrow \infty)$$

を満足するとする．ただし， c と d は定数とする．このとき，

- (i) $X_n + Y_n \xrightarrow{P} c + d$
- (ii) $X_n Y_n \xrightarrow{P} cd$
- (iii) $d \neq 0$ ならば， $X_n / Y_n \xrightarrow{P} c/d$

が成立する .

証明 (i) の証明 . $|(X_n + Y_n) - (c + d)| \leq |X_n - c| + |Y_n - d|$ から , どんな正の数 $\epsilon > 0$ に対しても

$$|X_n - c| \leq \frac{\epsilon}{2} \quad \text{かつ} \quad |Y_n - d| \leq \frac{\epsilon}{2} \quad (4.12)$$

ならば ,

$$|(X_n + Y_n) - (c + d)| \leq \epsilon$$

であるので

$$\{|X_n - c| \leq \frac{\epsilon}{2}\} \cap \{|Y_n - d| \leq \frac{\epsilon}{2}\} \subset \{|(X_n + Y_n) - (c + d)| \leq \epsilon\}$$

より , $n \uparrow \infty$ のとき ,

$$\mathbb{P}\{|(X_n + Y_n) - (c + d)| \leq \epsilon\} \geq \mathbb{P}\{|X_n - c| \leq \frac{\epsilon}{2}\} \cap \{|Y_n - d| \leq \frac{\epsilon}{2}\} \rightarrow 1$$

となる⁶ . なぜならば , $\mathbb{P}\{|X_n - c| \leq \frac{\epsilon}{2}\} \rightarrow 1$ と $\mathbb{P}\{|Y_n - d| \leq \frac{\epsilon}{2}\} \rightarrow 1$ から補題 4.2 を用いればわかる .

(ii) の証明 . $cd \neq 0$ として議論を進める . $c = 0$ または $d = 0$ ならば , 以下の議論は省略できる . $X_n Y_n - cd = (X_n - c)(Y_n - d) + d(X_n - c) + c(Y_n - d)$ に注意する . どんな正の数 $\epsilon > 0$ に対しても

$$\begin{aligned} \mathbb{P}\{|X_n Y_n - cd| > \epsilon\} &\leq \mathbb{P}\{|(X_n - c)(Y_n - d)| > \frac{\epsilon}{3}\} + \mathbb{P}\{|(X_n - c)| > \frac{\epsilon}{3|d|}\} \\ &\quad + \mathbb{P}\{|(Y_n - d)| > \frac{\epsilon}{3|c|}\} \end{aligned}$$

⁶(4.12) の対偶をとれば ,

$$|(X_n + Y_n) - (c + d)| \geq \epsilon \quad \text{ならば} \quad |X_n - c| \geq \frac{\epsilon}{2} \quad \text{また} \quad |Y_n - d| \geq \frac{\epsilon}{2}$$

より

$$\begin{aligned} \mathbb{P}\{|(X_n + Y_n) - (c + d)| \geq \epsilon\} &\leq \mathbb{P}\{|X_n - c| \geq \frac{\epsilon}{2}\} \cup \{|Y_n - d| \geq \frac{\epsilon}{2}\} \\ &\leq \mathbb{P}\{|X_n - c| \geq \frac{\epsilon}{2}\} + \mathbb{P}\{|Y_n - d| \geq \frac{\epsilon}{2}\} \end{aligned}$$

としても示せる .

となる⁷．どんな正の数 $\delta > 0$ に対しても

$$\begin{aligned} & \mathbb{P}\{|(X_n - c)(Y_n - d)| > \frac{\epsilon}{3}\} \\ &= \mathbb{P}\{|(X_n - c)(Y_n - d)| > \frac{\epsilon}{3} \text{ かつ } |Y_n - d| > \delta\} \\ & \quad + \mathbb{P}\{|(X_n - c)(Y_n - d)| > \frac{\epsilon}{3} \text{ かつ } |Y_n - d| \leq \delta\} \\ &\leq \mathbb{P}\{|Y_n - d| > \delta\} + \mathbb{P}\{|(X_n - c)(Y_n - d)| > \frac{\epsilon}{3} \text{ かつ } |Y_n - d| \leq \delta\} \\ &\leq \mathbb{P}\{|Y_n - d| > \delta\} + \mathbb{P}\{|X_n - c| > \frac{\epsilon}{3\delta}\} \rightarrow 0, \quad (n \rightarrow \infty) \end{aligned}$$

となることからわかる．

(iii) の証明． $1/Y_n \xrightarrow{P} 1/d$ を示せば (ii) よりわかる．十分小さな正の数 $\delta > 0$ に対して， $|Y_n - d| \leq \delta$ ならば， $|Y_n| > (1/2)|d|$ より

$$\mathbb{P}\{|Y_n| \leq \frac{1}{2}|d|\} \leq \mathbb{P}\{|Y_n - d| > \delta\} \rightarrow 0, \quad (n \rightarrow \infty)$$

となる⁸．また， $|Y_n| > (1/2)|d|$ のとき，

$$\left| \frac{1}{Y_n} - \frac{1}{d} \right| = \frac{|Y_n - d|}{|Y_n||d|} < \frac{2}{|d|^2}|Y_n - d|$$

⁷ $|X_n - c)(Y_n - d)| \leq \epsilon/3$ かつ $|d(X_n - c)| \leq \epsilon/3$ かつ $|c(Y_n - d)| \leq \epsilon/3$ ならば

$$|(X_n - c)(Y_n - d) + d(X_n - c) + c(Y_n - d)| \leq \epsilon$$

である．これの対偶を取れば，

$$|(X_n - c)(Y_n - d) + d(X_n - c) + c(Y_n - d)| > \epsilon$$

ならば， $|X_n - c)(Y_n - d)| > \epsilon/3$ または $|d(X_n - c)| > \epsilon/3$ または $|c(Y_n - d)| > \epsilon/3$ である．これより

$$\begin{aligned} & \mathbb{P}(|(X_n - c)(Y_n - d) + d(X_n - c) + c(Y_n - d)| > \epsilon) \\ & \leq \mathbb{P}(\{|(X_n - c)(Y_n - d)| > \epsilon/3\} \cup \{|d(X_n - c)| > \epsilon/3\} \cup \{|c(Y_n - d)| > \epsilon/3\}) \\ & \leq \mathbb{P}(|(X_n - c)(Y_n - d)| > \epsilon/3) + \mathbb{P}(|d(X_n - c)| > \epsilon/3) + \mathbb{P}(|c(Y_n - d)| > \epsilon/3) \end{aligned}$$

となる．

⁸ $|Y_n - d| \leq \delta$ ならば， $|Y_n| > (1/2)|d|$ の対偶を取れば， $|Y_n| \leq (1/2)|d|$ ならば， $|Y_n - d| > \delta$ である．したがって， $\mathbb{P}(|Y_n| \leq (1/2)|d|) \leq \mathbb{P}(|Y_n - d| > \delta)$ となる．

が成立する．これらを用いれば，

$$\begin{aligned}
 \mathbb{P} \left\{ \left| \frac{1}{Y_n} - \frac{1}{d} \right| > \epsilon \right\} &= \mathbb{P} \left\{ \left| \frac{1}{Y_n} - \frac{1}{d} \right| > \epsilon, |Y_n| > \frac{1}{2}|d| \right\} \\
 &\quad + \mathbb{P} \left\{ \left| \frac{1}{Y_n} - \frac{1}{d} \right| > \epsilon, |Y_n| \leq \frac{1}{2}|d| \right\} \\
 &\leq \mathbb{P} \left\{ |Y_n - d| > \frac{|d|^2}{2}\epsilon \right\} + \mathbb{P} \left\{ |Y_n| \leq \frac{1}{2}|d| \right\} \\
 &\rightarrow 0, \quad (n \rightarrow \infty)
 \end{aligned} \tag{4.13}$$

より示せた． □

命題 4.11 (連続写像定理): g を実数値連続関数とする．このとき, $Y_n \xrightarrow{P} b$ (b は定数) ならば, $n \uparrow \infty$ のとき,

$$g(Y_n) \xrightarrow{P} g(b)$$

が成立する．

証明 どんな正の数 $\epsilon > 0$ に対してもある正の数 $\delta > 0$ が存在して,

$$|Y_n - b| \leq \delta \quad \text{ならば} \quad |g(Y_n) - g(b)| \leq \epsilon$$

を満足するので,

$$\mathbb{P}\{|Y_n - b| \leq \delta\} \leq \mathbb{P}\{|g(Y_n) - g(b)| \leq \epsilon\}$$

より, $n \uparrow \infty$ のとき,

$$\mathbb{P}\{|g(Y_n) - g(b)| > \epsilon\} \leq \mathbb{P}\{|Y_n - b| > \delta\} \rightarrow 0$$

より命題は示せた． □

命題 4.12 (Slutsky の定理) $\{X_n\}_{n=1}^{\infty}, \{A_n\}_{n=1}^{\infty}, \{B_n\}_{n=1}^{\infty}$ を確率変数列とし $X_n \xrightarrow{d} X$, $A_n \xrightarrow{P} a$, $B_n \xrightarrow{P} b$ を満足するとする．ただし, X は確率変数, a と b は定数とする．このとき,

$$A_n + X_n \xrightarrow{d} a + X$$

と

$$B_n X_n \xrightarrow{d} bX$$

が成立する．特に， $b = 0$ のとき，

$$B_n X_n \xrightarrow{P} 0$$

である．

が成立する⁹．

証明 まず，

$$A_n + X_n \xrightarrow{d} a + X$$

を示す．どんな正の数 $\epsilon > 0$ に対しても

$$\begin{aligned} F_{(A_n+X_n)}(x) &= \mathbb{P}\{A_n + X_n \leq x\} \\ &= \mathbb{P}\{A_n + X_n \leq x, A_n \geq a - \epsilon\} \\ &\quad + \mathbb{P}\{A_n + X_n \leq x, A_n < a - \epsilon\} \end{aligned} \quad (4.14)$$

となることに注意する．十分小さな ϵ をとり， $x \pm \epsilon$ が $F_{(a+X)}(\cdot) = \mathbb{P}\{a+X \leq \cdot\}$ の連続点になるようにとる¹⁰．(4.14) から

$$F_{(A_n+X_n)}(x) \leq \mathbb{P}\{X_n \leq x - a + \epsilon\} + \mathbb{P}\{|A_n - a| > \epsilon\}$$

となる．また， $F_{X+a}(\cdot)$ の連続点 x に対して，

$$\begin{aligned} F_{(X_n+a)}(x) &= \mathbb{P}\{X_n + a \leq x\} = F_{X_n}(x - a) \\ &\rightarrow F_X(x - a) = P(X \leq x - a) = F_{X+a}(x) \end{aligned}$$

となることより， $X_n + a \xrightarrow{d} X + a$ が成り立つ．したがって，

$$\begin{aligned} \limsup_{n \rightarrow \infty} F_{(X_n+A_n)}(x) &\leq \lim_{n \rightarrow \infty} \{\mathbb{P}\{X_n \leq x - a + \epsilon\} + \mathbb{P}\{|A_n - a| > \epsilon\}\} \\ &= \lim_{n \rightarrow \infty} F_{(X_n+a)}(x + \epsilon) + \lim_{n \rightarrow \infty} \mathbb{P}\{|A_n - a| > \epsilon\} \\ &= F_{(X+a)}(x + \epsilon) \end{aligned} \quad (4.15)$$

⁹ $b = 0$ のときは， $A_n + B_n X_n \xrightarrow{P} a$ となるが，注意 4.5 からわかるように， $A_n + B_n X_n \xrightarrow{d} a$ となるのがわかる．

¹⁰不連続点が高々可算個しかないことからこのような点がとれることが保障される．

となる．また，

$$\begin{aligned}
 1 - F_{(X_n + A_n)}(x) &= \mathbb{P}\{X_n + A_n > x\} \\
 &= \mathbb{P}\{X_n + A_n > x, -A_n \geq -a - \epsilon\} \\
 &\quad + \mathbb{P}\{X_n + A_n > x, -A_n < -a - \epsilon\} \\
 &\leq \mathbb{P}\{X_n > x - a - \epsilon\} + \mathbb{P}\{|A_n - a| > \epsilon\}
 \end{aligned} \tag{4.16}$$

から

$$\liminf_{n \rightarrow \infty} F_{(X_n + A_n)}(x) \geq \lim_{n \rightarrow \infty} \{F_{(X_n + a)}(x - \epsilon) - \mathbb{P}\{|A_n - a| \geq \epsilon\}\} = F_{(X + a)}(x - \epsilon) \tag{4.17}$$

となる．(4.15) と (4.17) から

$$F_{(X + a)}(x - \epsilon) \leq \liminf_{n \rightarrow \infty} F_{(X_n + A_n)}(x) \leq \limsup_{n \rightarrow \infty} F_{(X_n + A_n)}(x) \leq F_{(X + a)}(x + \epsilon)$$

を得る．したがって，

$$\lim_{n \rightarrow \infty} F_{(X_n + A_n)}(x) = F_{(X + a)}(x)$$

が成り立つ．

つぎに， $b = 0$ として， $B_n X_n \xrightarrow{P} 0$ を示す．任意の固定した正の数 k と ϵ に対して，

$$\begin{aligned}
 \mathbb{P}(|B_n X_n| > \epsilon) &= \mathbb{P}(|B_n X_n| > \epsilon, |B_n| \leq \epsilon/k) + \mathbb{P}(|B_n X_n| > \epsilon, |B_n| > \epsilon/k) \\
 &\leq \mathbb{P}(|X_n| > k) + \mathbb{P}(|B_n| > \epsilon/k)
 \end{aligned}$$

より

$$\limsup_{n \rightarrow \infty} \mathbb{P}(|B_n X_n| > \epsilon) \leq \mathbb{P}(|X| > k)$$

となる． k を十分大きく選べば，

$$\mathbb{P}(|B_n X_n| > \epsilon) \rightarrow 0 \quad (n \rightarrow \infty)$$

となり， $B_n X_n \xrightarrow{P} 0$ がわかる．

つぎに， $b \neq 0$ のとき，

$$B_n X_n - b X_n = (B_n - b) X_n$$

となるので, $B_n - b \xrightarrow{P} 0$ と $X_n \xrightarrow{d} X$ より, $b = 0$ のときの結果を利用すれば,

$$B_n X_n - b X_n = (B_n - b) X_n \xrightarrow{P} 0$$

となる. これより

$$B_n X_n = b X_n + (B_n - b) X_n \xrightarrow{d} b X$$

がわかる. □

命題 4.13 $X_n \xrightarrow{P} X$ ならば, $X_n \xrightarrow{d} X$ が成立する.

証明: $A_n = X_n - X$ とおく. 条件より, $A_n \xrightarrow{P} 0$ となり, $X_n = X + A_n$ に対して, Slutsky の定理を用いれば, この命題は示される. □

注意 4.5 上の命題の逆は一般には成立しないが, c をある定数とすると,

$$X_n \xrightarrow{d} c \Leftrightarrow X_n \xrightarrow{P} c$$

である. 実際, X_n と X の分布関数を H_n と H かけば, すべての $\epsilon > 0$ に対して,

$$\lim_{n \rightarrow \infty} H_n(c - \epsilon) = H(c - \epsilon) = 0$$

と

$$\lim_{n \rightarrow \infty} H_n(c + \epsilon) = H(c + \epsilon) = 1$$

から

$$\begin{aligned} \mathbb{P}(|X_n - c| > \epsilon) &= \mathbb{P}(X_n > c + \epsilon) + \mathbb{P}(X_n < c - \epsilon) \\ &\leq 1 - H_n(c + \epsilon) + H_n(c - \epsilon) \rightarrow 0 \end{aligned}$$

となる.

命題 4.14 $X_n \xrightarrow{d} X$ となるための必要十分条件は, すべての有界連続な関数 f に対し,

$$\mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)]$$

が成立することである.

証明 $X_n \xrightarrow{d} X$ が成立すると仮定する．与えられた任意の有界連続関数 f に対して，

$$M = \sup_x |f(x)|$$

とおけば， $M < \infty$ となる．このとき，与えられた正の数 $\epsilon > 0$ に対して， $\pm K$ が F_X の連続点になるように K を選び，

$$\mathbb{P}(|X| > K) < \frac{\epsilon}{M}$$

となるようにする．これは， F_X の不連続点は高々可算個であり， $\lim_{x \rightarrow \infty} \mathbb{P}(|X| > x) = 0$ より取れることがわかる．分布収束することにより， n を十分大きく取れば，

$$\mathbb{P}(|X_n| > K) < \frac{2\epsilon}{M}$$

とできる．これらの K と ϵ に対して，階段関数

$$g = \sum_{i=1}^k a_i \mathbb{1}_{(x_{i-1}, x_i]}, \quad -K = x_0 < \cdots < x_k = K,$$

をうまく取り，

$$\sup_{x \in [-K, K]} |f(x) - g(x)| < \epsilon$$

となるようにする．ただし，各 x_i は F_X の連続点になるように取る．このとき，十分大きな n に対して，

$$\begin{aligned} |\mathbb{E}[f(X_n)] - \mathbb{E}[f(X)]| &\leq |\mathbb{E}[f(X_n)\mathbb{1}\{|X_n| \leq K\}] - \mathbb{E}[f(X)\mathbb{1}\{|X| \leq K\}]| \\ &\quad + |\mathbb{E}[f(X_n)\mathbb{1}\{|X_n| > K\}]| + |\mathbb{E}[f(X)\mathbb{1}\{|X| > K\}]| \\ &\leq |\mathbb{E}[f(X_n)\mathbb{1}\{|X_n| \leq K\}] - \mathbb{E}[f(X)\mathbb{1}\{|X| \leq K\}]| + 3\epsilon \\ &\leq 3\epsilon + |\mathbb{E}[f(X_n)\mathbb{1}\{|X_n| \leq K\}] - \mathbb{E}[g(X_n)]| \\ &\quad + |\mathbb{E}[f(X)\mathbb{1}\{|X| \leq K\}] - \mathbb{E}[g(X)]| \\ &\quad + |\mathbb{E}[g(X_n)] - \mathbb{E}[g(X)]| \\ &\leq 5\epsilon + |\mathbb{E}[g(X_n)] - \mathbb{E}[g(X)]| \end{aligned}$$

となる．しかし， $X_n \xrightarrow{d} X$ であり，各 x_i は F_X の連続点なので，

$$\begin{aligned}\mathbb{E}[g(X_n)] &= \sum_{i=1}^k a_i [F_{X_n}(x_i) - F_{X_n}(x_{i-1})] \\ &\rightarrow \sum_{i=1}^k a_i [F_X(x_i) - F_X(x_{i-1})] = \mathbb{E}[g(X)]\end{aligned}$$

となる．したがって，

$$\mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)]$$

がわかる．

つぎに，すべての有界連続な関数 f に対し，

$$\mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)]$$

が成立すると仮定する． F_X の連続点 t と正の数 ϵ に対して，有界連続関数 f を

$$\mathbb{1}_{(-\infty, t]}(x) \leq f(x) \leq \mathbb{1}_{(-\infty, t+\epsilon]}(x)$$

となるように取る．たとえば， $x \leq t$ のとき， $f(x) = 1$ ， $x \geq t + \epsilon$ のとき， $f(x) = 0$ とし， x が t と $t + \epsilon$ の間のとき， f は線形関数とすればよい．このとき，

$$F_{X_n}(t) = \mathbb{E}[\mathbb{1}_{(-\infty, t]}(X_n)] \leq \mathbb{E}[f(X_n)]$$

と

$$\mathbb{E}[f(X)] \leq \mathbb{E}[\mathbb{1}_{(-\infty, t+\epsilon]}(X)] = F_X(t + \epsilon)$$

に注意をして， $n \rightarrow \infty$ とすれば，

$$\limsup_{n \rightarrow \infty} F_{X_n}(t) \leq \lim_{n \rightarrow \infty} \mathbb{E}[f(X_n)] = \mathbb{E}[f(X)] \leq \mathbb{E}[\mathbb{1}_{(-\infty, t+\epsilon]}(X)] = F_X(t + \epsilon)$$

となる． $\epsilon \downarrow 0$ とすれば， F_X の右連続性より

$$\limsup_{n \rightarrow \infty} F_{X_n}(t) \leq F_X(t)$$

となる．

つぎに, f を有界連続関数で

$$\mathbb{1}_{(-\infty, t-\epsilon]}(x) \leq f(x) \leq \mathbb{1}_{(-\infty, t]}(x)$$

となるように取る. このとき,

$$\begin{aligned} F_X(t-\epsilon) &= \mathbb{E}[\mathbb{1}_{(-\infty, t-\epsilon]}(X)] \leq \mathbb{E}[f(X)] = \lim_{n \rightarrow \infty} \mathbb{E}[f(X_n)] \\ &\leq \liminf_{n \rightarrow \infty} \mathbb{E}[\mathbb{1}_{(-\infty, t]}(X_n)] = \liminf_{n \rightarrow \infty} F_{X_n}(t) \end{aligned}$$

となる. t は F_X の連続点なので,

$$F_X(t) = F_X(t-) \leq \liminf_{n \rightarrow \infty} F_{X_n}(t)$$

となる. したがって,

$$F_X(t) \leq \liminf_{n \rightarrow \infty} F_{X_n}(t) \leq \limsup_{n \rightarrow \infty} F_{X_n}(t) \leq F_X(t)$$

より, $F_{X_n}(t) \rightarrow F_X(t)$ がわかる. □

命題 4.15 $\{X_n\}_{n=1}^{\infty}$, X は確率変数とし, $g(x)$ を実数値連続関数とする. このとき, つぎが成立する.

$$X_n \xrightarrow{d} X \quad \text{ならば,} \quad g(X_n) \xrightarrow{d} g(X)$$

.

証明 命題 4.14 から任意の有界連続関数 f に対して, $n \rightarrow \infty$ のとき,

$$\mathbb{E}[f(g(X_n))] \rightarrow \mathbb{E}[f(g(X))]$$

を示せばよい. $f \circ g$ も有界連続関数であることと仮定より上の式は明らか. □

系 4.3 $\{X_n\}_{n=1}^{\infty}$, $\{Y_n\}_{n=1}^{\infty}$ は確率変数とし, $X_n \xrightarrow{d} c$ かつ $Y_n \xrightarrow{d} d$ とする. ただし, c, d は定数である. このとき, つぎが成立する.

- (1) $X_n + Y_n \xrightarrow{d} c + d$.
- (2) $X_n Y_n \xrightarrow{d} cd$.

証明 Slutsky の定理より明らか. □

命題 4.16 (デルタ法) $\{X_n\}_{n=1}^{\infty}$, Z は確率変数, θ を定数とする. $n \rightarrow \infty$ のとき,

$$\sqrt{n}(X_n - \theta) \xrightarrow{d} Z$$

が成立すると仮定する. 実数値関数 $g(x)$ は $x = \theta$ で微分可能で微係数 $\dot{g}(\theta)$ を持ち, $\dot{g}(\theta) \neq 0$ ならば,

$$\sqrt{n}(g(X_n) - g(\theta)) \xrightarrow{d} \dot{g}(\theta)Z$$

が成立する.

証明 まず, 仮定と命題 4.12 から

$$X_n - \theta = \frac{1}{\sqrt{n}}\sqrt{n}(X_n - \theta) \xrightarrow{d} 0$$

となる. さらに, 注意 4.5 から $X_n - \theta \xrightarrow{P} 0$ となる. ここで $g(X_n)$ を $X_n = \theta$ のまわりでテーラー展開する:

$$\sqrt{n}(g(X_n) - g(\theta)) = \dot{g}(\theta)\sqrt{n}(X_n - \theta) + \sqrt{n}\text{Rem} \quad (4.18)$$

ここで

$$\lim_{X_n \rightarrow \theta} \frac{\text{Rem}}{|X_n - \theta|} = 0$$

である. これと仮定から

$$\frac{\text{Rem}}{|X_n - \theta|} \xrightarrow{P} 0$$

となる. また, $\sqrt{n}(X_n - \theta)$ は分布収束することと上のことに注意して, 再度命題 4.14 を用いると

$$\sqrt{n}\text{Rem} = \sqrt{n}(X_n - \theta) \frac{\text{Rem}}{|X_n - \theta|} \xrightarrow{d} 0$$

となり, 注意 4.5 から $\sqrt{n}\text{Rem} \xrightarrow{P} 0$ を得る. (4.18) に命題 4.14 を適用すれば命題は証明される. \square

例 4.7 $\sqrt{n}(X_n - \mu) \xrightarrow{d} X$ とし, X は正規分布 $N(0, \sigma^2)$ に従うとする. ここで, $\mu \neq 0, \sigma^2 > 0$ を仮定する. このとき,

$$\sqrt{n}\left(\frac{1}{X_n} - \frac{1}{\mu}\right) \xrightarrow{d} -\frac{1}{\mu^2}X$$

となる．さらに，正規分布の性質から $-(1/\mu^2)X$ は正規分布 $N(0, \sigma^2/\mu^4)$ に従うことがわかる．

5 大数の法則と中心極限定理

定理 4.1 (大数の (弱) 法則) X_1, X_2, \dots は独立同一分布に従う確率変数列で $\mathbb{E}[|X_1|] < \infty$ とする． $n \rightarrow \infty$ のとき，

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{P} \mu \quad (4.19)$$

が成立する．ただし， $\mathbb{E}[X_1] = \mu$ である．

証明 証明は略． □

注意 4.6 X_1, X_2, \dots は独立同一分布に従う確率変数列で $\mathbb{E}[X_1^2] < \infty$ とする．いま， $\mathbb{E}[X_1] = \mu$ ， $\text{VAR}[X_1] = \sigma^2 < \infty$ と書くことにする．このとき，任意の正数 ϵ に対して， $n \rightarrow \infty$ のとき，

$$\mathbb{P}(|\bar{X}_n - \mu| > \epsilon) \leq \frac{1}{\epsilon^2} \mathbb{E}[(\bar{X}_n - \mu)^2] = \frac{1}{\epsilon^2} \text{VAR}[\bar{X}_n] = \frac{1}{\epsilon^2} \frac{1}{n} \text{VAR}[X_1] \rightarrow 0$$

を得る．ただし，上の不等号はチェビシエフの不等式を利用した． $\mathbb{E}[X_1^2] < \infty$ を仮定すれば，チェビシエフの不等式より (4.19) はわかるが，大数の弱法則は $\mathbb{E}[|X_1|] < \infty$ ならば，(4.19) が成立することを主張している．

例 4.8 X_1, X_2, \dots は独立同一分布に従う確率変数列で $\mathbb{E}[X_1] = \mu$ ， $\text{VAR}[X_1] = \sigma^2 < \infty$ とする．このとき，

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \xrightarrow{P} \sigma^2$$

を示そう．まず，

$$\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = \frac{n}{n-1} \left(\frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}_n^2 \right)$$

に注意する． $\mathbb{E}[|X_1|] < \infty$ から大数の弱法則を用いれば， $\bar{X}_n \xrightarrow{P} \mu$ になる．さらに，命題 4.11 から $\bar{X}_n^2 \xrightarrow{P} \mu^2$ を得る．また， $\mathbb{E}[X_1^2] = \text{VAR}[X_1] + \{\mathbb{E}[X_1]\}^2 < \infty$ から大数の弱法則を用いれば，

$$\frac{1}{n} \sum_{i=1}^n X_i^2 \xrightarrow{P} \mathbb{E}[X_1^2]$$

となる．最後に，系 4.3 から $n \rightarrow \infty$ のとき，

$$\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \xrightarrow{P} \mathbb{E}[X_1^2] - \mu^2 = \text{VAR}[X_1]$$

を得る．

定理 4.2 (中心極限定理) X_1, X_2, \dots は独立同一分布に従う確率変数列で $\mathbb{E}[X_1] = \mu$, $\text{VAR}[X_1] = \sigma^2 < \infty$ とし，

$$Z_n = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n (X_i - \mu) = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$$

とする． $n \rightarrow \infty$ のとき，

$$Z_n \xrightarrow{d} N(0, 1)$$

が成立する．すなわち，任意の実数 x に対し，

$$\lim_{n \rightarrow \infty} \mathbb{P}(Z_n \leq x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

が成立する．

証明 証明は略． □

例 4.9 X_1, X_2, \dots, X_n は独立同一に母数 p , $0 < p < 1$ のベルヌーイ分布に従うとする．すると $T_n = \sum_{i=1}^n X_i$ は母数 n と p の二項分布に従う．したがって，非負の整数 t_1, t_2 ($t_1 < t_2$) に対して，

$$\mathbb{P}(t_1 \leq T_n \leq t_2) = \sum_{x=t_1}^{t_2} f_{T_n}(x)$$

となる . ただし ,

$$f_{T_n}(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0, 1, 2, \dots, n$$

である . この確率の近似値を中心極限定理を利用して求めよう . $\mathbb{E}[X_1] = p$, $\text{VAR}[X_1] = p(1-p)$ と中心極限定理から

$$\mathbb{P}\left(\frac{\sqrt{n}((1/n)T_n - p)}{\sqrt{p(1-p)}} \leq x\right) = \mathbb{P}\left(\frac{T_n - np}{\sqrt{np(1-p)}} \leq x\right) \rightarrow \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

となる . これより

$$\begin{aligned} \mathbb{P}(t_1 \leq T_n \leq t_2) &= \mathbb{P}(t_1 - (1/2) < T_n \leq t_2 + (1/2)) \\ &= \mathbb{P}(T_n \leq t_2 + (1/2)) - \mathbb{P}(T_n \leq t_1 - (1/2)) \\ &= \mathbb{P}\left(\frac{T_n - np}{\sqrt{np(1-p)}} \leq \frac{t_2 - np + (1/2)}{\sqrt{np(1-p)}}\right) \\ &\quad - \mathbb{P}\left(\frac{T_n - np}{\sqrt{np(1-p)}} \leq \frac{t_1 - np - (1/2)}{\sqrt{np(1-p)}}\right) \\ &\approx \Phi\left(\frac{t_2 - np + (1/2)}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{t_1 - np - (1/2)}{\sqrt{np(1-p)}}\right) \end{aligned}$$

となる . ただし ,

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$$

とした .

いま , $n = 25$ と $p = 0.2$ と $t_1 = 3$, $t_2 = 5$ とし , $\mathbb{P}(3 \leq T_{25} \leq 5)$ の近似値を求めよう :

$$\begin{aligned} \mathbb{P}(3 \leq T_{25} \leq 5) &= \mathbb{P}(3.0 - 0.5 < T_{25} \leq 5.0 + 0.5) \\ &\approx \Phi\left(\frac{5.0 - 5.0 + 0.5}{\sqrt{25 \cdot 0.2 \cdot 0.8}}\right) - \Phi\left(\frac{3.0 - 5.0 - 0.5}{\sqrt{25 \cdot 0.2 \cdot 0.8}}\right) \\ &= \Phi(0.25) - \Phi(-1.25) \doteq 0.599 - 0.106 \\ &= 0.493 \end{aligned}$$

となる . なお , 厳密は確率は 0.5184642 である .

```
> pbinom(5,25,0.2)-pbinom(2,25,0.2)
[1] 0.5184642
> dbinom(3,25,0.2)+dbinom(4,25,0.2)+dbinom(5,25,0.2)
[1] 0.5184642
```

のように R で計算すればよい。

6 演習問題

問題 4.1 母集団分布がそれぞれつぎの場合について標本の大きさが n のランダム標本の同時確率密度関数または同時確率関数を書け。

- (1) 母数 p ($0 < p < 1$) のベルヌーイ分布
- (2) 母数 λ ($\lambda > 0$) のポアソン分布
- (3) 区間 (a, b) 上の一様分布．ただし, $a < b$ である．
- (4) 平均 μ , 分散 σ^2 ($0 < \sigma < \infty$) の正規分布．
- (5) 母数 λ ($\lambda > 0$) の指数分布

問題 4.2 平均 μ , 分散 σ^2 ($0 < \sigma < \infty$) の母集団分布からの標本の大きさが n のランダム標本を X_1, X_2, \dots, X_n ($n \geq 4$) とする．つぎの統計量の期待値と分散を求めよ．

- (1) $T_1 = X_1$
- (2) $T_2 = \frac{X_1 + X_2}{2}$
- (3) $T_3 = \sum_{i=1}^n X_i$
- (4) $T_4 = \frac{1}{n} T_3$
- (5) $T_5 = 12$

問題 4.3 平均 μ , 分散 σ^2 ($0 < \sigma < \infty$) の母集団分布からの標本の大きさが n のランダム標本を X_1, X_2, \dots, X_n とする．統計量

$$T(a_1, a_2, \dots, a_n) = a_1 X_1 + a_2 X_2 + \dots + a_n X_n$$

を考える．ただし, $a_i, i = 1, 2, \dots, n$ は既知の定数とする．簡単に $T(a_1, a_2, \dots, a_n)$ を T と書くことにする．

- (1) T の平均 $\mathbb{E}[T]$ と分散 $\text{VAR}[T]$ を求めよ．
- (2) $a_1 + a_2 + \dots + a_n = 1$ のとき

$$\sum_{i=1}^n a_i^2 = \sum_{i=1}^n \left(a_i - \frac{1}{n} \right)^2 + \frac{1}{n}$$

を示せ．

- (3) $a_1 + a_2 + \cdots + a_n = 1$ という条件のもとで T の分散を最小にする a_1, a_2, \dots, a_n を求めよ.

問題 4.4

X_1, X_2 は区間 $(0, 2)$ 上の一様分布からの標本の大きさが 2 のランダム標本とする.
統計量

$$T_1 = X_1 + X_2, \quad T_2 = X_1 - X_2$$

を考える.

- (1) T_1 の期待値 $\mathbb{E}[T_1]$ と分散 $\text{VAR}[T_1]$ を求めよ.
- (2) T_2 の期待値 $\mathbb{E}[T_2]$ と分散 $\text{VAR}[T_2]$ を求めよ.
- (3) T_1, T_2 の同時確率密度関数を求めよ.
- (4) T_1 の確率密度関数を求めよ.

問題 4.5

$$\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}_n^2 = \frac{1}{2n^2} \sum_{i=1}^n \sum_{j=1}^n (X_i - X_j)^2$$

を示せ. ただし, $\bar{X}_n = (1/n) \sum_{i=1}^n X_i$ である.

問題 4.6

以下を示せ.

- (1) Z が標準正規分布に従うとき, $X = Z^2$ は自由度 1 のカイ自乗分布に従う.
- (2) W が母数 r, λ のガンマ分布に従うとは W が確率密度関数

$$f_W(w) = \begin{cases} \frac{\lambda^r}{\Gamma(r)} (\lambda w)^{r-1} e^{-\lambda w} & w > 0, \\ 0 & (\text{その他}) \end{cases}$$

を持つときをいう. ただし, $r > 0, \lambda > 0$ である. W の積率母関数は

$$M_W(t) = \left(\frac{\lambda}{\lambda - t} \right)^r, \quad t < \lambda$$

となることを示せ.

- (3) X が自由度 p のカイ自乗分布に従うとき, X の積率母関数は

$$M_X(t) = \left(\frac{1}{1 - 2t} \right)^{p/2}, \quad t < \frac{1}{2}$$

となることを示せ.

- (4) X_1 と X_2 は互いに独立とし, 各 $X_i, i = 1, 2$ は自由度 p_i カイ自乗分布に従うとき, $X_1 + X_2$ は自由度 $p_1 + p_2$ のカイ自乗分布に従うことを示せ.

問題 4.7

X と Y は独立にそれぞれ正規分布 $N(\mu, \sigma_X^2)$ と $N(\mu, \sigma_Y^2)$ に従うとき,

$$\text{COV} \left[\frac{\sigma_Y^2}{\sigma_X^2 + \sigma_Y^2} X + \frac{\sigma_X^2}{\sigma_X^2 + \sigma_Y^2} Y, X - Y \right] = 0$$

を示せ.

問題 4.8 X_1, X_2, \dots, X_n は正規分布 $N(\mu, 1)$ からの標本の大きさ n のランダム標本とする。ただし, $n \geq 2$ とする。統計量

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2, \quad \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

を考える。

- (1) $n \geq 3$ のとき, 等式

$$(n-1)S_n^2 = (n-2)S_{n-1}^2 + \left(\frac{n-1}{n}\right)(X_n - \bar{X}_{n-1})^2$$

を示せ。 $n = 2$ のときは

$$S_2^2 = \frac{1}{2}(X_2 - X_1)^2$$

を示せ。

- (2) $X_1 - X_2$ の分布を求めよ。
 (3) $X_1 + X_2$ と $X_1 - X_2$ は独立であることを示せ。
 (4) S_2^2 は自由度 1 のカイ自乗分布に従うことを示せ。
 (5) $n = k$ のとき, $(k-1)S_k^2$ は自由度 $k-1$ のカイ自乗分布に従い, S_k^2 と \bar{X}_k は独立であると仮定したとき, 以下を示せ。ただし, $k \geq 2$ である。

(5a) $X_{k+1} - \bar{X}_k$ の期待値 $\mathbb{E}[X_{k+1} - \bar{X}_k]$ と分散 $\text{VAR}[X_{k+1} - \bar{X}_k]$ を求めよ。

(5b) $\sqrt{\frac{k}{k+1}}(X_{k+1} - \bar{X}_k)$ は標準正規分布 $N(0, 1)$ に従うことを示せ。

(5c) $\frac{k}{k+1}(X_{k+1} - \bar{X}_k)^2$ は自由度 1 のカイ自乗分布に従うことを示せ。

(5d) $n = k+1$ のとき,

$$kS_{k+1}^2 = (k-1)S_k^2 + \left(\frac{k}{k+1}\right)(X_{k+1} - \bar{X}_k)^2$$

は自由度 k のカイ自乗分布に従うことを示せ。

- (5e) X_{k+1} と \bar{X}_k はそれぞれ独立で正規分布 $N(\mu, 1)$ と $N(\mu, 1/k)$ に従うことに注意して,

$$\text{COV}\left[\frac{k}{k+1}\bar{X}_k + \frac{1}{k+1}X_{k+1}, \bar{X}_k - X_{k+1}\right] = 0$$

を示せ。したがって, $\frac{k}{k+1}\bar{X}_k + \frac{1}{k+1}X_{k+1}$ と $\bar{X}_k - X_{k+1}$ は正規分布に従うので, $\frac{k}{k+1}\bar{X}_k + \frac{1}{k+1}X_{k+1}$ と $\bar{X}_k - X_{k+1}$ は独立となる。さらに,

$$\bar{X}_{k+1} = \frac{k}{k+1}\bar{X}_k + \frac{1}{k+1}X_{k+1}$$

から S_{k+1}^2 と \bar{X}_{k+1} は独立であることがわかる。

問題 4.9 X_1, X_2, X_3, X_4 を確率密度関数

$$f_X(x) = \begin{cases} 2x & 0 < x < 1, \\ 0 & (\text{その他の場合}) \end{cases}$$

からの標本の大きさが 4 のランダム標本とし, $X_{(1)}, X_{(2)}, X_{(3)}, X_{(4)}$, をその順序統計量とする。

- (1) $X_{(4)}$ の確率密度関数を求めよ。
 (2) $X_{(1)}$ の確率密度関数を求めよ。

問題 4.10 X_1, X_2, \dots, X_n を確率密度関数 $f_X(x)$ からの標本の大きさが $n (n \geq 2)$ のランダム標本とし, $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ をその順序統計量とする. また, $F_X(x) = \int_{-\infty}^x f_X(t) dt$ とする. このとき,

$$R = X_{(n)} - X_{(1)}, \quad T = \frac{X_{(n)} + X_{(1)}}{2}$$

としたとき, R と T の同時確率密度関数は

$$f_{R,T}(r, t) = \begin{cases} n(n-1) [F_X(t + \frac{r}{2}) - F_X(t - \frac{r}{2})]^{n-2} f_X(t - \frac{r}{2}) f_X(t + \frac{r}{2}) & r > 0, \\ 0 & (\text{その他の場合}) \end{cases}$$

で与えられることを示せ.

問題 4.11 X_1, X_2, \dots, X_n を $(0, 1)$ 上の一様分布¹¹からの標本の大きさが $n (n \geq 2)$ のランダム標本とし, $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ をその順序統計量とする. さらに,

$$R = X_{(n)} - X_{(1)}, \quad T = \frac{X_{(n)} + X_{(1)}}{2}$$

とする.

- (1) R と T の同時確率密度関数は

$$f_{R,T}(r, t) = \begin{cases} n(n-1)r^{n-2} & 0 < r < 1, \frac{r}{2} < t < 1 - \frac{r}{2} \\ 0 & (\text{その他の場合}) \end{cases}$$

で与えられることを示せ.

- (2) R の周辺確率密度関数は

$$f_R(r) = \begin{cases} n(n-1)r^{n-2}(1-r) & 0 < r < 1, \\ 0 & (\text{その他の場合}) \end{cases}$$

となることを示せ. さらに,

$$\int_{-\infty}^{\infty} f_R(r) dr = 1$$

を示せ.

問題 4.12 X_1, X_2, \dots, X_n を母数 $\lambda (\lambda > 0)$ の指数分布¹²からの標本の大きさが $n (n \geq 2)$ のランダム標本とし, $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ をその順序統計量とする.

¹¹ 確率密度関数は

$$f_X(x) = \begin{cases} 1 & 0 < x < 1, \\ 0 & (\text{その他の場合}) \end{cases}$$

¹² 確率密度関数は

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0, \\ 0 & (\text{その他の場合}) \end{cases}$$

で与えられる.

- (1) $X_{(1)}$ の確率密度関数を求めよ .
- (2) $X_{(n)}$ の確率密度関数を求めよ .
- (3) $X_{(1)}$ と $X_{(n)}$ の同時確率密度関数を求めよ .
- (4) $R = X_{(n)} - X_{(1)}$ と $T = (X_{(1)} + X_{(n)})/2$ の確率密度関数を求めよ .
- (5) R の周辺確率密度関数を求めよ .

問題 4.13 $\{X_n\}_{n=1}^{\infty}$ を確率変数列とし, 各 X_n は確率関数

$$f_{X_n}(x) = \begin{cases} \frac{1}{n} & (x = c + n), \\ 1 - \frac{1}{n} & (x = c), \\ 0 & (\text{その他}) \end{cases}$$

を持つとする . ただし, c は定数とする . このとき, 以下の問いに答えよ .

- (1) $X_n \xrightarrow{P} c$ を示せ .
- (2) 期待値の定義に従い, $\lim_{n \rightarrow \infty} \mathbb{E}[X_n]$ を求め, $\mathbb{E}[X_n] \not\rightarrow c$ であることを確認せよ .

問題 4.14 確率変数列 $\{X_n\}_{n=1}^{\infty}$ において, 各 X_n は母数 n のポアソン分布に従うとする . すなわち,

$$\mathbb{P}(X_n = k) = \frac{n^k}{k!} e^{-n}, \quad k = 0, 1, \dots$$

である .

- (1) $\mathbb{E}[X_1] = 1$ を示せ . ただし, $e = \sum_{k=0}^{\infty} \frac{1}{k!}$ を用いてよい .
- (2) $\mathbb{E}[X_1(X_1 - 1)] = 1$ を示せ .
- (3) $\text{VAR}[X_1] = \mathbb{E}[X_1^2] - \{\mathbb{E}[X_1]\}^2 = 1$ を示せ .
- (4) Z_1 と Z_2 が独立に母数 1 のポアソン分布に従うとき, $Z_1 + Z_2$ は母数 2 のポアソン分布に従うことを示せ .
ヒント: $\mathbb{P}(Z_1 + Z_2 = k) = \sum_{\ell=0}^k \mathbb{P}(Z_1 = \ell, Z_2 = k - \ell) = \sum_{\ell=0}^k \mathbb{P}(Z_1 = \ell) \mathbb{P}(Z_2 = k - \ell)$ となることと二項定理 $2^k = \sum_{\ell=0}^k \binom{k}{\ell}$ を用いる .
- (5) $X_n/n \xrightarrow{P} 1$ を大数の法則を用いて示せ .
- (6) 中心極限定理を用いて,

$$\frac{X_n - n}{\sqrt{n}} \xrightarrow{d} N(0, 1)$$

を示せ . ヒント: Z_1, Z_2, \dots, Z_n が独立に母数 1 のポアソン分布に従うとき, 上の問いの結果から $Z_1 + Z_2 + \dots + Z_n$ が母数 n のポアソン分布に従うことを利用する .

問題 4.15 X_1, X_2, \dots, X_n を平均 μ , 分散 σ^2 の分布からの標本の大きさが n のランダム標本とし,

$$\bar{X}_n = \frac{1}{n} \sum_{\ell=1}^n X_{\ell}, \quad S_n^2 = \frac{1}{n-1} \sum_{\ell=1}^n (X_{\ell} - \bar{X}_n)^2$$

とする . ただし, $\sigma > 0$ とする .

- (1) 各 X_ℓ ($\ell = 1, 2, \dots, n$) の分布が正規分布のとき,

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$$

は $N(0, 1)$ に従うことを示せ.

- (2) 中心極限定理を用いて,

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0, 1)$$

を示せ.

注意: 各 X_ℓ ($\ell = 1, 2, \dots, n$) の分布が必ずしも正規分布でなくともよい.

- (3)

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sqrt{S_n^2}} \xrightarrow{d} N(0, 1)$$

を示せ. ただし, $S_n^2 \xrightarrow{P} \sigma^2$ は用いてよい.

問題 4.16 X_1, X_2, \dots, X_n を平均 μ , 分散 σ^2 の分布からの標本の大きさが n のランダム標本とし,

$$\bar{X}_n = \frac{1}{n} \sum_{\ell=1}^n X_\ell,$$

とする. ただし, $\sigma > 0$ である.

- (1) $\mathbb{E}[\bar{X}_n]$ と $\text{VAR}[\bar{X}_n]$ を求めよ.
 (2) つぎの不等式をみたすために, n をいくつ以上にすればよいかをチェビシェフの不等式を用いて調べよ.

$$\mathbb{P}(|\bar{X}_n - \mu| < \frac{\sigma}{2}) \geq 0.99 \quad (*)$$

ヒント: 確率変数 Y の分散が存在 ($\text{VAR}[Y] = \tau^2$ ($\tau > 0$)) するならば, 任意の正の数 a に対して,

$$\mathbb{P}(|Y - \mathbb{E}[Y]| \geq a\tau) \leq \frac{1}{a^2}, \quad \mathbb{P}(|Y - \mathbb{E}[Y]| < a\tau) \geq 1 - \frac{1}{a^2},$$

第5章 データの縮約

母集団分布から無作為標本は未知の母数に推測するための情報をもっているが、統計量は、その情報を整理要約している。このときに、未知の母数についての情報を失っているかどうかは興味あるところである。統計量は必要な情報を失っているかどうかを定式化する方法を学ぶ。

1 十分統計量

X_1, X_2, \dots, X_n を確率 (密度) 関数 $f(x|\theta)$ を持つ分布からのランダム標本とする。ただし、 θ は母数空間 $\Theta \subset \mathbb{R}^k$ (k は自然数) の元とする。また、 $T = T(X_1, X_2, \dots, X_n)$ を統計量とする。統計量は n 個の確率変数の「情報」をひとつの確率変数に縮約している。縮約をするときに、どのような「情報」が失われているかを知りたい。別のいい方をすれば、推測に必要な「情報」が失われていない保障があるかどうかを知りたい。このようなことをどのように定式化するかをここでは考えていく。

\mathcal{X} を標本空間とし、 T を統計量とすれば、 T は標本空間を分割していると考えることができる。このことを簡単な例でみてみよう。

例 5.1 X_1, X_2, X_3 はベルヌーイ試行 $Bi(1, p), 0 < p < 1$ からのランダム標本とする。したがって、標本空間は

$$\mathcal{X} = \{(0, 0, 0), (0, 0, 1), (0, 1, 0), (1, 0, 0), (0, 1, 1), (1, 0, 1), (1, 1, 0), (1, 1, 1)\}$$

となる。 $S(x_1, x_2, x_3) = x_1 + x_2 + x_3$ とすれば、 $S(\cdot)$ の値域は $0, 1, 2, 3$ と

なり, \mathcal{X} を分割している :

$S(\cdot)$ の値	標本点
0	(0, 0, 0)
1	(0, 0, 1), (0, 1, 0), (1, 0, 0)
2	(0, 1, 1), (1, 0, 1), (1, 1, 0)
3	(1, 1, 1)

たとえば, $S = 1$ とすれば, 「1」が1回出現したが, 何回目に出現したかという「情報」は失われる.

つぎに, $T = T(X_1, X_2, X_3) = X_1X_2 + X_3$ という統計量を考えて, S と T をそれぞれ与えたときの (X_1, X_2, X_3) の条件付き確率分布をしらべてみよう.

標本点	T の値	$f_{X_1, X_2, X_3 T}(\cdot)$	S 値	$f_{X_1, X_2, X_3 S}(\cdot)$
(0, 0, 0)	0	$\frac{1-p}{1+p}$	0	1
(0, 0, 1)	1	$\frac{1-p}{1+2p}$	1	$\frac{1}{3}$
(0, 1, 0)	0	$\frac{p}{1+p}$	1	$\frac{1}{3}$
(1, 0, 0)	0	$\frac{p}{1+p}$	1	$\frac{1}{3}$
(0, 1, 1)	1	$\frac{p}{1+2p}$	2	$\frac{1}{3}$
(1, 0, 1)	1	$\frac{p}{1+2p}$	2	$\frac{1}{3}$
(1, 1, 0)	1	$\frac{p}{1+2p}$	2	$\frac{1}{3}$
(1, 1, 1)	2	1	3	1

表から S を与えたときの (X_1, X_2, X_3) の条件付き分布は p には依存しないが, T を与えたときの (X_1, X_2, X_3) の条件付き分布は p には依存ことがわかる.

1.1 十分統計量の定義

定義 5.1 X_1, X_2, \dots, X_n は確率(密度)関数 $f(x|\theta)$ を持つ分布からの標本の大きさ n のランダム標本とする. ただし, $\theta \in \Theta$ とする. 統計量 $S = S(X_1, X_2, \dots, X_n)$ が θ の十分統計量であるとは, $S = s$ を与えたときの (X_1, X_2, \dots, X_n) の条件付き分布がどんな s の値に対しても θ に依存しないときをいう.

例 5.2 X_1, X_2 は正規分布 $N(\theta, 1)$ からの標本の大きさ 2 のランダム標本とする. $S = S(X_1, X_2) = X_1 + X_2$ は十分統計量であることを示そう. その

ために，変換

$$\begin{cases} S = X_1 + X_2 \\ R = X_1 - X_2 \end{cases}$$

を考える． (S, R) と (X_1, X_2) は一対一対応である．また，正規分布の性質から S と R は独立となるので，

$$f_{R|S}(r|s) = \frac{f_{R,S}(r, s)}{f_S(s)} = \frac{f_R(r)f_S(s)}{f_S(s)} = f_R(r)$$

となる．しかし，

$$f_R(r) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2}} e^{-r^2/4}$$

となるので， S を与えたときの R の条件付き分布は θ に依存しない．さらに， (S, R) と (X_1, X_2) は一対一対応なので， S を与えたときの (X_1, X_2) の条件付き分布は θ に依存しないことがわかる．

統計量が複数あるときも十分性を拡張して定義できる．

定義 5.2 X_1, X_2, \dots, X_n は確率（密度）関数 $f(x|\theta)$ を持つ分布からの標本の大きさ n のランダム標本とする．ただし， $\theta \in \Theta$ とする．統計量 $S_1 = S_1(X_1, X_2, \dots, X_n)$, $S_2 = S_2(X_1, X_2, \dots, X_n)$, \dots , $S_\ell = S_\ell(X_1, X_2, \dots, X_n)$ が十分統計量であるとは， $S_1 = s_1, S_2 = s_2, \dots, S_\ell = s_\ell$ を与えたときの (X_1, X_2, \dots, X_n) の条件付き分布がどんな s_1, s_2, \dots, s_ℓ の値に対しても θ に依存しないときをいう．

例 5.3 X_1, X_2, \dots, X_n をベルヌーイ試行

$$f_\theta(z) = \theta^z(1-\theta)^{1-z} \mathbb{1}_{\{0,1\}}(z), \quad z \in \mathbb{R}, \quad \theta \in (0, 1)$$

からの標本の大きさ n のランダム標本とし， $\mathbf{X} = (X_1, X_2, \dots, X_n)$ とする． $S(\mathbf{X}) = \sum_{i=1}^n X_i$ なる統計量は，直観から θ に対する情報をすべて含んでいると予想される．したがって，十分統計量となることが予想される．これを示すために， $\mathbf{x} = (x_1, x_2, \dots, x_n)$ とし，

$$\mathbb{P}(\mathbf{X} = \mathbf{x} | S = s) = \frac{\mathbb{P}(\mathbf{X} = \mathbf{x}, S = s)}{\mathbb{P}(S = s)}$$

が $\theta \in (0, 1)$ に依存しないことを示せばよい.

$$\mathbb{P}(S = s) = \binom{n}{s} \theta^s (1 - \theta)^{n-s} \mathbb{1}_{\{0, 1, \dots, n\}}(s)$$

と

$$\begin{aligned} \mathbb{P}(X = \mathbf{x}, S = s) &= \prod_{i=1}^n \mathbb{P}(X_i = x_i) \\ &= \prod_{i=1}^n \theta^{x_i} (1 - \theta)^{1-x_i} \mathbb{1}_{\{0, 1\}}(x_i) \\ &= \theta^s (1 - \theta)^{n-s} \prod_{i=1}^n \mathbb{1}_{\{0, 1\}}(x_i) \end{aligned}$$

となることに注意する. $B_s = \{(x_1, x_2, \dots, x_n) : x_i = 0, 1, \sum_{i=1}^n x_i = s\}$ とおけば,

$$\mathbb{P}(X = \mathbf{x} | S = s) = \frac{1}{\binom{n}{s}} \mathbb{1}_{B_s}(\mathbf{x})$$

となり, θ に依存しないので, $S(X)$ は $\theta \in (0, 1)$ に対する十分統計量である.

1.2 分解定理

定義に従って十分統計量を見つけるにはあらかじめ十分統計量と思われるものが事前にわかっているなければならない. つぎの定理を用いると比較的容易に十分統計量を見つけることができる.

命題 5.1 σ -有限な測度 ν によって優越される $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ 上の確率分布族を $\mathcal{P} = \{\mathbb{P}_\theta : \theta \in \Theta\}$ とし, X を \mathbb{P}_θ からのランダム標本とする. このとき, $S(X)$ が $\theta \in \Theta$ に対して十分であるための必要十分条件は $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ 上の非負のボレロ可測関数 h と S の値域上の関数 g_θ (\mathbb{P}_θ に依存) が存在し,

$$\frac{d\mathbb{P}_\theta}{d\nu}(\mathbf{x}) = g_\theta(S(\mathbf{x}))h(\mathbf{x}) \quad (5.1)$$

と書けることである.

証明 まず，証明は離散型分布の場合は比較的簡単であるので，その場合について証明を与える． σ -有限な測度 ν によって優越される $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ 上の確率分布族に対する証明は後で示す．

はじめに S が十分であると仮定する．このとき，

$$\begin{aligned}\mathbb{P}_\theta(\mathbf{X} = \mathbf{x}) &= \sum_t \mathbb{P}_\theta(\mathbf{X} = \mathbf{x}, S = s) \\ &= \mathbb{P}_\theta(\mathbf{X} = \mathbf{x}, S = S(\mathbf{x})) \\ &= \mathbb{P}_\theta(S = S(\mathbf{x}))\mathbb{P}_\theta(\mathbf{X} = \mathbf{x}|S = S(\mathbf{x}))\end{aligned}$$

となる¹． S が十分統計量であることから $\mathbb{P}_\theta(\mathbf{X} = \mathbf{x}|S = S(\mathbf{x}))$ は θ に依存しないので，

$$\begin{aligned}g_\theta(S(\mathbf{x})) &= \mathbb{P}_\theta(S = S(\mathbf{x})) \\ h(\mathbf{x}) &= \mathbb{P}_\theta(\mathbf{X} = \mathbf{x}|S = S(\mathbf{x}))\end{aligned}$$

とおけばよい．

つぎに，(5.1) が成立すると仮定する． $s = S(\mathbf{x})$ とおく．このとき，

$$\begin{aligned}\mathbb{P}_\theta(S = s) &= \sum_{\mathbf{y}: S(\mathbf{y})=s} \mathbb{P}_\theta(\mathbf{X} = \mathbf{y}) \\ &= \sum_{\mathbf{y}: S(\mathbf{y})=s} g_\theta(S(\mathbf{y}))h(\mathbf{y}) \\ &= g_\theta(s) \sum_{\mathbf{y}: S(\mathbf{y})=s} h(\mathbf{y})\end{aligned}$$

となる．従って，

$$\begin{aligned}\mathbb{P}_\theta(\mathbf{X} = \mathbf{x}|S = s) &= \frac{\mathbb{P}_\theta(\mathbf{X} = \mathbf{x}, S = s)}{\mathbb{P}_\theta(S = s)} \\ &= \frac{g_\theta(s)h(\mathbf{x})}{g_\theta(s) \sum_{\mathbf{y}: S(\mathbf{y})=s} h(\mathbf{y})} \\ &= \frac{h(\mathbf{x})}{\sum_{\mathbf{y}: S(\mathbf{y})=s} h(\mathbf{y})}\end{aligned}$$

¹— 一番目の等式は

$$\{\mathbf{X} = \mathbf{x}\} = \{\mathbf{X} = \mathbf{x}\} \cap (\cup_s \{S = s\}) = \cup_s (\{\mathbf{X} = \mathbf{x}\} \cap \{S = s\})$$

からわかる．また，二番目の $S(\mathbf{x}) = t$ を満たさない \mathbf{x} との積事象は空事象なので， $T(\mathbf{x}) = t$ 以外の和の t に関する項の事象は空事象となるので，和はなくなることがわかる．

となり, θ に依存しないことがわかる. □

例 5.4 : X_1, X_2, \dots, X_n を確率密度関数

$$f_X(x|\theta) = \frac{1}{\theta} \mathbb{1}_{(0, \theta)}(x)$$

からのランダム標本とする. ただし, $\theta > 0$ とする. $\mathbf{X} = (X_1, X_2, \dots, X_n)$ の同時確率密度関数は

$$\begin{aligned} f(\mathbf{x}|\theta) &= \frac{1}{\theta^n} \mathbb{1}_{(0, \theta)^n}(\mathbf{x}) \\ &= \frac{1}{\theta^n} \mathbb{1}\{\max_{1 \leq i \leq n} x_i \leq \theta\} \mathbb{1}\{\min_{1 \leq i \leq n} x_i \geq 0\} \\ &= g_\theta(\max(x_1, x_2, \dots, x_n)) h(\mathbf{x}) \end{aligned}$$

となる. ただし, $\mathbf{x} = (x_1, x_2, \dots, x_n)$ である. 従って,

$$X_{(n)} = \max(X_1, X_2, \dots, X_n)$$

は $\theta \in (0, \infty)$ の十分統計量である.

例 5.5 : $\mathbf{X} = (X_1, X_2, \dots, X_n)$ の分布が k 母数指数分布族に属するとする. すなわち, その同時確率関数または確率密度関数が

$$f(\mathbf{x}|\theta) = \exp \left[\sum_{i=1}^k c_i(\theta) T_i(\mathbf{x}) - d(\theta) + S(\mathbf{x}) \right] \mathbb{1}\{\mathbf{x} \in A\}$$

で与えられる. このとき, $h(\mathbf{x}) = \exp[S(\mathbf{x})] \mathbb{1}\{\mathbf{x} \in A\}$ とみれば, 因数分解定理から

$$T = (T_1(\mathbf{X}), T_2(\mathbf{X}), \dots, T_k(\mathbf{X}))$$

は θ の十分統計量となる.

例 5.6 X_1, X_2, \dots, X_n を正規分布 $N(\theta, 1)$ からの大きさ n のランダム標本とする. このとき,

$$\begin{aligned} \prod_{i=1}^n f(x_i|\theta) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(x_i - \theta)\right] = \frac{1}{(2\pi)^{n/2}} \exp\left[-\frac{1}{2} \sum_{i=1}^n (x_i - \theta)\right] \\ &= \frac{1}{(2\pi)^{n/2}} \exp\left[\theta \sum_{i=1}^n x_i - \frac{n}{2}\theta\right] \exp\left[-\frac{1}{2} \sum_{i=1}^n x_i^2\right] \end{aligned}$$

となる．したがって， $S(X_1, X_2, \dots, X_n) = \sum_{i=1}^n X_i$ として，

$$g(s|\theta) = \exp\left[\theta s - \frac{n}{2}\theta\right],$$

$$h(x_1, x_2, \dots, x_n) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2}\sum_{i=1}^n x_i^2\right]$$

とすれば，分解定理から $S(X_1, X_2, \dots, X_n) = \sum_{i=1}^n X_i$ は十分統計量となることがわかる．

例 5.7 X_1, X_2, \dots, X_n を正規分布 $N(\mu, \sigma^2)$ からの大きさ n のランダム標本とする．ただし， $\theta = (\mu, \sigma^2)$ である．

$$\begin{aligned} \prod_{i=1}^n f(x_i|\theta) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma}\right)^2\right] \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2}\sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma}\right)^2\right] \\ &= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n x_i^2 - 2\mu\sum_{i=1}^n x_i + n\mu^2\right)\right] \end{aligned}$$

となる．したがって，

$$S_1(X_1, X_2, \dots, X_n) = \sum_{i=1}^n X_i, \quad S_2(X_1, X_2, \dots, X_n) = \sum_{i=1}^n X_i^2$$

とし，

$$g(s_1, s_2|\theta) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2}(s_2 - 2\mu s_1 + n\mu^2)\right],$$

$$h(x_1, x_2, \dots, x_n) = 1$$

とおけば， $\sum_{i=1}^n X_i$ と $\sum_{i=1}^n X_i^2$ は十分統計量になる．さらに，

$$\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right) \longleftrightarrow \left(\bar{X}_n, \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2\right)$$

は一対一対応なので， $(\bar{X}_n, \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2)$ も十分統計量であることがわかる．十分統計量が表現のしかたは一通りでないことに注意する．

例 5.8 X_1, X_2, \dots, X_n を一様分布

$$f(x|\theta) = \frac{1}{\theta_2 - \theta_1} \mathbb{1}_{(\theta_1, \theta_2)}(x)$$

からの大きさ n のランダム標本とする。ただし, $\theta = (\theta_1, \theta_2), \theta_2 > \theta_1$ である。このとき,

$$\begin{aligned} \prod_{i=1}^n f(x_i|\theta) &= \prod_{i=1}^n \frac{1}{\theta_2 - \theta_1} \mathbb{1}_{(\theta_1, \theta_2)}(x_i) \\ &= \frac{1}{(\theta_2 - \theta_1)^n} \prod_{i=1}^n \mathbb{1}_{(\theta_1, \theta_2)}(x_i) \\ &= \frac{1}{(\theta_2 - \theta_1)^n} \prod_{i=1}^n \mathbb{1}_{(\theta_1, y_n)}(x_i) \mathbb{1}_{(y_1, \theta_2)}(x_n) \end{aligned}$$

となる²。ただし, $y_1 = \min[x_1, x_2, \dots, x_n], y_n = \max[x_1, x_2, \dots, x_n]$ である。したがって, $(\min[X_1, X_2, \dots, X_n], \max[X_1, X_2, \dots, X_n])$ が十分統計量となる。

定理 5.1 最尤推定量は十分統計量を通してのみ標本に依存する。

証明

$$\begin{aligned} S_1 &= S_1(X_1, X_2, \dots, X_n), \\ S_2 &= S_2(X_1, X_2, \dots, X_n), \\ &\vdots \\ S_\ell &= S_\ell(X_1, X_2, \dots, X_n) \end{aligned}$$

を十分統計量とする。尤度関数は

$$L(\theta|x_1, x_2, \dots, x_n) = \prod_{i=1}^n f(x_i|\theta) = g(s_1, s_2, \dots, s_n|\theta)h(x_1, x_2, \dots, x_n)$$

と書ける。ただし, $s_i = S_i(x_1, x_2, \dots, x_n), i = 1, 2, \dots, \ell$ とした。尤度関数 $L(\theta|x_1, x_2, \dots, x_n)$ の θ に関する最大化は $g(s_1, s_2, \dots, s_n|\theta)$ を θ に関する最大化することと同値である。よって, 定理は証明された。□

2

$\theta_1 < x_i < \theta_2 \iff \theta_1 < \min[x_1, x_2, \dots, x_n] < \max[x_1, x_2, \dots, x_n] < \theta_2$
からわかる。

2 演習問題

問題 5.1

X_1, X_2 は正規分布 $N(0, 1)$ からの標本の大きさが 2 のランダム標本とする.

$$\begin{cases} S = X_1 + X_2 \\ R = X_1 - X_2 \end{cases}$$

としたとき, つぎの問いに答えよ.

- (1) S と R の共分散 $\text{COV}(S, R) = \mathbb{E}[(S - \mathbb{E}[S])(R - \mathbb{E}[R])]$ を求めよ.
- (2) (S, R) の同時確率密度関数を求めよ.
- (3) R の確率密度関数を求めよ.
- (4) S と R は独立かどうかを調べよ.

問題 5.2

確率変数 X を確率関数

$$f(x|\theta) = \left(\frac{\theta}{2}\right)^{|x|} (1-\theta)^{1-|x|}, \quad x = -1, 0, 1$$

からの標本の大きさが 1 のランダム標本とする. ただし, $0 < \theta < 1$ とする. ふたつの統計量

$$S = S(X) = |X|, \quad T = T(X) = \begin{cases} 2, & (X = 1) \\ 0, & \text{その他} \end{cases}$$

を考える.

- (1) S の確率分布および S が与えられたときの X の条件付確率分布を求め, S は θ の十分統計量かどうかを調べよ.
- (2) S は θ の不偏推定量かどうかを調べよ.
- (3) T は θ の不偏推定量かどうかを調べよ.
- (4) S と T の MSE (平均 2 乗誤差) を比較せよ. (横軸を θ とし, 縦軸を MSE の値として, S と T の MSE をグラフに描き比較すること)

第6章 点推定法

母集団分布から無作為標本に基づき、母集団分布の未知の母数をひとつの値で推定する方法とその推定精度の評価法について学ぶ。

1 点推定量の性質

1.1 近隣度 (closeness)

X_1, X_2, \dots, X_n を確率 (密度) 関数 $f_X(x|\theta)$ からの標本の大きさが n のランダム標本とする。 $\Theta \subset \mathbb{R}^k$ を θ の母数空間とし、確率 (密度) 関数 $f_X(x|\theta)$ は未知母数 θ のみに依存するものとする。 $\tau: \Theta \rightarrow \mathbb{R}$ とし、 $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ とおく。

定義 6.1 $t(\mathbf{X})$ と $t'(\mathbf{X})$ を $\tau(\theta)$ の推定量とする。 $t'(\mathbf{X})$ が $t(\mathbf{X})$ より集中した $\tau(\theta)$ の推定量であるとは、すべての $\lambda > 0$ に対して

$$\mathbb{P}_\theta[|t'(\mathbf{X}) - \tau(\theta)| \leq \lambda] \geq \mathbb{P}_\theta[|t(\mathbf{X}) - \tau(\theta)| \leq \lambda], \quad \theta \in \Theta$$

を満足することをいう。 $t^* = t^*(\mathbf{X}) - \tau(\theta)$ が最も集中しているとは、他のどんな推定量より集中していることをいう。

注意 6.1 t^* は最も望ましい推定量であるが、ほとんどの場合は存在しない。

定義 6.2 (Pitman's closeness) $t'(\mathbf{X})$ が $t(\mathbf{X})$ より Pitman の意味で $\tau(\theta)$ に近い推定量であるとは

$$\mathbb{P}_\theta[|t'(\mathbf{X}) - \tau(\theta)| < |t(\mathbf{X}) - \tau(\theta)|] > \frac{1}{2}, \quad \theta \in \Theta$$

を満足することである。

$t^*(\mathbf{X})$ が Pitman の意味で $\tau(\theta)$ の最近隣推定量であるとは、どんな推定量 $t(\mathbf{X})$ よりも Pitman の意味で $\tau(\theta)$ に近いものをいう。

注意 6.2 これもなかなか存在しない。

1.2 平均 2 乗誤差

定義 6.3 $t(\mathbf{X})$ を $\tau(\theta)$ の推定量とする．推定量 $t(\mathbf{X})$ の平均 2 乗誤差を

$$\mathbb{E}_\theta[(t(\mathbf{X}) - \tau(\theta))^2]$$

で定義する．これを $MSE(\theta, t)$ と記すことにする．

注意 6.3 \mathbb{E}_θ の θ は考えている標本が分布族の中のどこから来ているかを明示するために用いている．すなわち

$$\begin{aligned} \mathbb{E}_\theta[(t(\mathbf{X}) - \tau(\theta))^2] &= \int \cdots \int [t(x_1, x_2, \dots, x_n) - \tau(\theta)]^2 f(x_1|\theta) f(x_2|\theta) \times \cdots \times f(x_n|\theta) \\ &\quad \times dx_1 dx_2 \cdots dx_n \end{aligned}$$

である．

注意 6.4 $t_1(\mathbf{X})$ と $t_2(\mathbf{X})$ の平均 2 乗誤差はともに θ の関数となる．

例 6.1 $\tau(\theta) = \theta$, $\theta \in \Theta \subset \mathbb{R}$ とする． $\theta_0 \in \Theta$ を固定し,

$$t_{\theta_0}(x_1, x_2, \dots, x_n) = \theta_0$$

なる推定量 $t_{\theta_0}(\mathbf{X})$ を考える． $t_{\theta_0}(\mathbf{X})$ の平均 2 乗誤差は

$$MSE(\theta, t_{\theta_0}) = \mathbb{E}_\theta[(t_{\theta_0}(\mathbf{X}) - \theta)^2] = \mathbb{E}_\theta[(\theta_0 - \theta)^2] = (\theta_0 - \theta)^2$$

となる．すなわち,

$$MSE(\theta_0, t_{\theta_0}) = 0$$

なる．

いま，すべての推定量 $t(\mathbf{X})$ に対して

$$MSE(\theta, t^*) \leq MSE(\theta, t), \quad \forall \theta \in \Theta$$

を満足するような推定量 $t^*(\mathbf{X})$ が存在するならば,

$$MSE(\theta, t^*) = 0$$

とならなければいけない¹．これは常に θ を正しく推定できることある．したがって， θ がわかっていることと同じである．

¹ $t(\mathbf{X})$ として $t_{\theta_0}(\mathbf{X})$ とすればよい．

θ に関して一様に MSE を最小にする推定量を見つけることができない理由はすべての推定量の集合は大きすぎるからである．したがって，対象とする推定量の集合を小さくしよう．

定義 6.4 推定量 $t(\mathbf{X})$ が $\tau(\theta)$ の不偏推定量であるとは，すべての $\theta \in \Theta$ に対して

$$\mathbb{E}_\theta[t(\mathbf{X})] = \tau(\theta)$$

を満足することである．

例 6.2 $t_{\theta_0}(\mathbf{X})$ は θ の不偏推定量ではない．なぜならば， $\theta \neq \theta_0$ のとき

$$\mathbb{E}_\theta[t_{\theta_0}(\mathbf{X})] = \theta_0 \neq \theta$$

からわかる．

定理 6.1 $t(\mathbf{X})$ は $\tau(\theta)$ の推定量とし， $\mathbb{E}_\theta|t(\mathbf{X})|^2 < \infty$ とする．このとき，

$$MSE(\theta, t) = \text{VAR}(t(\mathbf{X})) + \{\mathbb{E}_\theta[t(\mathbf{X})] - \tau(\theta)\}^2$$

が成立する．さらに， $t(\mathbf{X})$ が $\tau(\theta)$ の不偏推定量ならば，

$$MSE(\theta, t) = \text{VAR}(t(\mathbf{X}))$$

が成立する．

証明

$$\begin{aligned} MSE(\theta, t) &= \mathbb{E}_\theta[\{t(\mathbf{X}) - \mathbb{E}_\theta[t(\mathbf{X})] + \mathbb{E}_\theta[t(\mathbf{X})] - \tau(\theta)\}^2] \\ &= \mathbb{E}_\theta[\{t(\mathbf{X}) - \mathbb{E}_\theta[t(\mathbf{X})]\}^2] \\ &\quad + 2(\mathbb{E}_\theta[t(\mathbf{X})] - \tau(\theta))\mathbb{E}_\theta[t(\mathbf{X}) - \mathbb{E}_\theta[t(\mathbf{X})]] \\ &\quad + \{\mathbb{E}_\theta[t(\mathbf{X})] - \tau(\theta)\}^2 \\ &= \text{VAR}(t(\mathbf{X})) + \{\mathbb{E}_\theta[t(\mathbf{X})] - \tau(\theta)\}^2 \end{aligned}$$

□

定義 6.5 $BIAS = \mathbb{E}_\theta[t(\mathbf{X}) - \tau(\theta)]$ を推定量 $t(\mathbf{X})$ のバイアスという．

例 6.3 $X_1, X_2, \dots, X_n (n \geq 2)$ を $N(\mu, \sigma^2)$ からのランダム標本とする .

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i, \quad U^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \quad S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

とおく .

$\mathbb{E}_{\mu, \sigma^2}[\bar{X}_n] = \mu$ より \bar{X}_n は μ の不偏推定量である . また , \bar{X}_n の MSE ($\tau(\mu, \sigma^2) = \mu$) は

$$MSE(\mu, \bar{X}_n) = \mathbb{E}_{\mu, \sigma^2}[(\bar{X}_n - \mu)^2] = \frac{\sigma^2}{n}$$

である .

一方 ,

$$\mathbb{E}_{\mu, \sigma^2}[S^2] = \mathbb{E}_{\mu, \sigma^2}\left[\frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 - (\bar{X}_n - \mu)^2\right] = \sigma^2 - \frac{\sigma^2}{n} = \frac{n-1}{n} \sigma^2$$

となり , σ^2 の最尤推定量は σ^2 の不偏推定量ではない . しかし ,

$$\begin{aligned} \mathbb{E}_{\mu, \sigma^2}[U^2] &= \mathbb{E}_{\mu, \sigma^2}\left[\frac{1}{n-1} \sum_{i=1}^n (X_i - \mu)^2 - \frac{n}{n-1} (\bar{X}_n - \mu)^2\right] \\ &= \frac{n}{n-1} \sigma^2 - \frac{1}{n-1} \sigma^2 = \sigma^2 \end{aligned}$$

となり , U^2 は σ^2 の不偏推定量である .

例 6.4 X_1, X_2, \dots, X_n をベルヌーイ試行² $\text{Bi}(1, p)$ からのランダム標本とする . ただし , $0 < p < 1$ とする . 標本平均 \bar{X}_n は p の不偏推定量となる . $\tau(p) = p$ として , \bar{X}_n の MSE を求めよう .

$$MSE(p, \bar{X}_n) = \mathbb{E}[(\bar{X}_n - p)^2] = \text{VAR}[\bar{X}_n] = \frac{p(1-p)}{n}$$

となる . いま , $Y = \sum_{i=1}^n X_i$ として ,

$$\hat{p}_B = \frac{Y + \alpha}{\alpha + \beta + n}$$

²すなわち ,

$$\mathbb{P}(X_1 = x) = p^x (1-p)^{1-x}, \quad x = 0, 1$$

なる推定量を考えよう。ただし、 α と β は (n に依存する) 定数とする。定理 6.1 を利用して、

$$\begin{aligned} MSE(p, \hat{p}_B) &= \text{VAR}[\hat{p}_B] + \{\mathbb{E}[\hat{p}_B] - p\}^2 \\ &= \text{VAR}\left[\frac{Y + \alpha}{\alpha + \beta + n}\right] + \left[\frac{np + \alpha}{\alpha + \beta + n} - p\right]^2 \\ &= \frac{1}{(\alpha + \beta + n)^2} \text{VAR}[Y] + \left[\frac{np + \alpha}{\alpha + \beta + n} - p\right]^2 \\ &= \frac{np(1-p)}{(\alpha + \beta + n)^2} + \left[\frac{np + \alpha}{\alpha + \beta + n} - p\right]^2 \end{aligned}$$

となる。ここで $\alpha = \beta = \sqrt{n/4}$ とおけば、

$$MSE(p, \hat{p}_B) = \frac{n}{4(\sqrt{n} + n)^2}$$

となる³。

1.3 推定量の一致性

定義 6.6 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ とする。 $\tau(\theta)$ の推定量の列 $\{t_n(\mathbf{X})\}_{n=1}^{\infty}$ が平均 2 乗誤差の意味で $\tau(\theta)$ の一致推定量であるとは、すべての $\theta \in \Theta$ に対して、

$$\lim_{n \rightarrow \infty} \mathbb{E}_{\theta}[(t_n(\mathbf{X}) - \tau(\theta))^2] = 0$$

を満足することである。

注意 6.5 推定量 $t_n(\mathbf{X})$ が平均 2 乗誤差の意味で一致性を持てば、定理 6.1 から

$$\lim_{n \rightarrow \infty} \mathbb{E}_{\theta}[t_n(\mathbf{X})] = \tau(\theta), \quad \lim_{n \rightarrow \infty} \text{VAR}_{\theta}[t_n(\mathbf{X})] = 0$$

³なぜならば、

$$\begin{aligned} MSE(p, \hat{p}_B) &= \frac{np(1-p)}{(\sqrt{n} + n)^2} + \left\{ \frac{np + \sqrt{n/4}}{\sqrt{n} + n} - p \right\}^2 \\ &= \frac{np(1-p)}{(\sqrt{n} + n)^2} + \left\{ \frac{np + \sqrt{n/4} - p\sqrt{n} - np}{\sqrt{n} + n} \right\}^2 \\ &= \frac{np(1-p)}{(\sqrt{n} + n)^2} + \frac{n/4 + p^2n - 2\sqrt{n/4} \cdot p\sqrt{n}}{(\sqrt{n} + n)^2} = \frac{n}{4(\sqrt{n} + n)^2} \end{aligned}$$

からわかる。

となる .

例 6.5 X_1, X_2, \dots, X_n を平均 μ , 分散 σ^2 の分布からのランダム標本とする . ただし , $\sigma < \infty$ とする . $\{\bar{X}_n\}_{n=1}^\infty$ を μ の推定量の列とすれば ,

$$\mathbb{E}_{\mu, \sigma^2}[\{\bar{X}_n - \mu\}^2] = \text{VAR}[\bar{X}_n] = \frac{\sigma^2}{n}$$

よって , \bar{X}_n は平均 2 誤差の意味で一致性を持つ .

定義 6.7 $\tau(\theta)$ の推定量の列 $\{t_n(\mathbf{X})\}_{n=1}^\infty$ が弱一致性を持つとは , どんな正の数 ϵ に対しても

$$\lim_{n \rightarrow \infty} \mathbb{P}_\theta[|t_n(\mathbf{X}) - \tau(\theta)| < \epsilon] = 1, \quad \forall \theta \in \Theta$$

が成立することである . 簡単に , このような推定量 $t_n(\mathbf{X})$ を弱一致推定量という .

注意 6.6 $t_n(\mathbf{X})$ が平均 2 乗誤差の意味で一致推定量であれば , $t_n(\mathbf{X})$ は弱一致推定量である . なぜならば , Markov の不等式から

$$\begin{aligned} \mathbb{P}_\theta[|t_n(\mathbf{X}) - \tau(\theta)| < \epsilon] &= 1 - \mathbb{P}_\theta[|t_n(\mathbf{X}) - \tau(\theta)| \geq \epsilon] \\ &= 1 - \mathbb{P}_\theta[|t_n(\mathbf{X}) - \tau(\theta)|^2 \geq \epsilon^2] \\ &\geq 1 - \frac{1}{\epsilon^2} \mathbb{E}_\theta[|t_n(\mathbf{X}) - \tau(\theta)|^2] \nearrow 1, \quad n \rightarrow \infty \end{aligned}$$

からわかる .

2 最尤法

確率ベクトル $\mathbf{X} = (X_1, X_2, \dots, X_n)$ の統計モデルを $\mathcal{P} = \{P_\theta : \theta \in \Theta \subset \mathbb{R}^k\}$ とする . \mathcal{P} に含まれる P_θ に対応する確率密度関数もしくは確率関数を $p(\mathbf{x}|\theta)$ と記すことにする . ただし , $\mathbf{x} = (x_1, x_2, \dots, x_n)$ である . $\mathbf{X} = \mathbf{x}$ が観測されたときの尤度関数を

$$L_n(\theta|\mathbf{x}) = p(\mathbf{x}|\theta), \quad \theta \in \Theta$$

で定めることにする . $L_n(\cdot|\mathbf{x})$ は標本空間から $\{\theta \mapsto p(\mathbf{x}|\theta) : \mathbf{x} \in S\}$ なる関数族への対応となる . \mathbf{x} があたられたとき , $L_n(\theta|\mathbf{x})$ は θ の関数とみなす . こ

れを簡単に $L_n(\theta)$ と書くことにする． $L_n(\theta)$ は x が与えられたとき，いろいろな θ の「確からしさ」もしくは「尤もらしさ」を表現するものである．

特に， X_1, X_2, \dots, X_n が独立同一に確率密度関数 $f(x|\theta)$ に従うならば，尤度関数は

$$L_n(\theta|\mathbf{x}) = \prod_{i=1}^n f(x_i|\theta)$$

で与えられる．

最尤法とは，与えられたデータを実現させるために「尤もらしい」母数の値を母数の推定値として用いる手法である．すなわち， $X = x$ が与えられたとき，尤度関数を最大にする値 $\hat{\theta}(x)$ を見つけることである：

$$L_n(\hat{\theta}(x)|x) = p(x|\hat{\theta}(x)) = \max\{p(x|\theta) : \theta \in \Theta\} = \max\{L_n(\theta|x) : \theta \in \Theta\}$$

$\hat{\theta}(x)$ を θ の最尤推定値といい， $\hat{\theta}(X)$ を θ の最尤推定量という．

例 6.6 確率変数 X が正規分布 $N(\theta, \sigma^2)$ に従うとする．ただし， σ^2 は既知とする．このとき，尤度関数は

$$L_1(\theta|x) = \frac{1}{\sigma} \psi\left(\frac{x-\theta}{\sigma}\right)$$

となる．ただし， $\psi(x) = (1/\sqrt{2\pi})e^{-x^2/2}$ である．このとき，最大は

$$\hat{\theta}(x) = x$$

のとき唯一達成される．したがって， $\hat{\theta}(X) = X$ は最尤推定量となる．

つぎに， X_1, X_2, \dots, X_n は独立同一に正規分布 $N(\theta, \sigma^2)$ に従うとする．ここでも σ^2 は既知とする．このとき，尤度関数は

$$\begin{aligned} L_n(\theta|\mathbf{x}) &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left[-\frac{1}{2} \sum_{i=1}^n \frac{(x_i - \theta)^2}{\sigma^2}\right] \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left[-\frac{1}{2} \frac{(\bar{x}_n - \theta)^2}{\sigma^2/n}\right] \exp\left[-\frac{1}{2} \sum_{i=1}^n \frac{(x_i - \bar{x}_n)^2}{\sigma^2}\right] \end{aligned}$$

となる．よって，最大は

$$\hat{\theta}(\mathbf{x}) = \bar{x}_n, \quad \bar{x}_n = \frac{1}{n}(x_1 + x_2 + \dots + x_n)$$

で達成される．したがって，最尤推定量は $\hat{\theta}(X) = \bar{X}_n$ となる．ただし， $\bar{X}_n = (1/n)(X_1 + X_2 + \dots + X_n)$ である．

尤度関数に対数をとったものを対数尤度とよび，

$$l_n(\theta) = \log L_n(\theta | \mathbf{x})$$

と記す⁴．特に， $X_i, i = 1, 2, \dots, n$ ，が独立同一に確率密度関数 $f(x|\theta)$ に従う場合には

$$l_n(\theta) = \log \prod_{i=1}^n f(x_i|\theta) = \sum_{i=1}^n \log f(x_i|\theta)$$

となる．

もし， Θ が開集合で $l_n(\theta)$ が θ に関して微分可能で $\hat{\theta}(\mathbf{x})$ が存在するならば， $\hat{\theta}(\mathbf{x})$ は方程式

$$\frac{\partial}{\partial \theta_i} l_n(\theta) = 0, \quad i = 1, 2, \dots, k$$

をみtas. この方程式を尤度方程式という．

例 6.7 標識 1, 2, 3 のどれかをもつ個体から構成される母集団を考える．それぞれの標識の出現確率は Hardy-Weinberg 比率で与えられるとする：

$$p(1|\theta) = \theta^2, \quad p(2|\theta) = 2\theta(1-\theta), \quad p(3|\theta) = (1-\theta)^2, \quad 0 < \theta < 1$$

たとえば，3 つの個体を観測し， $x_1 = 1, x_2 = 2, x_3 = 1$ を得たとする．このとき

$$L_3(\theta | \mathbf{x}) = p(1|\theta)p(2|\theta)p(1|\theta) = 2\theta^5(1-\theta)$$

となる．尤度方程式は

$$\frac{\partial}{\partial \theta} l_3(\theta) = \frac{5}{\theta} - \frac{1}{1-\theta} = 0$$

となり，唯一の解 $\hat{\theta} = 5/6$ を得る．これは

$$\frac{\partial^2}{\partial \theta^2} l_3(\theta) = -\frac{5}{\theta^2} - \frac{1}{(1-\theta)^2} < 0, \quad 0 < \theta < 1$$

よりわかる．

一般に， n 個の観測 x_1, x_2, \dots, x_n を得たとする．いま

$$n_j = \#\{x_i = j : i = 1, 2, \dots, n\}, \quad j = 1, 2, 3$$

⁴形式的に， $0/0 = 0, 0 \times \infty = 0$ とする．

とする．尤度関数は

$$L_n(\theta | \mathbf{x}) = \theta^{2n_1} \{2\theta(1-\theta)\}^{n_2} (1-\theta)^{2n_3} = 2^{n_2} \theta^{2n_1+n_2} (1-\theta)^{n_2+2n_3}$$

より

$$\frac{\partial}{\partial \theta} l_n(\theta) = \frac{2n_1 + n_2}{\theta} - \frac{n_2 + 2n_3}{1-\theta} = \frac{1}{\theta(1-\theta)} \{(2n_1 + n_2) - 2(n_1 + n_2 + n_3)\theta\}$$

より， $2n_1 + n_2 > 0$ ， $n_2 + 2n_3 > 0$ のとき，最尤推定値は唯一存在して，

$$\hat{\theta}(\mathbf{x}) = \frac{2n_1 + n_2}{2(n_1 + n_2 + n_3)}$$

となる．もし， $2n_1 + n_2 = 0$ のとき，尤度関数は

$$2^{n_2} (1-\theta)^{(2n_1+n_2)+(n_2+2n_3)} = 2^{n_2} (1-\theta)^{n_2+2n_3}$$

となり， $\theta = 0$ のとき，尤度関数は最大になり， $\Theta = (0, 1)$ なので，最尤推定値は存在しない．また， $n_2 + 2n_3 = 0$ のときは，尤度関数は $2^{n_2} \theta^{2n_1+n_2}$ となり，最尤推定値は存在しない．

例 6.8 X_1, X_2, \dots, X_n は独立同一分布に従い，各 $X_i, i = 1, 2, \dots, n$ は $1, 2, \dots, k$ の値をとり，その確率は $\theta_j = \mathbb{P}\{X_i = j\}, j = 1, 2, \dots, k$ ，で与えられるとする．ここで， $n \geq k-1$ を仮定する．いま， $N_j = \sum_{i=1}^n \mathbb{1}\{X_i = j\}$ おく． $\mathbf{X} = \mathbf{x}$ が与えられたとする．このとき， $n_j = \sum_{i=1}^n \mathbb{1}\{x_i = j\}$ とおけば，対数尤度は

$$l_n(\theta) = \log L_n(\theta | \mathbf{x}) = \log p(\mathbf{x} | \theta) = \sum_{j=1}^k n_j \log \theta_j$$

となる．ただし， $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ で

$$\theta_k = 1 - \sum_{j=1}^{k-1} \theta_j \quad (6.1)$$

である．

まず， $n_j > 0, j = 1, 2, \dots, k$ を仮定する．このとき， θ_j のどれかがゼロならば， $p(\mathbf{x} | \theta) = 0$ となる．したがって，最尤推定値は $\theta_j > 0$ となるので，上

の仮定のもとでは，最尤推定値は $[0, 1]^k$ の内点である．したがって，最尤推定値は尤度方程式

$$\frac{\partial}{\partial \theta_j} l_n(\theta) = \frac{\partial}{\partial \theta_j} \sum_{\ell=1}^k n_\ell \log \theta_\ell = \sum_{\ell=1}^k \frac{n_\ell}{\theta_\ell} \frac{\partial \theta_\ell}{\partial \theta_j} = 0, \quad j = 1, 2, \dots, k-1$$

となる．(6.1) から $\partial \theta_k / \partial \theta_j = -1, j = 1, 2, \dots, k-1$ ，となる．よって，尤度方程式は

$$\frac{\hat{\theta}_k}{\hat{\theta}_j} = \frac{n_k}{n_j}, \quad j = 1, 2, \dots, k-1$$

となる．これを再度 (6.1) に代入すれば， $\hat{\theta}_k = n_k/n$ となり，

$$\hat{\theta}_j = \frac{n_j}{n}, \quad j = 1, 2, \dots, k$$

となる．ただし， $n = \sum_{\ell=1}^k n_\ell$ とした．つぎに， $\theta_j = n_j/n, j = 1, 2, \dots, k$ が実際に $l_n(\theta)$ を最大にしていることを確認するために， $l_n(\theta)$ は $(\theta_1, \theta_2, \dots, \theta_{k-1})$ に関して concave であることを示す． $1 \leq r \leq k-1, 1 \leq j \leq k-1$ に対して，

$$\frac{\partial}{\partial \theta_r} \frac{\partial}{\partial \theta_j} l_n(\theta) = \frac{\partial}{\partial \theta_r} \left(\frac{n_j}{\theta_j} - \frac{n_k}{\theta_k} \right) = \begin{cases} -\left(\frac{n_r}{\theta_r^2} + \frac{n_k}{\theta_k^2} \right) < 0, & r = j \\ -\frac{n_k}{\theta_k^2} < 0, & r \neq j \end{cases}$$

となる．

ある j に対して， $n_j = 0$ のとき， $\hat{\theta}_j = n_j/n$ が最尤推定値であることも確認することができる．

定理 6.2 $T(X)$ を未知母数 θ の十分統計量とする．このとき， θ の最尤推定量が一意に存在するならば， θ の最尤推定量は T の関数である．

証明 $p(x|\theta)$ を確率関数または確率密度関数とする．因子分解定理から θ と T を通してのみ x に依存する関数 g と x のみに依存する関数 h が存在して，

$$p(x|\theta) = h(x)g(T(x)|\theta)$$

と書ける．これより θ に関して $p(x|\theta)$ を最大化することは θ に関して $g(T(x)|\theta)$ を最大化することと同値になる．また， $T(x) = t$ と与えられたとき， $g(t|\theta)$ が 2 つ以上 θ で最大になるとすれば，それに対応する x において $g(T(x)|\theta)$ も

2 つ以上の θ で最大化されるので、仮定と矛盾する．したがって、 $g(t|\theta)$ を θ について最大にする点はひとつである． \square

定理 6.3 関数 $g: \mathbb{R} \rightarrow \mathbb{R}$ が 1 対 1 のとき、 $\hat{\theta}$ が θ の最尤推定量であれば、 $g(\hat{\theta})$ は $g(\theta)$ の最尤推定量である．

証明 g は 1 対 1 だから、逆関数 g^{-1} が存在し、 $\tau = g(\theta)$ のとき、 $\theta = g^{-1}(\tau)$ となる．これより

$$L_n(\theta|\mathbf{x}) = p(\mathbf{x}|\theta) = p(\mathbf{x}|g^{-1}(\tau)) = \tilde{L}_n(\tau|\mathbf{x})$$

と書けるので、

$$\sup_{\tau} \tilde{L}_n(\tau|\mathbf{x}) = \sup_{\tau} L_n(g^{-1}(\tau)|\mathbf{x}) = \sup_{\theta} L_n(\theta|\mathbf{x})$$

よって、 $\hat{\theta} = g^{-1}(\hat{\tau})$ のとき最大化される．したがって、 $\hat{\tau} = g(\hat{\theta})$ のとき最大化される． \square

例 6.9 (指数分布族の最尤推定量) $\mathbf{X} = (X_1, X_2, \dots, X_n)'$ は k 母数指数分布族

$$f(x|\boldsymbol{\eta}) = S(x) \exp\left[\sum_{j=1}^k \eta_j T_j(x) - A(\boldsymbol{\eta})\right]$$

からの大きさ n のランダム標本とする．ただし、 $\boldsymbol{\eta} = (\eta_1, \eta_2, \dots, \eta_k)$ とし、自然母数空間 $\boldsymbol{\eta} \in \mathcal{E}$ は \mathbb{R}^k の開集合とする．である．このとき、 \mathbf{X} の同時確率(密度)関数は

$$\begin{aligned} p(\mathbf{x}|\boldsymbol{\eta}) &= \prod_{i=1}^n f(x_i|\boldsymbol{\eta}) \\ &= \prod_{i=1}^n S(x_i) \exp\left[\sum_{j=1}^k n\eta_j \bar{T}_j(\mathbf{x}) - nA(\boldsymbol{\eta})\right] \end{aligned}$$

となる．ただし、 $\bar{T}_j(\mathbf{x}) = \sum_{i=1}^n T_j(x_i)$ 、 $j = 1, 2, \dots, k$ である．したがって、対数尤度は

$$l_n(\boldsymbol{\eta}) = n \sum_{j=1}^k \eta_j \bar{T}_j(\mathbf{x}) - nA(\boldsymbol{\eta}) + (\text{定数項})$$

となる．系 ?? を用いれば，

$$\frac{\partial}{\partial \eta_j} l_n(\boldsymbol{\eta}) = n\bar{T}_j(\mathbf{x}) - n \frac{\partial}{\partial \eta_j} A(\boldsymbol{\eta}) = n\bar{T}_j(\mathbf{x}) - n\mathbb{E}_\eta[T_j(X)]$$

と

$$\frac{\partial^2}{\partial \eta_j \partial \eta_m} l_n(\boldsymbol{\eta}) = -n \frac{\partial^2}{\partial \eta_j \partial \eta_m} A(\boldsymbol{\eta}) = -n \text{COV}_\eta(T_j(X), T_m(X))$$

となるので，行列 $((\partial^2/\partial \boldsymbol{\eta}' \boldsymbol{\eta})A(\boldsymbol{\eta}))$ は負の定符号となる．したがって，尤度方程式

$$\frac{1}{n} \sum_{i=1}^n T_j(x_i) = \frac{\partial}{\partial \eta_j} A(\boldsymbol{\eta})$$

または

$$\frac{1}{n} \sum_{i=1}^n T_j(x_i) = \mathbb{E}_\eta[T_j(X)], \quad j = 1, 2, \dots, k$$

は唯一の解を持ち，これは $\boldsymbol{\eta}$ の最尤推定値となる．

3 不偏推定と Cramér-Rao の定理

この節では，簡単のために， $\theta \in \Theta \subset \mathbb{R}$ とする． X_1, X_2, \dots, X_n を確率（密度）関数 $f(x|\theta)$ からのランダム標本とし，標本 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ に基づく $\tau(\theta)$ の推定量を $T = T(\mathbf{X})$ とする． T の平均自乗誤差 (MSE) は

$$\mathbb{E}_\theta[\{T(\mathbf{X}) - \tau(\theta)\}^2] = \text{VAR}_\theta[T(\mathbf{X})] + \{\mathbb{E}_\theta[T(\mathbf{X})] - \tau(\theta)\}^2$$

と書けた．さらに， T が $\tau(\theta)$ の不偏推定量であれば，

$$\mathbb{E}_\theta[\{T(\mathbf{X}) - \tau(\theta)\}^2] = \text{VAR}_\theta[T(\mathbf{X})]$$

となる．

定義 6.8 (一様最小分散不偏推定量 Uniformly Minimum Variance Unbiased Estimator) $\theta \in \Theta \subset \mathbb{R}$ とする． X_1, X_2, \dots, X_n を確率（密度）関数 $f(x|\theta)$ からのランダム標本とし， $\tau(\theta)$ の推定量 $T^* = T^*(\mathbf{X})$ が $\tau(\theta)$ の一様最小分散不偏推定量（略して UMVUE）であるとはつぎを満足することである：

- (i) $\mathbb{E}_\theta[T^*] = \tau(\theta)$ (すべての $\theta \in \Theta$)
- (ii) すべての $\theta \in \Theta$ に対して， $\tau(\theta)$ の任意の不偏推定量 $T = T(\mathbf{X})$ に対し

$$\text{VAR}_\theta[T^*] \leq \text{VAR}_\theta[T], \quad (\text{すべての } \theta \in \Theta)$$

3.1 不偏推定量の分散の下限

この節では、各 X_i ($i = 1, 2, \dots, n$) は連続型分布とし、確率密度関数 $f(x|\theta)$ を持つものとして議論を進める。各 X_i ($i = 1, 2, \dots, n$) が離散型分布の場合には、積分を和の記号に置き換えればよい。

定理 6.4 $\theta \in \Theta \subset \mathbb{R}$ とする。 X_1, X_2, \dots, X_n を確率密度関数 $f(x|\theta)$ からのランダム標本とし、以下の仮定をおく：

- (A1) 集合 $\{x : f(x|\theta) > 0\}$ は θ に依存しない。
- (A2) すべての x と θ に対し、 $(d/d\theta)f(x|\theta)$ は存在する。
- (A3) $\mathbb{E}_\theta[|T(\mathbf{X})|] < \infty$ なる推定量 T に対し

$$\begin{aligned} \frac{d}{d\theta} \int \cdots \int T(x_1, x_2, \dots, x_n) \prod_{i=1}^n f(x_i|\theta) dx_i \\ = \int \cdots \int T(x_1, x_2, \dots, x_n) \frac{d}{d\theta} \prod_{i=1}^n f(x_i|\theta) dx_i. \end{aligned}$$

- (A4) すべての $\theta \in \Theta$ に対し

$$0 < \mathbb{E}_\theta \left[\left\{ \frac{d}{d\theta} \log f(X_1|\theta) \right\}^2 \right] < \infty.$$

- (A5) $\tau(\theta)$ は微分可能で $\tau(\theta) = \mathbb{E}_\theta[T(\mathbf{X})]$.

このとき、 θ に関して一様に

$$\text{VAR}_\theta[T] \geq \frac{[\tau'(\theta)]^2}{n \mathbb{E}_\theta \left[\left(\frac{d}{d\theta} \log f(X_1|\theta) \right)^2 \right]}.$$

が成立する。また、等号成立は

$$\frac{d}{d\theta} \log f(\mathbf{X}|\theta) = K(\theta, n)[T(\mathbf{X}) - \tau(\theta)] \quad (6.2)$$

のときに限る。

証明 $f(\mathbf{x}|\theta) = \prod_{i=1}^n f(x_i|\theta)$ と $d\mathbf{x} = dx_1 dx_2 \cdots dx_n$ と記すことにする.

$$\begin{aligned}
 \tau'(\theta) &= \frac{d}{d\theta} \int \cdots \int T(\mathbf{x}) f(\mathbf{x}|\theta) d\mathbf{x} \\
 &= \int \cdots \int T(\mathbf{x}) \frac{d}{d\theta} f(\mathbf{x}|\theta) d\mathbf{x} - \tau(\theta) \frac{d}{d\theta} \int \cdots \int f(\mathbf{x}|\theta) d\mathbf{x} \\
 &= \int \cdots \int [T(\mathbf{x}) - \tau(\theta)] \frac{d}{d\theta} f(\mathbf{x}|\theta) d\mathbf{x} \\
 &= \int \cdots \int [T(\mathbf{x}) - \tau(\theta)] \left(\frac{d}{d\theta} \log f(\mathbf{x}|\theta) \right) f(\mathbf{x}|\theta) d\mathbf{x} \\
 &= \mathbb{E}_\theta \left[\{T(\mathbf{X}) - \tau(\theta)\} \left(\frac{d}{d\theta} \log f(\mathbf{X}|\theta) \right) \right]
 \end{aligned}$$

となる. Cauchy-Schwarz の不等式から

$$[\tau'(\theta)]^2 \leq \mathbb{E}_\theta[\{T(\mathbf{X}) - \tau(\theta)\}^2] \mathbb{E}_\theta \left[\left(\frac{d}{d\theta} \log f(\mathbf{X}|\theta) \right)^2 \right]$$

よって

$$\text{VAR}_\theta[T] \geq \frac{[\tau'(\theta)]^2}{\mathbb{E}_\theta \left[\left(\frac{d}{d\theta} \log f(\mathbf{X}|\theta) \right)^2 \right]}.$$

しかし

$$\begin{aligned}
 \mathbb{E}_\theta \left[\left(\frac{d}{d\theta} \log f(\mathbf{X}|\theta) \right)^2 \right] &= \mathbb{E}_\theta \left[\left\{ \sum_{i=1}^n \frac{d}{d\theta} \log f(X_i|\theta) \right\}^2 \right] \\
 &= \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}_\theta \left[\frac{d}{d\theta} \log f(X_i|\theta) \frac{d}{d\theta} \log f(X_j|\theta) \right] \\
 &= n \mathbb{E}_\theta \left[\left\{ \frac{d}{d\theta} \log f(X_1|\theta) \right\}^2 \right].
 \end{aligned}$$

なぜならば, $i \neq j$ のとき X_i と X_j は独立であることと

$$\begin{aligned}
 \mathbb{E}_\theta \left[\frac{d}{d\theta} \log f(X_1|\theta) \right] &= \int \left(\frac{d}{d\theta} \log f(x|\theta) \right) f(x|\theta) dx = \int \frac{d}{d\theta} f(x|\theta) dx \\
 &= \frac{d}{d\theta} \int f(x|\theta) dx = 0
 \end{aligned}$$

最後に Cauchy-Schwarz の不等式の等号成立条件から

$$\frac{d}{d\theta} \log f(\mathbf{X}|\theta) = K(\theta, n)[t(\mathbf{X}) - \tau(\theta)] + b(\theta) \quad (6.3)$$

が等号成立条件となることがわかる。しかし、両辺の期待値をとれば、

$$0 = \mathbb{E}_\theta \left[\frac{d}{d\theta} \log f(\mathbf{X}|\theta) \right] = K(\theta, n)\mathbb{E}_\theta [t(\mathbf{X}) - \tau(\theta)] + b(\theta) = b(\theta)$$

となるので、 $b(\theta) = 0$ がわかる。よって、等号成立条件は (6.2) であることが示せた。□

注意 6.7 T が $\tau(\theta)$ の一様最小分散不偏推定量ならば、一意である。これは以下の議論からわかる。 T' を別の $\tau(\theta)$ の一様最小分散不偏推定量とし、 $T^* = (1/2)(T + T')$ とおく。すると、 $\mathbb{E}[T^*] = \tau(\theta)$ となり、 T^* も $\tau(\theta)$ の不偏推定量となる。したがって、すべての θ に対して、

$$\text{VAR}_\theta(T^*) \geq \text{VAR}_\theta(T) \quad (6.4)$$

となる。しかし、

$$\begin{aligned} \text{VAR}_\theta(T^*) &= \text{VAR}_\theta((1/2)T + (1/2)T') \\ &= \frac{1}{4}\text{VAR}_\theta(T) + \frac{1}{4}\text{VAR}_\theta(T') + \frac{1}{2}\text{COV}_\theta(T, T') \\ &\leq \frac{1}{4}\text{VAR}_\theta(T) + \frac{1}{4}\text{VAR}_\theta(T') + \frac{1}{2}\sqrt{\text{VAR}_\theta(T)}\sqrt{\text{VAR}_\theta(T')} \\ &= \text{VAR}_\theta(T) \end{aligned} \quad (6.5)$$

となる。不等号は Cauchy-Schwartz の不等式からわかる。また、最後の等号は $\text{VAR}_\theta(T) = \text{VAR}_\theta(T')$ よりわかる。(6.4) と (6.5) から

$$\text{VAR}_\theta(T) = \text{VAR}_\theta(T^*) \quad (6.6)$$

が成立することがわかる。したがって、(6.5) の Cauchy-Schwartz の不等式は等号が成立している。Cauchy-Schwartz の不等式の等号成立条件から、ある定数 $a(\theta)$ と $b(\theta)$ が存在して

$$T' = a(\theta)T + b(\theta)$$

と書ける。しかし,

$$\begin{aligned}\text{COV}_\theta[T, T'] &= \text{COV}_\theta[T, a(\theta)T + b(\theta)] \\ &= \text{COV}_\theta[T, a(\theta)T] \\ &= a(\theta)\text{VAR}_\theta[T]\end{aligned}$$

となる。しかし, (6.6) において, 等号が成立するので, $\text{COV}_\theta[T, T'] = \text{VAR}_\theta[T]$ となるので, $a(\theta) = 1$ がわかる。さらに,

$$\tau(\theta) = \mathbb{E}_\theta[T'] = \mathbb{E}_\theta[T] + b(\theta) = \tau(\theta) + b(\theta)$$

から $b(\theta) = 0$ となる。したがって, $T = T'$ となることがわかる。

例 6.10 X_1, X_2, \dots, X_n を指数分布 $f(x|\theta) = \theta e^{-\theta x} \mathbf{1}_{(0, \infty)}(x)$ からの大きさ n のランダム標本とする。ただし, $\theta > 0$ である。

まず, $\tau(\theta) = \theta$ とする。 θ の推定量 $T = T(X_1, X_2, \dots, X_n)$ の Cramér-Rao の下限は

$$\text{VAR}_\theta(T) \geq \frac{1}{n \mathbb{E}_\theta \left[\left\{ \frac{d}{d\theta} \log f(X_1|\theta) \right\}^2 \right]}$$

となる。さらに,

$$\frac{d}{d\theta} \log f(x|\theta) = \frac{1}{\theta} - x$$

より

$$\mathbb{E}_\theta \left[\left\{ \frac{d}{d\theta} \log f(X_1|\theta) \right\}^2 \right] = \mathbb{E}_\theta \left[\left\{ \frac{1}{\theta} - X_1 \right\}^2 \right] = \text{VAR}(X_1) = \frac{1}{\theta^2}$$

となる⁵。したがって,

$$\text{VAR}_\theta[T] \geq \frac{1}{n(1/\theta^2)} = \frac{\theta^2}{n}$$

⁵

$$\mathbb{E}(X_1) = \theta \int_0^\infty x \theta e^{-\theta x} dx = (1/\theta) \int_0^\infty y e^{-y} dy = (1/\theta)\Gamma(2) = (1/\theta)$$

と

$$\mathbb{E}(X_1^2) = (1/\theta^2) \int_0^\infty y^2 e^{-y} dy = (1/\theta^2)\Gamma(3) = (1/\theta^2)2!$$

からわかる。

となる .

つぎに , $\tau(\theta) = 1/\theta$ とする . $\tau(\theta)$ の推定量 $S = S(X_1, X_2, \dots, X_n)$ の Cramér - Rao の下限は

$$\text{VAR}_\theta(S) \geq \frac{\{\tau'(\theta)\}^2}{n(1/\theta^2)} = \frac{1}{n\theta^2}$$

となる . さらに ,

$$\sum_{i=1}^n \frac{d}{d\theta} \log f(X_i|\theta) = \sum_{i=1}^n \frac{d}{d\theta} (\log \theta - \theta X_i) = \sum_{i=1}^n \left(\frac{1}{\theta} - X_i \right) = -n \left(\bar{X}_n - \frac{1}{\theta} \right)$$

より , (6.2) において $K(\theta, n) = -n$ とおけば , $S = \bar{X}_n$ は $1/\theta$ の UMVUE となる .

例 6.11 X_1, X_2, \dots, X_n をポアソン分布 $f(x|\theta) = e^{-\theta}\theta^x/x!(x=0, 1, \dots)$ からの大きさ n のランダム標本とする . ただし , $\theta > 0$ である .

$\tau(\theta) = e^{-\theta} = \mathbb{P}[X_1 = 0]$ とする .

$$\frac{d}{d\theta} \log f(x|\theta) = -1 + \frac{x}{\theta}$$

に注意すれば ,

$$\mathbb{E}_\theta \left[\left\{ \frac{\partial}{\partial \theta} \log f(X_1|\theta) \right\}^2 \right] = \mathbb{E}_\theta \left[\left\{ \frac{X_1}{\theta} - 1 \right\}^2 \right] = \frac{1}{\theta^2} \mathbb{E}_\theta [(X_1 - \theta)^2] = \frac{1}{\theta}$$

となる⁶ . よって , $\tau(\theta)$ の推定量推定量 $T = T(X_1, X_2, \dots, X_n)$ の Cramér-Rao) の下限は

$$\text{VAR}_\theta(T) \geq \frac{(-e^{-\theta})^2}{n(1/\theta)} = \frac{\theta e^{-2\theta}}{n}$$

となる .

いま , $\tau(\theta)$ の推定量推定量

$$T_0 = \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{X_i = 0\}$$

⁶ポアソン分布の平均と分散は θ になることからわかる .

を考える．

$$\mathbb{E}_\theta[T_0] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}_\theta[\mathbb{1}\{X_i = 0\}] = \frac{1}{n} \sum_{i=1}^n [1 \cdot \mathbb{P}_\theta(X_i = 0) + 0 \cdot \mathbb{P}_\theta(X_i \neq 0)] = e^{-\theta}$$

となる．したがって， T_0 は $\tau(\theta) = e^{-\theta}$ の不偏推定量となる．また，

$$\begin{aligned} \text{VAR}_\theta[T_0] &= \mathbb{E}_\theta \left[\left\{ \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{X_i = 0\} \right\}^2 \right] - \{\mathbb{E}_\theta(T_0)\}^2 \\ &= \sum_{i=1}^n \mathbb{E}_\theta \left[\frac{1}{n^2} \mathbb{1}\{X_i = 0\} \right] + \sum_{i \neq j} \mathbb{E}_\theta \left[\frac{1}{n^2} \mathbb{1}\{X_i = 0\} \mathbb{1}\{X_j = 0\} \right] - e^{-2\theta} \\ &= \frac{1}{n} e^{-\theta} + \frac{n(n-1)}{n^2} e^{-2\theta} - e^{-2\theta} = \frac{1}{n} e^{-\theta} (1 - e^{-\theta}) \end{aligned}$$

となる． $\theta > 0$ に対し

$$\frac{1}{n} e^{-\theta} (1 - e^{-\theta}) \geq \frac{1}{n} \theta e^{-2\theta}$$

である⁷．

つぎに， $\tau(\theta) = \theta$ とおく．

$$\sum_{i=1}^n \frac{d}{d\theta} \log f(X_i | \theta) = \sum_{i=1}^n \left(-1 + \frac{X_i}{\theta} \right) = \frac{n}{\theta} (\bar{X}_n - \theta)$$

となり，(6.2) において $K(\theta, n) = n/\theta$ とおけば， \bar{X}_n は $\tau(\theta) = \theta$ の UMVUE となることがわかる．

4 十分性と完備性

定理 6.5 (Rao-Blackwell) X_1, X_2, \dots, X_n を確率 (密度) 関数 $f(x|\theta)$, $\theta \in \Theta$ からの大きさ n ($n \geq 2$) のランダム標本とし， $S_i = S_i(X_1, X_2, \dots, X_n)$, $i = 1, 2, \dots, k$ を十分統計量とする． $T = T(X_1, X_2, \dots, X_n)$ を $\tau(\theta)$ の不偏推定量で $\text{VAR}(T) < \infty$ なるものとする．推定量 $T' = T'(X_1, X_2, \dots, X_n)$ を $T' = \mathbb{E}[T | S_1, \dots, S_k]$ で定義する．このとき，

⁷これは $e^{-\theta}(1 - e^{-\theta}) - \theta e^{-2\theta} = e^{-\theta}[1 - e^{-\theta} - \theta e^{-\theta}]$ となるので， $f(\theta) = 1 - e^{-\theta} - \theta e^{-\theta}$ として評価すれば， $f(\theta) \geq 0$ となることよりわかる．

- (i) T' は統計量で S_1, \dots, S_k を通してのみ (X_1, X_2, \dots, X_n) に依存する .
したがって, $T' = T'(S_1, \dots, S_k)$ を記すことにする .
- (ii) $\mathbb{E}_\theta[T'] = \tau(\theta)$ となる . すなわち, T' は $\tau(\theta)$ の不偏推定量である .
- (iii) すべての $\theta \in \Theta$ に対し

$$\text{VAR}(T') \leq \text{VAR}(T)$$

が成立する . ある θ に対し, $\mathbb{P}(T = T') \neq 1$ ならば,

$$\text{VAR}(T') < \text{VAR}(T)$$

が成立する .

証明 S_1, \dots, S_k の十分性から S_1, \dots, S_k を与えたときの (X_1, X_2, \dots, X_n) の条件付き分布は θ に依存しない . したがって, S_1, \dots, S_k の十分性から S_1, \dots, S_k を与えたときの T の条件付き分布は θ に依存しないので, $\mathbb{E}_\theta[T | S_1, \dots, S_k]$ は θ に依存しないので, 統計量となる . また, これは S_1, \dots, S_k の関数となることは明らかである .

$$T' = \mathbb{E}_\theta[\mathbb{E}[T | S_1, \dots, S_k]] = \mathbb{E}_\theta[T] = \tau(\theta)$$

となる . したがって, T' は $\tau(\theta)$ の不偏推定量となる .

$$\begin{aligned} \text{VAR}_\theta[T] &= \mathbb{E}_\theta[(T - \mathbb{E}_\theta(T))^2] \\ &= \mathbb{E}_\theta[(T - T' + T' - \mathbb{E}_\theta(T))^2] \\ &= \mathbb{E}_\theta[\{T - T'\}^2] + 2\mathbb{E}_\theta[(T - T')(T' - \mathbb{E}_\theta(T))] + \text{VAR}_\theta(T') \end{aligned}$$

となる . しかし,

$$\begin{aligned} \mathbb{E}_\theta[(T - T')(T' - \mathbb{E}_\theta(T))] &= \mathbb{E}_\theta[\mathbb{E}[(T - T')(T' - \mathbb{E}_\theta(T)) | S_1, \dots, S_k]] \\ &= \mathbb{E}_\theta[(T' - \mathbb{E}_\theta(T))\mathbb{E}[T - T' | S_1, \dots, S_k]] = 0 \end{aligned}$$

となる . よって,

$$\text{VAR}_\theta[T] = \mathbb{E}_\theta[\{T - T'\}^2] + \text{VAR}_\theta(T') \geq \text{VAR}_\theta(T')$$

を得る．さらに，

$$\mathbb{E}_\theta\{[T - T']^2\} = 0 \iff \mathbb{P}_\theta(T - T' = 0) = 1$$

に注意すればよい． □

注意 6.8 Rao-Blackwell の定理の証明の最後の部分は以下のように確認できる． W を確率変数とし， $\mathbb{E}[W^2] < \infty$ とする．このとき，

$$\mathbb{E}[W^2] = 0 \iff P(W = 0) = 1$$

を示せばよい．まず， \Leftarrow は自明である．したがって， \Rightarrow を示せばよい．任意の正数 ϵ に対し

$$\mathbb{E}[W^2] \geq \mathbb{E}[W^2 \mathbf{1}\{W \geq \epsilon\}] \geq \epsilon^2 \mathbb{E}[\mathbf{1}\{W \geq \epsilon\}] = \epsilon^2 \mathbb{P}\{W \geq \epsilon\}$$

となることより， $\mathbb{P}\{W \geq \epsilon\} = 0$ となる．同様にすれば， $\mathbb{P}\{W \leq -\epsilon\} = 0$ を得る．したがって， $\mathbb{P}\{-\epsilon \leq W \leq \epsilon\} = 1$ となる．いま， $A_n = \{-1/n \leq W \leq 1/n\}$ ($n = 1, 2, \dots$) とおけば， $A_1 \supset A_2 \supset \dots$ となる．また， $\{W = 0\} = \bigcap_{n=1}^{\infty} A_n$ であり， $\mathbb{P}_\theta(A_n) = 1$ である．これらより

$$\mathbb{P}(W = 0) = \mathbb{P}\{\bigcap_{n=1}^{\infty} A_n\} = \lim_{n \rightarrow \infty} \mathbb{P}(A_n) = 1$$

となることよりわかる．

例 6.12 X_1, X_2, \dots, X_n をベルヌーイ試行 $f(x|\theta) = \theta^x(1-\theta)^{1-x} \mathbf{1}_{\{0,1\}}(x)$ からの大きさ n ($n \geq 2$) のランダム標本とし， $\tau(\theta) = \theta$ の推定を考える．明らかに， $T(X_1, X_2, \dots, X_n) = X_1$ は θ の不偏推定量になる．しかし，この推定量は効率がわるいので，十分統計量 $S = \sum_{i=1}^n X_i$ を用いて，修正することを考える．

$$\tilde{T} = \mathbb{E}[X_1 | S]$$

とおく．この推定量を具体的に計算してみよう．

$$\begin{aligned}
 \mathbb{P}[X_1 = 0 \mid \sum_{i=1}^n X_i = s] &= \frac{\mathbb{P}[X_1 = 0 \text{ かつ } \sum_{i=1}^n X_i = s]}{\mathbb{P}[\sum_{i=1}^n X_i = s]} \\
 &= \frac{\mathbb{P}[X_1 = 0 \text{ かつ } \sum_{i=2}^n X_i = s]}{\mathbb{P}[\sum_{i=1}^n X_i = s]} \\
 &= \frac{\mathbb{P}[X_1 = 0] \mathbb{P}[\sum_{i=2}^n X_i = s]}{\mathbb{P}[\sum_{i=1}^n X_i = s]} \\
 &= \frac{(1-\theta) \binom{n-1}{s} \theta^s (1-\theta)^{n-1-s}}{\binom{n}{s} \theta^s (1-\theta)^{n-s}} = \frac{n-s}{n}
 \end{aligned}$$

となる．また，

$$\begin{aligned}
 \mathbb{P}[X_1 = 1 \mid \sum_{i=1}^n X_i = s] &= \frac{\mathbb{P}[X_1 = 1 \text{ かつ } \sum_{i=1}^n X_i = s]}{\mathbb{P}[\sum_{i=1}^n X_i = s]} \\
 &= \frac{\mathbb{P}[X_1 = 1 \text{ かつ } \sum_{i=2}^n X_i = s-1]}{\mathbb{P}[\sum_{i=1}^n X_i = s]} \\
 &= \frac{\mathbb{P}[X_1 = 1] \mathbb{P}[\sum_{i=2}^n X_i = s-1]}{\mathbb{P}[\sum_{i=1}^n X_i = s]} \\
 &= \frac{\theta \binom{n-1}{s-1} \theta^{s-1} (1-\theta)^{n-1-(s-1)}}{\binom{n}{s} \theta^{s-1} (1-\theta)^{n-s}} = \frac{s}{n}
 \end{aligned}$$

となる．したがって

$$\tilde{T} = \mathbb{E}_\theta[X_1 \mid S] = 0 \cdot \mathbb{P}[X_1 = 0 \mid S] + 1 \cdot \mathbb{P}[X_1 = 1 \mid S] = \frac{S}{n}$$

定義 6.9 (完備性) X_1, X_2, \dots, X_n を確率 (密度) 関数 $f(x \mid \theta)$, $\theta \in \Theta$ からの大きさ n のランダム標本とし, $T = T(X_1, X_2, \dots, X_n)$ を統計量とする．統計量 T の標本分布の確率 (密度) 関数の族が完備であるとは, 次の条件をみたすことである．統計量 T の値域上で定義された可測関数 g が, すべての $\theta \in \Theta$ に対し, $\mathbb{E}_\theta[g(T)] = 0$ をみたすならば, $\mathbb{P}_\theta(g(T) = 0) = 1$ が成立する．

さらに、統計量 T が完備であるとは、統計量 T の標本分布の確率(密度)関数の族が完備であることをいう。

例 6.13 X_1, X_2, \dots, X_n をベルヌーイ試行 $f(x|\theta) = \theta^x(1-\theta)^{1-x}\mathbb{1}_{\{0,1\}}(x)$ からの大きさ n ($n \geq 2$) のランダム標本とする。ただし、 $0 < \theta < 1$ とする。たとえば、統計量 $\tilde{T} = X_1 - X_2$ は完備ではない。なぜならば、 $\mathbb{E}_\theta[X_1 - X_2] = 0$ であるが、 $\mathbb{P}(X_1 - X_2 = 0) \neq 1$ であることよりわかる。

次に、 $T = \sum_{i=1}^n X_i$ を考える。 $g(T)$ をすべての $\theta \in (0, 1)$ に対し $\mathbb{E}_\theta[g(T)] = 0$ をみたす可測関数とする。 T が完備であることを示すために、 $t = 0, 1, \dots, n$ に対して $g(t) = 0$ を示せばよい。いま、 $0 < \theta < 1$ に対し

$$0 = \mathbb{E}_\theta[g(T)] = \sum_{t=0}^n \binom{n}{t} \theta^t (1-\theta)^{n-t} = (1-\theta)^n \sum_{t=0}^n g(t) \binom{n}{t} \left(\frac{\theta}{1-\theta}\right)^t$$

に注意すれば、すべての $0 < \theta < 1$ に対し、

$$\sum_{t=0}^n g(t) \binom{n}{t} \left(\frac{\theta}{1-\theta}\right)^t = 0$$

となることがわかる。または

$$\sum_{t=0}^n g(t) \binom{n}{t} \alpha^t = 0, \quad \alpha = \frac{\theta}{1-\theta}$$

となる。 $\{\alpha^t, t = 0, 1, \dots, n\}$ は一次独立なので、

$$g(t) \binom{n}{t} = 0, \quad t = 0, 1, \dots, n$$

となる。 $\binom{n}{t} \neq 0$ なので、 $g(t) = 0$ ($t = 0, 1, \dots, n$) となる。

例 6.14 X_1, X_2, \dots, X_n を一様分布 $f(x) = \mathbb{1}_{(0,1)}(x)$ からの大きさ n のランダム標本とする。ただし、 $\Theta = \{\theta, \theta > 0\}$ である。 $Y_n = \max(X_1, X_2, \dots, X_n)$ が完備であることを示そう。そのために、すべての $\theta \in \Theta$ に対し、 $\mathbb{E}_\theta[g(Y_n)] = 0$ ならば、 $\mathbb{P}_\theta\{g(Y_n) = 0\} = 1$ を示せばよい。

$$\mathbb{E}_\theta[g(Y_n)] = \int g(y) f_{Y_n}(y) dy = \int g(y) \theta^{-n} n y^{n-1} \mathbb{1}_{(0,\theta)}(y) dy$$

となり, すべての $\theta \in \Theta$ に対し

$$\mathbb{E}_\theta[g(Y_n)] = 0 \iff \frac{n}{\theta^n} \int_0^\theta g(y)y^{n-1} dy = 0$$

となる. したがって, θ に関して微分すれば

$$g(y)\theta^{n-1} = 0 \iff g(y) = 0$$

となることがわかる.

定理 6.6 Lehmann-Sheffée X_1, X_2, \dots, X_n を $f(x|\theta)$ からの大きさ n のランダム標本とし, $\tau: \Theta \rightarrow \mathbb{R}$ とする. $S = S(X_1, X_2, \dots, X_n)$ が完備十分統計量で $T^* = T^*(S)$ が $\tau(\theta)$ の不偏推定量ならば, T^* は $\tau(\theta)$ の UMVUE である.

証明 $T' = T'(S)$ を $\tau(\theta)$ の不偏推定量とする. すると任意の $\theta \in \Theta$ に対し

$$\mathbb{E}_\theta[T' - T^*] = 0$$

となる. S の完備性から

$$\mathbb{P}_\theta[T'(S) = T^*(S)] = 0$$

となる. S の関数である不偏推定量は唯一である.

次に, T を任意の $\tau(\theta)$ の推定量とする. Rao-Blackwell の定理から $\mathbb{E}[T|S]$ は $\tau(\theta)$ の不偏推定量で S の関数となるので

$$T^* = \mathbb{E}[T|S], \quad a.s.$$

となる. さらに, Rao-Blackwell の定理から

$$\text{VAR}_\theta[T^*] \leq \text{VAR}_\theta[T]$$

となる. よって, T^* は UMVUE である. □

例 6.15 X_1, X_2, \dots, X_n をポアソン分布 $f(x|\theta) = e^{-\theta}\theta^x/x!$ ($x = 0, 1, \dots$) からの大きさ n ($n \geq 2$) のランダム標本とする. $\sum_{i=1}^n X_i$ は母数 $n\theta$ のポアソ

ン分布に従う．いま， $T = \sum_{i=1}^n X_i$ とする．

$$\begin{aligned} \mathbb{E}_\theta[g(T)] &= \sum_{t=0}^{\infty} g(t) \frac{e^{-n\theta} (n\theta)^t}{t!} = e^{-n\theta} \sum_{t=0}^{\infty} g(t) \frac{n^t}{t!} \theta^t = 0 \\ &\Leftrightarrow \sum_{t=0}^{\infty} g(t) \frac{n^t}{t!} \theta^t = 0 \quad \theta \text{ の多項式は一次独立} \\ &\Leftrightarrow g(t) = 0 \end{aligned}$$

となる．したがって， $\sum_{i=1}^n X_i$ は完備統計量である．

$\tau(\theta) = e^{-\theta} = \mathbb{P}_\theta\{X_1 = 0\}$ の推定を考えよう． $\mathbb{1}_{\{0\}}(X_1)$ は $e^{-\theta}$ の不偏推定量なので，Lehmann-Sheffé の定理から $T^* = \mathbb{E}[\mathbb{1}_{\{0\}}(X_1) | \sum_{i=1}^n X_i]$ は $\tau(\theta)$ の UMVUE となる． T^* がどのような推定量になるかを計算しよう．

$$\begin{aligned} \mathbb{P}\{X_1 = 0 \mid \sum_{i=1}^n X_i = s\} &= \frac{\mathbb{P}\{X_1 = 0 \text{ かつ } \sum_{i=1}^n X_i = s\}}{\mathbb{P}\{\sum_{i=1}^n X_i = s\}} \\ &= \frac{\mathbb{P}\{X_1 = 0 \text{ かつ } \sum_{i=2}^n X_i = s\}}{\mathbb{P}\{\sum_{i=1}^n X_i = s\}} \\ &= \frac{\mathbb{P}\{X_1 = 0\} \mathbb{P}\{\sum_{i=2}^n X_i = s\}}{\mathbb{P}\{\sum_{i=1}^n X_i = s\}} \\ &= \frac{e^{-\theta} e^{-(n-1)\theta} [(n-1)\theta]^s / s!}{e^{-n\theta} [n\theta]^s / s!} = \left(\frac{n-1}{n}\right)^s \end{aligned}$$

となる．したがって

$$\mathbb{E}[\mathbb{1}_{\{0\}}(X_1) | \sum_{i=1}^n X_i = s] = \mathbb{P}\{X_1 = 0 | \sum_{i=1}^n X_i = s\} = \left(\frac{n-1}{n}\right)^s$$

となる．故に

$$T^* = \left(\frac{n-1}{n}\right)^{\sum_{i=1}^n X_i}$$

は $\tau(\theta) = e^{-\theta}$ の UMVUE である．

5 ベイズ推定量

6 演習問題

問題 6.1 X_1, X_2, \dots, X_m を正規分布 $N(\theta, \sigma^2)$ からの標本の大きさが m のランダム標本とし, Y_1, Y_2, \dots, Y_n を正規分布 $N(\theta, \tau^2)$ からの標本の大きさが n のランダム標本とする. ここで, X_1, X_2, \dots, X_m と Y_1, Y_2, \dots, Y_n は互いに独立する. また, σ^2 と τ^2 は既知とする. 定数 a ($0 \leq a \leq 1$) に対して,

$$\hat{\theta}_a = a\bar{X}_m + (1-a)\bar{Y}_n,$$

とする. ただし, $\bar{X}_m = (1/m)\sum_{i=1}^m X_i$ と $\bar{Y}_n = (1/n)\sum_{i=1}^n Y_i$ である.

- (1) $\hat{\theta}_a$ は θ の不偏推定量であることを示せ.
- (2) $\hat{\theta}_a$ の MSE を求め, $\hat{\theta}_a$ の MSE を最小にする a は

$$a = \frac{\tau^2}{n} \bigg/ \left(\frac{\sigma^2}{m} + \frac{\tau^2}{n} \right)$$

となることを示せ.

問題 6.2 X_1, X_2, \dots, X_m と Y_1, Y_2, \dots, Y_n をそれぞれ平均 μ_1 , 分散 σ^2 の正規分布と平均 μ_2 , 分散 σ^2 の正規分布からのランダム標本とする. ここで, X_1, X_2, \dots, X_m と Y_1, Y_2, \dots, Y_n は独立であるとする.

- (1) μ_1, μ_2, σ^2 の対数尤度関数から μ_1, μ_2, σ^2 の最尤推定量は以下で与えられることを示せ.

$$\hat{\mu}_1 = \bar{X}, \quad \hat{\mu}_2 = \bar{Y}, \quad \hat{\sigma}^2 = \frac{1}{m+n} \left\{ \sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2 \right\}.$$

- (2) 最尤推定量のバイアスを求めよ. さらに, σ^2 の推定量

$$c \left\{ \sum_{i=1}^m (X_i - \bar{X})^2 + \sum_{j=1}^n (Y_j - \bar{Y})^2 \right\}.$$

が不偏推定になるように定数 c を定めよ.

問題 6.3 X_1, X_2, \dots, X_n を確率密度関数

$$f(x; \theta) = \frac{1}{\sigma} \exp\{-(x - \mu)/\sigma\} \mathbb{1}_{(\mu, \infty)}(x)$$

をもつ分布からのランダム標本とする. ただし, $-\infty < \mu < \infty$, $\sigma > 0$ である.

- (1) $\mathbb{E}(X - \mu)$ と $\text{VAR}(X)$ を求めよ.
- (2) μ は既知として, σ^2 の最尤推定量が $(\bar{X} - \mu)^2$ で与えられることを示せ. ただし, $\bar{X} = \sum X_i/n$ である.
- (3) 最尤推定量 $(\bar{X} - \mu)^2$ のバイアスを計算せよ (ヒント: $\sum (X_i - \mu)$ は母数 n , σ^{-1} のガンマ分布に従うことに注意せよ.)

(4)

$$\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

は σ^2 の不偏推定量になることを示せ .

- (5) μ は既知とし, $\mathbb{P}[X_1 > t] (t > \mu)$ の最尤推定量を求めよ . (ヒント : 確率を σ と t で表現し, それが σ の 1 対 1 の関数になっていることを利用せよ .)

問題 6.4 確率変数 X は次のような確率関数 $f(x|\theta)$ を持つとする .

X	-2	-1	0	1	2	計
$f(x \theta)$	$\frac{\theta}{6}$	$\frac{\theta}{3}$	$1-\theta$	$\frac{\theta}{3}$	$\frac{\theta}{6}$	1

ただし, $0 < \theta < 1$ である . 統計量

$$S = S(X) = |X|, \quad T = T(X) = \frac{3}{4}S$$

を考える .

- (1) S の確率分布を求めよ .
- (2) S の確率分布および S が与えられたときの X の条件付確率分布を求め, S は θ の十分統計量かどうかを調べよ .
- (3) θ の推定量 T の $\text{MSE}(T, \theta)$ (平均 2 乗誤差) を求めよ . (横軸を θ とし, 縦軸を $\text{MSE}(T, \theta)$ の値として, $\text{MSE}(T, \theta)$ のグラフに描くこと)

問題 6.5 X_1, X_2, \dots, X_n は確率密度関数

$$f(x|\theta) = \theta x^{\theta-1}, \quad 0 < x < 1, \quad 0 < \theta < \infty \quad (6.7)$$

からの大きさ n のランダム標本とする .

- (1) $X_1 = x_1, \dots, X_n = x_n$ を観測したときの尤度関数を述べよ .
- (2) θ の最尤推定量を求めよ . さらに, $\tau = \theta^{-1}$ の最尤推定量は

$$\hat{\tau}_n = -\frac{1}{n} \sum_{i=1}^n \log X_i$$

で与えられることを示せ .

- (3) 確率変数 X が確率密度関数 (6.7) を持つ分布に従うとき, X と $T = -\theta \log X$ の分布関数を求めよ . さらに, T の積率母関数 $M_T(r)$ を求めよ .
- (4) $\hat{\tau}_n$ は τ の不偏推定量であることを示した上でその分散を求めよ .
- (5) $\hat{\tau}_n$ は τ の一様最小分散不偏推定量であることを示せ .

ヒント

- 適当な正則条件のもとで $g(\theta)$ の不偏推定量 U (標本の大きさが n のランダム標本に基づく推定量) の分散の下限が

$$\text{VAR}_\theta[U] \geq \frac{\left(\frac{\partial g}{\partial \theta}(\theta)\right)^2}{n\mathbb{E}\left[\left(\frac{\partial}{\partial \theta} \log f(X_1|\theta)\right)^2\right]}$$

で与えられる.

- 母数 $(n, 1)$ のガンマ分布の確率密度関数と積率母関数は

$$f(x|n, 1) = \frac{1}{\Gamma(n)} x^{n-1} e^{-x} I_{(0, \infty)}(x) \quad \text{と} \quad M_X(r) = \left(\frac{1}{1-r}\right)^n, \quad (r < 1)$$

で与えられる.

- $u_{\alpha/2}$ と $l_{\alpha/2}$ を母数 $(n, 1)$ のガンマ分布の上側 $(\alpha/2) \times 100\%$ 点と上側 $(1-\alpha/2) \times 100\%$ とする. すなわち,

$$\int_{u_{\alpha/2}}^{\infty} \frac{1}{\Gamma(n)} x^{n-1} e^{-x} dx = \frac{\alpha}{2}, \quad \text{と} \quad \int_0^{l_{\alpha/2}} \frac{1}{\Gamma(n)} x^{n-1} e^{-x} dx = \frac{\alpha}{2}$$

とする.

問題 6.6

$X_1, X_2, \dots, X_n (n \geq 2)$ をポアソン分布

$$f_X(x|\lambda) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \dots, \quad \lambda > 0$$

からの大きさ n のランダム標本とする.

- (1) $\bar{X}_n = (1/n) \sum_{i=1}^n X_i$ としたとき, \bar{X}_n は λ の不偏推定量であることを示せ (ヒント: $\mathbb{E}[X_1] = \lambda$)
- (2) \bar{X}_n の分散を求めよ (ヒント: $\mathbb{E}[X_1(X_1 - 1)] = \lambda^2$)
- (3) $\mathbb{E}\{[(\partial/\partial \lambda) \log f_X(x|\lambda)]^2\}$ を求め, λ の不偏推定量の分散の下限を求めよ (ヒント: クラメル・ラオの下限は $1/(n\mathbb{E}\{[(\partial/\partial \lambda) \log f_X(x|\lambda)]^2\})$ である.)
- (4) λ の一様最小分散不偏推定量を求めよ.

第7章 検定法

母集団分布から無作為標本に基づき，母集団分布の未知の母数の検定法とその推定精度の評価法について学ぶ．

1 検定論の枠組み

この節では統計的決定理論の用語を用いて検定論の基本的な考え方と枠組みを説明する．

記号

Θ : 母数空間， \mathbb{R}^m の部分空間

\mathcal{X} : 標本空間， \mathbb{R}^n の部分空間でランダム標本 $X = (X_1, X_2, \dots, X_n)$ の値域である．

$F_X(x|\theta)$: X の累積分布関数．ただし， $\theta \in \Theta$ ， $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ である．特に， X_1, X_2, \dots, X_n が互いに独立の場合は

$$F_X(x|\theta) = \prod_{i=1}^n F_{X_i}(x_i|\theta)$$

となる．ここで， $F_{X_i}(x_i|\theta)$ を X_i の累積分布関数とした．

検定問題では，母数空間 Θ が互いに共通部分を持たない二つの部分集合 Θ_0, Θ_1 に分けられた場合，すなわち

$$\Theta = \Theta_0 \cup \Theta_1, \quad \Theta_0 \cap \Theta_1 = \emptyset$$

を考え，ランダム標本 X に基づいて未知の母数 θ が Θ_0 と Θ_1 のどちらに属しているかを判断する．未知の母数 θ が Θ_0 に属しているとする仮説を帰無仮説といい， $H_0 : \theta \in \Theta_0$ と書く．逆に， θ が Θ_1 に属していると仮説を対立仮説といい， $H_1 : \theta \in \Theta_1$ と書くことにする．検定問題では二つの仮説を同等と

みなさず、中心となる仮説を帰無仮説と呼び、相対する仮説を対立仮説と呼ぶ。以下では帰無仮説と対立仮説をあわせて

$$H_0: \theta \in \Theta_0 \quad \text{vs.} \quad H_1: \theta \in \Theta_1$$

と簡潔に表すことにする。

帰無仮説に対応する母数空間が1点集合 $\Theta_0 = \{\theta_0\}$ であるとき帰無仮説を単純帰無仮説であるという。同様に対立仮説に対応する母数空間が1点集合 $\Theta_1 = \{\theta_1\}$ であるとき対立仮説を単純対立仮説であるという。単純仮説でないものを複合仮説という。

例 7.1 コインを投げる実験を考える。表の出る確率を p とする。このとき、コインにゆがみがないということを帰無仮説とすれば、これは単純帰無仮説 $H_0: p = 1/2$ となる。

θ が1次元のときには片側検定と両側検定という用語もしばしば用いる。 θ が実数の検定問題が

$$H_0: \theta = \theta_0 \quad \text{vs.} \quad H_1: \theta \neq \theta_0$$

の形のとき両側検定といい、

$$H_0: \theta \leq \theta_0 \quad \text{vs.} \quad H_1: \theta > \theta_0$$

のとき片側検定という。ただし、 θ_0 は固定された値である。また、 θ は2次元で $\theta = (\mu, \sigma^2)$ であり、興味のある母数は μ の方で、検定問題とし、

$$H_0: \mu = \mu_0 \quad \text{vs.} \quad H_1: \mu \neq \mu_0 \quad (\sigma^2 \text{ は任意の正の数})$$

を考えることがある。ただし、 μ_0 は固定した値である。この場合、 μ に関しては1点の帰無仮説ではあるが、 σ^2 は未知であるために帰無仮説は複合仮説

$$\Theta_0 = \{(\mu_0, \sigma^2) : \sigma^2 > 0\}$$

となる。 σ^2 は未知であるが検定問題にとってさしあたり興味がない母数である。このような母数のことを局外母数または擾乱母数とよぶ。

ランダム標本 X に基づいて帰無仮説と対立仮説のいずれが正しいかを判断したいとき、取り得る決定(行動)は H_0 が正しいと判断するか、 H_1 が正しい

と判断するか of the どちらかである。 H_0 が正しいと判断することを「帰無仮説 H_0 を受容する」といい、逆に H_1 が正しいと判断することを「帰無仮説 H_0 を棄却する」という。以下では、帰無仮説 H_0 を受容する決定を 0 で表し、帰無仮説 H_0 を棄却する決定を 1 で表すことにする。また、 $\mathcal{D} = \{0, 1\}$ を決定（行動）空間とよぶことにする。

帰無仮説が正しいとき ($\theta \in \Theta_0$) に帰無仮説を棄却する誤りを第 1 種の誤りとよび、対立仮説が正しいとき ($\theta \in \Theta_1$) に帰無仮説を受容する誤りを第 2 種の誤りと呼ぶ。

$\Theta \setminus \mathcal{D}$	0	1
$\theta \in \Theta_0$	正しい	第 1 種の誤り
$\theta \in \Theta_1$	第 2 種の誤り	正しい

伝統的な検定論の考え方では、第 1 種の誤りを重視し、第 1 種の誤りを犯す確率をある与えられた限界 α ($0 < \alpha < 1$) 以下に押さえた上で、第 2 種の誤りを犯す確率をできるだけ小さくすることを目指す。この与えられた α の値を有意水準という。

検定問題において関数 $\psi: \mathcal{X} \rightarrow [0, 1]$ で、 $X = x$ を観測したとき確率 $\psi(x)$ で帰無仮説を棄却し、確率 $1 - \psi(x)$ で帰無仮説を受容するものを検定関数と呼ぶ。特に、 $\psi: \mathcal{X} \rightarrow \{0, 1\}$ を非確率化検定関数と呼ぶ。そうでないものを確率化検定関数という。

非確率化検定関数 ψ が与えられたときは、標本空間を ψ の値によって分割することになる。 $W = \{x: \psi(x) = 1\}$ と $W^c = \mathcal{X} \setminus W$ とおく。このとき、 W を棄却域とよび、 W^c を受容域とよぶ。標本空間の分割は特定の統計量 $T(X)$ の値によって定義されることが多い。たとえば、棄却域と受容域が

$$W = \{x: T(x) > c\}, \quad W^c = \{x: T(x) \leq c\}$$

のような具合である。ただし、 c はある定数である。このとき、この棄却域に対応する検定関数は、

$$\psi(x) = \mathbb{1}\{T(x) > c\}$$

となる。このような統計量 T のことを検定統計量といい、 c を棄却限界という。

いま、検定関数の $\{0, 1\}$ -値損失関数 $L(\theta, d)$ を考えよう。ただし、 $\theta \in \Theta$ と

$d \in \mathcal{D}$ である .

$$L(\theta, 0) = \begin{cases} 0, & (\theta \in \Theta_0) \\ 1, & (\theta \in \Theta_1) \end{cases}$$

$$L(\theta, 1) = 1 - L(\theta, 0) = \begin{cases} 1, & (\theta \in \Theta_0) \\ 0, & (\theta \in \Theta_1) \end{cases}$$

と定める . いま , $d: \mathcal{X} \rightarrow \{0, 1\}$ とし , $\mathbf{X} = \mathbf{x}$ が与えられたとき , 確率 $\psi(\mathbf{x})$ で $d(\mathbf{x}) = 1$ をとり , 確率 $1 - \psi(\mathbf{x})$ で $d(\mathbf{x}) = 0$ を取るとする . このとき , リスク関数を

$$\mathbf{R}(\theta, \psi) = \mathbb{E}_\theta[L(\theta, d(\mathbf{X}))]$$

で定める . $\theta \in \Theta_0$ のとき ,

$$\begin{aligned} \mathbf{R}(\theta, \psi) &= \mathbb{E}_\theta[\mathbb{E}[\mathbb{1}\{d(\mathbf{x}) = 1\} | \mathbf{X} = \mathbf{x}]] \\ &= \mathbb{E}_\theta[1 \times \mathbb{P}(d(\mathbf{x}) = 1 | \mathbf{X} = \mathbf{x}) + 0 \times \mathbb{P}(d(\mathbf{x}) = 0 | \mathbf{X} = \mathbf{x})] \\ &= \mathbb{E}_\theta[\psi(\mathbf{X})] \end{aligned}$$

となる . また , $\theta \in \Theta_1$ のとき ,

$$\begin{aligned} \mathbf{R}(\theta, \psi) &= \mathbb{E}_\theta[\mathbb{1}\{d(\mathbf{X}) = 0\}] = \mathbb{E}_\theta[\mathbb{E}[\mathbb{1}\{d(\mathbf{x}) = 0\} | \mathbf{X} = \mathbf{x}]] \\ &= \mathbb{E}_\theta[1 \times \mathbb{P}(d(\mathbf{x}) = 0 | \mathbf{X} = \mathbf{x}) + 0 \times \mathbb{P}(d(\mathbf{x}) = 1 | \mathbf{X} = \mathbf{x})] \\ &= \mathbb{E}_\theta[1 - \psi(\mathbf{X})] = 1 - \mathbb{E}_\theta[\psi(\mathbf{X})] \end{aligned}$$

となる . ここで

$$\beta_\psi(\theta) = \mathbb{P}_\theta[d(\mathbf{X}) = 1] = \mathbb{E}_\theta[\psi(\mathbf{X})]$$

とおけば ,

$$\mathbf{R}(\theta, \psi) = \begin{cases} \beta_\psi(\theta), & (\theta \in \Theta_0) \\ 1 - \beta_\psi(\theta), & (\theta \in \Theta_1) \end{cases}$$

となる . $\beta_\psi(\theta)$ のことを検出力関数 (または簡単に検出力および検定力) とよぶ .

例 7.2 ある工場で大量生産される製品の不良率を θ ($0 < \theta < 1$) とする . θ がある限界 θ_0 以下であれば生産工程は正常であるとし , θ_0 を越えた場合は生産工程は異常があるものとする . この場合 , 検定問題は

$$H_0: \theta \in \Theta_0 = (0, \theta_0) \quad \text{vs.} \quad H_1: \theta \in \Theta_1 = (\theta_0, 1)$$

となる．ここで 10 個の製品を検査しその中で不良品の数に対応する確率変数を X とし，もし X が 1 以上ならば帰無仮説 H_0 を棄却するような検定関数を考えよう．すなわち

$$\psi(x) = \begin{cases} 1, & (x \geq 1) \\ 0, & (x = 0) \end{cases}, \quad x = 0, 1, \dots, 10$$

となる． X は二項分布 $Bin(10, \theta)$ に従うと考えてよいので検出力関数は

$$\beta_\psi(\theta) = \mathbb{P}_\theta[X \geq 1] = 1 - \mathbb{P}_\theta[X = 0] = 1 - (1 - \theta)^{10}$$

で与えられる．

$\theta \in \Theta_0$ に対し $\beta_\psi(\theta)$ は第 1 種の誤りの確率であるから，有意水準 α に対し

$$\beta_\psi(\theta) \leq \alpha, \quad \theta \in \Theta_0$$

ならば検定関数 ψ は有意水準 α の検定の一つである．

$$\sup_{\theta \in \Theta_0} \beta_\psi(\theta)$$

を検定関数 ψ のサイズとよぶ．

例 7.3 X_1, X_2, \dots, X_n を正規分布 $N(\theta, 1)$ からのランダム標本とし，検定問題 $H_0: \theta = 0$, vs. $H_1: \theta \neq 0$ を考える．このとき，

$$\psi(\mathbf{x}) = \mathbb{1}\{|\bar{x}_n| > c\}$$

なる検定関数が考えられる．すなわち， $T(\mathbf{X}) = |\bar{X}_n|$ である．簡便にかけば

$$T(\mathbf{X}) > c \quad \text{ならば} \quad H_0 \text{ を棄却}$$

である．この場合の ψ の検出力は $\bar{X}_n \sim N(\theta, 1/n)$ となることから

$$\begin{aligned} \beta_\psi(\theta) &= \mathbb{P}_\theta[T(\mathbf{X}) > c] = \mathbb{P}_\theta[|\bar{X}_n| > c] = 1 - P_\theta[-c < \bar{X}_n < c] \\ &= 1 - \Phi(\sqrt{n}(c - \theta)) + \Phi(-\sqrt{n}(c + \theta)) \end{aligned}$$

となる．ただし， $\Phi(x) = \int_{-\infty}^x (1/\sqrt{2\pi})e^{-t^2/2} dt$ である．

例 7.4 X_1, X_2, \dots, X_n を正規分布 $N(\theta, 5^2)$ からのランダム標本とし, 検定問題 $H_0: \theta \leq 17$, vs. $H_1: \theta > 17$ を考える.

たとえば, 棄却域 $W = \{(x_1, x_2, \dots, x_n) : \bar{x}_n > 17 + 5/\sqrt{n}\}$ とし, 非確率化検定関数

$$\psi(\mathbf{x}) = \begin{cases} 1, & (\mathbf{x} \in W) \\ 0, & (\mathbf{x} \in W^c) \end{cases}$$

を考える. ただし, $\mathbf{x} = (x_1, x_2, \dots, x_n)$ である.

検定関数 ψ のサイズは

$$\begin{aligned} \alpha &= \sup_{\theta \leq 17} \mathbf{R}(\theta, \psi) = \sup_{\theta \leq 17} \mathbb{E}_\theta[\mathbb{1}_W(\mathbf{x})] = \sup_{\theta \leq 17} \mathbb{P}_\theta \left[\bar{X}_n > 17 + \frac{5}{\sqrt{n}} \right] \\ &= \sup_{\theta \leq 17} \mathbb{P}_\theta \left[\frac{\bar{X}_n - \theta}{5/\sqrt{n}} > \frac{17 + \frac{5}{\sqrt{n}} - \theta}{5/\sqrt{n}} \right] = \sup_{\theta \leq 17} \mathbb{P}_\theta \left[Z > \frac{17 + \frac{5}{\sqrt{n}} - \theta}{5/\sqrt{n}} \right] \\ &= \mathbb{P}_{17} \left[Z > \frac{17 + \frac{5}{\sqrt{n}} - 17}{5/\sqrt{n}} \right] = 1 - \Phi(1) \approx 0.159 \end{aligned}$$

となる¹. ただし, Z は標準正規分布に従う確率変数とし,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$$

である.

一方, 検定関数 ψ の検出力は

$$\begin{aligned} \beta_\psi(\theta) &= \mathbb{P}_\theta \left[\bar{X}_n > 17 + \frac{5}{\sqrt{n}} \right] = \mathbb{P}_\theta \left[\frac{\bar{X}_n - \theta}{5/\sqrt{n}} > \frac{17 + \frac{5}{\sqrt{n}} - \theta}{5/\sqrt{n}} \right] \\ &= \left\{ 1 - \Phi \left(\frac{17 + \frac{5}{\sqrt{n}} - \theta}{5/\sqrt{n}} \right) \right\} \end{aligned}$$

となる.

例 7.5 X_1, X_2, \dots, X_n をベルヌーイ試行

$$f(x|\theta) = \theta^x (1-\theta)^{1-x} \mathbb{1}_{\{0,1\}}(x), \quad 0 < \theta < 1$$

¹ $\Phi(x)$ は狭義増加関数であることに注意.

からのランダム標本とし，検定問題 $H_0 : \theta \leq 1/2$, vs. $H_1 : \theta > 1/2$ を考える．標本空間 \mathcal{X} を領域 A, B, C に分割する．

$$A = \{(x_1, x_2, \dots, x_{10}) : \sum_{i=1}^{10} x_i < 5\}$$

$$B = \{(x_1, x_2, \dots, x_{10}) : \sum_{i=1}^{10} x_i = 5\}$$

$$A = \{(x_1, x_2, \dots, x_{10}) : \sum_{i=1}^{10} x_i > 5\}$$

さらに，検定関数を

$$\psi(x_1, x_2, \dots, x_{10}) = \begin{cases} 1, & ((x_1, x_2, \dots, x_{10}) \in C) \\ 1/2, & ((x_1, x_2, \dots, x_{10}) \in B) \\ 0, & ((x_1, x_2, \dots, x_{10}) \in A) \end{cases}$$

とする． $\sum_{i=1}^{10} X_i$ は二項分布 $Bin(10, \theta)$ に従うことに注意すれば，検定関数 ψ のサイズは

$$\begin{aligned} \alpha &= \sup_{\theta \leq 1/2} \mathbb{E}_\theta[\psi(X_1, X_2, \dots, X_n)] \\ &= \sup_{\theta \leq 1/2} \left[\sum_{i=6}^{10} \binom{10}{i} \theta^i (1-\theta)^{10-i} + \frac{1}{2} \binom{10}{5} \theta^5 (1-\theta)^5 \right] \\ &= \sum_{i=6}^{10} \binom{10}{i} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{10-i} + \frac{1}{2} \binom{10}{5} \left(\frac{1}{2}\right)^{10} \end{aligned}$$

となる．また，検出力は $\theta > 1/2$ に対し

$$\begin{aligned} \beta_\psi(\theta) &= \mathbb{E}_\theta[\psi(X_1, X_2, \dots, X_{10})] \\ &= \sum_{i=6}^{10} \binom{10}{i} \theta^i (1-\theta)^{10-i} + \frac{1}{2} \binom{10}{5} \theta^5 (1-\theta)^5 \end{aligned}$$

となる．

2 最強力検定とネイマン・ピアソンの補題

X_1, X_2, \dots, X_n を確率密度関数または確率関数 $f(x|\theta)$ からの大きさ n のランダム標本とする．ただし, $\Theta = \{\theta_0, \theta_1\}$ とする．いま, 検定問題

$$\text{帰無仮説 } H_0 : \theta = \theta_0, \text{ v.s. 対立仮説 } H_1 : \theta = \theta_1$$

を考える．望ましい検定関数 ψ の $\beta_\psi(\theta_0) = \mathbb{P}\{H_0 \text{を棄却} | H_0 \text{は真}\}$ は小さい値 (理想的なは 0) で, $\beta_\psi(\theta_1) = \mathbb{P}\{H_0 \text{を棄却} | H_1 \text{は真}\}$ は大きな値 (理想的なは 1) である．また, $\beta_\psi(\theta_0)$ は第 1 種の誤りであり, $1 - \beta_\psi(\theta_1)$ は第 2 種の誤りであった．しかし, このふたつの誤りの確率を同時に小さくすることはできない．

定義 7.1 帰無仮説 $H_0 : \theta = \theta_0$ v.s. 対立仮説 $H_1 : \theta = \theta_1$ の検定する検定関数 ψ が有意水準 α ($0 < \alpha < 1$) の最強力検定であるとはつぎをみたすことである．

- (1) $\beta_\psi(\theta_0) = \alpha$
- (2) どんな有意水準が α の検定関数 $\tilde{\psi}$ (すなわち, $\beta_{\tilde{\psi}}(\theta_0) \leq \alpha$) に対しても

$$\beta_\psi(\theta_1) \geq \beta_{\tilde{\psi}}(\theta_1)$$

定理 7.1 (Neyman-Pearson の基本定理): \mathbb{P}_0 と \mathbb{P}_1 を確率分布とし, それぞれはある測度 μ^2 に関する確率密度関数 p_0 と p_1 を持つこととする．また, \mathbb{P}_0 と \mathbb{P}_1 に対応する期待値を \mathbb{E}_0 と \mathbb{E}_1 記す．

(i) 存在: 検定問題 $H_0 : p_0, H_1 : p_1$ に対して, ある検定関数 ψ と定数 k と γ が存在し,

$$\mathbb{E}_0\psi(X) = \alpha \tag{7.1}$$

と

$$\psi(x) = \begin{cases} 1, & (p_1(x) > kp_0(x) \text{ のとき}) \\ \gamma & (p_1(x) = kp_0(x) \text{ のとき}) \\ 0, & (p_1(x) < kp_0(x) \text{ のとき}) \end{cases} \tag{7.2}$$

²一般には, μ を $P_0 + P_1$ とすればよい．

を満足する .

(ii) 最強力検定のための十分条件 : ある検定関数がある定数 k に対して , (7.1) と (7.2) を満足するならば , その検定関数は最強力検定になる .

(iii) 最強力検定のための必要条件 : もし , ψ が検定問題 $H_0 : p_0, H_1 : p_1$ の有意水準 α の最強力検定ならば , ψ はある k に対して (7.2) を満足する . また , 検定のサイズが α より小さく検出力が 1 なる検定が存在しなければ , ψ は (7.1) も満足する .

証明 : (i) 実数 c に対し , $g(c) = \mathbb{P}_0\{p_1(X) > cp_0(X)\}$ とおく . 確率は \mathbb{P}_0 のもとで計算されるので , 上の式の不等式は $\{x : p_0(x) > 0\}$ 上で定義されればよい . すると , $1 - g(c)$ は確率変数 $p_1(X)/p_0(X)$ の累積分布関数となる³ . さらに , $g(c)$ は非増加 , 右連続で

$$g(c-0) - g(c) = \mathbb{P}_0\left\{\frac{p_1(X)}{p_0(X)} = c\right\}, \quad g(-\infty) = 1, \quad g(+\infty) = 0$$

となる . 与えられた α に対して k を

$$g(k-0) \leq \alpha \leq g(k)$$

を満足するようにとり⁴ , 検定関数 ψ を

$$\psi(x) = \begin{cases} 1, & (p_1(x) > kp_0(x) \text{ のとき}) \\ \frac{\alpha - g(k)}{g(k-0) - g(k)}, & (p_1(x) = kp_0(x) \text{ のとき}) \\ 0, & (p_1(x) < kp_0(x) \text{ のとき}) \end{cases}$$

で定義する . 上の式の 2 番目の項は $g(k-0) - g(k) = 0$ の場合以外は定義される . また , $g(k-0) - \alpha(k) = 0$ の場合は $\mathbb{P}_0\{p_1(X) = kp_0(X)\} = 0$ となるので , ψ はほとんどいたるところで定義される . さらに , ψ のサイズは

$$\mathbb{E}_0\psi(X) = \mathbb{P}_0\left\{\frac{p_1(X)}{p_0(X)} > k\right\} + \frac{\alpha - g(k)}{g(k-0) - \alpha(k)} \mathbb{P}_0\left\{\frac{p_1(X)}{p_0(X)} = k\right\} = \alpha$$

となるように k をとればよい⁵ .

³ $\mathbb{P}_0\{p_0(X) = 0\} = 0$ より .

⁴ 与えられた α に対し唯一の k が対応する .

⁵ これは g の非増加性と $g(-\infty) = 1, g(+\infty) = 0$ から保障される . もし , α を挟んで g がジャンプしていれば , 上の式の右辺の 2 項目がジャンプする確率の内分点になっていることに注意 .

(ii) $\mathbb{E}_0\psi^*(X) \leq \alpha$ なるどんな検定関数 ψ^* にたいしても

$$\mathbb{E}_1\psi(X) \geq \mathbb{E}_1\psi^*(X)$$

が成立することを示せばよい。いま,

$$S^+ = \{x : \psi(x) - \psi^*(x) > 0\}, \quad S^- = \{x : \psi(x) - \psi^*(x) < 0\}$$

とおく。 $x \in S^+$ ならば⁶, $p_1(x) \geq kp_0(x)$ となる。同様に, $x \in S^-$ ならば⁷, $p_1(x) < kp_0(x)$ となる。これらから $S^* \cup S^-$ 上で

$$(\psi(x) - \psi^*(x))(p_1(x) - kp_0(x)) \geq 0$$

となる。したがって

$$\begin{aligned} & \int (\psi(x) - \psi^*(x))(p_1(x) - kp_0(x)) d\mu \\ &= \int_{S^* \cup S^-} (\psi(x) - \psi^*(x))(p_1(x) - kp_0(x)) d\mu \geq 0 \end{aligned}$$

となる。これから

$$\begin{aligned} & \int (\psi(x) - \psi^*(x))p_1(x) d\mu \\ & \geq k \int (\psi(x) - \psi^*(x))p_0(x) d\mu = k\{\mathbb{E}_0[\psi(X) - \psi^*(X)]\} \geq 0 \end{aligned}$$

なる。 ψ と ψ^* はともに最強力検定なので $\mathbb{E}_0\psi(X) = \alpha$, $\mathbb{E}_0\psi^*(X) = \alpha$ となることより最後の不等号はわかる。よって, (ii) は証明された。

(iii) ψ^* をサイズが α 以下の最強力検定とする。また, ψ を (7.1) と (7.2) を満足する検定とする。

$$S = (S^+ \cup S^-) \cap \{x : p_1(x) \neq kp_0(x)\}$$

⁶ $\psi(x) > \psi^*(x)$ とならば, $\psi(x) \neq 0$ となるので,

$$\psi(x) \neq 0 \implies p_1(x) \geq kp_0(x)$$

となる。

⁷ $\psi(x) < \psi^*(x)$ とならば, $\psi(x) \neq 1$ となり,

$$\psi(x) \neq 1 \implies p_1(x) \leq kp_0(x)$$

となる。

とする．すなわち， $S^+ \cup S^-$ 上で ψ と ψ^* の値が異なる点である．さらに， $\mu(S) > 0$ と仮定する．すると

$$\int_{S^+ \cup S^-} (\psi - \psi^*)(p_1 - kp_0) d\mu = \int_S (\psi - \psi^*)(p_1 - kp_0) d\mu > 0$$

となる．これより

$$\int_{S^+ \cup S^-} (\psi - \psi^*)p_1 d\mu > k \int_{S^+ \cup S^-} (\psi - \psi^*)p_0 d\mu = k(\mathbb{E}_0\psi(X) - \mathbb{E}_0\psi^*(X)) \geq 0$$

となる．最後の等号は $(S^+ \cup S^-)^c$ 上では $\psi(x) = \psi^*(x)$ からわかる．また，最後の不等号は $\mathbb{E}_0\psi(X) = \alpha$ ， $\mathbb{E}_0\psi^*(X) \leq \alpha$ からわかる．これより $\mathbb{E}_1\psi(X) - \mathbb{E}_1\psi^*(X) > 0$ となり， ψ が最強力検定であることに矛盾する．したがって， $\mu(S) = 0$ となる． \square

例 7.6 X_1, X_2, \dots, X_n を確率密度関数 $f(x|\theta) = \theta e^{-\theta x} \mathbb{1}_{(0, \infty)}(x)$ からの大きさ n のランダム標本とする．ただし， $\theta \in \Theta = \{\theta_0, \theta_1\}$ で $0 < \theta_i$ ($i = 0, 1$) かつ $\theta_0 < \theta_1$ である．検定問題

$$H_0 : \theta = \theta_0, \quad H_1 : \theta = \theta_1$$

の検定を考える．いま

$$\begin{aligned} L_0(x_1, x_2, \dots, x_n) &= L_0 = \prod_{i=1}^n f(x_i|\theta_0) \\ &= \theta_0^n \exp\{-\theta_0 \sum_{i=1}^n x_i\} \prod_{i=1}^n \mathbb{1}_{(0, \infty)}(x_i) \\ L_1(x_1, x_2, \dots, x_n) &= L_1 = \prod_{i=1}^n f(x_i|\theta_1) \\ &= \theta_1^n \exp\{-\theta_1 \sum_{i=1}^n x_i\} \prod_{i=1}^n \mathbb{1}_{(0, \infty)}(x_i) \end{aligned}$$

とおく．Neyman-Pearson の基本定理より $x_i > 0$ ($i = 1, 2, \dots, n$) に対し

$$\lambda = \frac{L_0}{L_1} < k^* \iff \left(\frac{\theta_0}{\theta_1}\right)^n \exp\{-(\theta_0 - \theta_1) \sum_{i=1}^n x_i\} < k^*$$

ならば H_0 を棄却すれば最強力検定を得られる．これは

$$\sum_{i=1}^n x_i < \frac{1}{\theta_1 - \theta_0} \log \left[\left(\frac{\theta_1}{\theta_0}\right)^n k^* \right] = k'$$

ならば H_0 を棄却することと同値である．つぎに有意水準 α に対して k' を定める．すなわち

$$\alpha = \mathbb{P}[H_0 \text{ 棄却} \mid H_0 \text{ が真}] = \mathbb{P}_{\theta_0} \left[\sum_{i=1}^n X_i < k' \right]$$

θ_0 のもとで $\sum_{i=1}^n X_i$ は母数 n と θ_0 のガンマ分布に従うので

$$\alpha = \mathbb{P}_{\theta_0} \left[\sum_{i=1}^n X_i \leq k' \right] = \int_0^{k'} \frac{1}{\Gamma(n)} \theta_0^n x^{n-1} e^{-\theta_0 x} dx \quad (7.3)$$

を満足するように k' を定められたよい．(7.3) の解 k' を用いて検定関数

$$\psi(x_1, x_2, \dots, x_n) = \begin{cases} 1, & (\sum_{i=1}^n x_i \leq k') \\ 0, & (\sum_{i=1}^n x_i > k') \end{cases}$$

とし，

$$d(x_1, x_2, \dots, x_n) = \begin{cases} 1, & (\text{確率 } \psi(x_1, x_2, \dots, x_n)) \\ 0, & (\text{確率 } 1 - \psi(x_1, x_2, \dots, x_n)) \end{cases}$$

とすればよい．

例 7.7 X_1, X_2, \dots, X_n を確率関数 $f(x|\theta) = \theta^x(1-\theta)^{1-x}\mathbb{1}_{\{0,1\}}(x)$ からの大きさ n のランダム標本とする．ただし， $\theta \in \Theta = \{\theta_0, \theta_1\}$ で $0 < \theta_i < 1$ ($i = 1, 2$) かつ $\theta_0 < \theta_1$ である．検定問題

$$H_0: \theta = \theta_0 \quad \text{vs.} \quad H_1: \theta = \theta_1$$

の検定を考える．いま

$$L_0(x_1, x_2, \dots, x_n) = L_0 = \theta_0^{\sum_{i=1}^n x_i} (1 - \theta_0)^{n - \sum_{i=1}^n x_i} \prod_{i=1}^n \mathbb{1}_{\{0,1\}}(x_i)$$

$$L_1(x_1, x_2, \dots, x_n) = L_1 = \theta_1^{\sum_{i=1}^n x_i} (1 - \theta_1)^{n - \sum_{i=1}^n x_i} \prod_{i=1}^n \mathbb{1}_{\{0,1\}}(x_i)$$

とおく．したがって，

$$\lambda \leq k^* \iff \theta_0^{\sum_{i=1}^n x_i} (1 - \theta_0)^{n - \sum_{i=1}^n x_i} / \theta_1^{\sum_{i=1}^n x_i} (1 - \theta_1)^{n - \sum_{i=1}^n x_i} \leq k^*$$

$$\iff \left[\frac{\theta_0(1 - \theta_1)}{\theta_1(1 - \theta_0)} \right]^{\sum_{i=1}^n x_i} \left(\frac{1 - \theta_0}{1 - \theta_1} \right)^n \leq k^*$$

$$\iff \sum_{i=1}^n x_i \geq k'$$

となる⁸ . よって

$$\sum_{i=1}^n x_i \geq k' \quad \text{ならば, } H_0 \text{ を棄却}$$

とすれば, 最強力検定が得られる .

ここで, $\theta_0 = 1/4, \theta_1 = 3/4, n = 10$ の場合について $\alpha = 0.05$ としたときに k' を求めよう .

$$\alpha = \mathbb{P}_{\theta=1/4}[H_0 \text{ を棄却}] = \mathbb{P}_{\theta=1/4}\left[\sum_{i=1}^{10} X_i \geq k'\right] = \sum_{i=k'}^{10} \binom{10}{i} \left(\frac{1}{4}\right)^i \left(\frac{3}{4}\right)^{n-i}$$

となる . $k' = 6$ ならば $\alpha = 0.0197$ となり, $k' = 5$ ならば $\alpha = 0.0781$ となるので,

$$\psi(x_1, x_2, \dots, x_{10}) = \begin{cases} 1, & (\sum_{i=1}^{10} x_i \geq 6) \\ \frac{0.05-0.00197}{0.0584}, & (\sum_{i=1}^{10} x_i = 5) \\ 0, & (\sum_{i=1}^{10} x_i \leq 4) \end{cases}$$

とし⁹,

$$d(x_1, x_2, \dots, x_{10}) = \begin{cases} 1, & (\text{確率 } \psi(\mathbf{x})) \\ 0, & (\text{確率 } 1 - \psi(\mathbf{x})) \end{cases}$$

とすれば, 有意水準 0.05 の最強力検定が得られる .

3 一様最強力検定

4 演習問題

問題 7.1 X を確率密度関数 $f(x|\theta), \theta \in \{0, 1\}$, からの大きさ 1 のランダム標本とする . ただし, $f(x|\theta)$ は

x	1	2	3	4	合計
$f(x 0)$	0.01	0.01	0.01	0.97	1
$f(x 1)$	0.06	0.05	0.04	0.85	1

⁸ $\theta_0 < \theta_1$ なので,

$$\frac{\theta_0}{1 - \theta_0} < \frac{\theta_1}{1 - \theta_1}$$

に注意する .

⁹ $\mathbb{P}(\sum_{i=1}^5 X_i = 5) = \mathbb{P}(\sum_{i=1}^5 X_i \leq 6) - \mathbb{P}(\sum_{i=1}^5 X_i \leq 4) = 0.781 - 0.0197 = 0.0584$ からわかる .

である．このとき，検定問題 $H_0 : \theta = 0$, v.s. $H_1 : \theta = 1$ を考える．

- (1) 有意水準 0.02 の非確率化検定関数をすべて求めよ．
- (2) 上の問題で求めた検定関数の第二種の誤りの確率を求め，そのうちで最も望ましい検定関数を指定せよ．

問題 7.2 X_1, X_2 をポアソン分布

$$f_X(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2, \dots, \quad \lambda > 0$$

からの大きさ 2 のランダム標本とする．ただし， $\lambda > 0$ である．このとき，検定問題 $H_0 : \lambda \leq 1$, v.s. $H_1 : \lambda > 1$ を考える．

- (1) 検定関数

$$\psi_1(x_1, x_2) = \begin{cases} 1, & (x_1 \geq 2) \\ 0, & (x_1 < 2) \end{cases}$$

および

$$\psi_2(x_1, x_2) = \begin{cases} 1, & (x_1 + x_2 > 3) \\ 0.65, & (x_1 + x_2 = 3) \\ 0, & (x_1 + x_2 < 3) \end{cases}$$

の検出力関数を求めよ．

- (2) 検定関数 $\psi_1(x_1, x_2)$ と $\psi_2(x_1, x_2)$ のサイズを求めよ．

問題 7.3 表の出る確率が θ ($0 < \theta < 1$) なるコインを投げる実験を考える．帰無仮説 $H_0 : \theta = 1/2$ ，対立仮説 $H_1 : \theta = 3/4$ の検定を考える．そのためにコインを繰り返し投げたときに何回に初めて表がでるかを観測する実験を二人の人が行なうことにする． X_1 と X_2 をそれぞれの人が行う実験の結果に対応させる．したがって， X_1 と X_2 は独立に

$$P(X_1 = n) = \theta(1 - \theta)^{n-1}, \quad n = 1, 2, \dots$$

なる確率分布に従う．

- (1) 棄却域

$$G = \{(n_1, n_2) \in \mathbb{N} \times \mathbb{N} : 2 \leq n_1 + n_2 \leq 3\}$$

であたえられた検定のサイズをもとめよ．

- (2) (1) において棄却域があたえられた検定の検出力を求めよ．

第8章 信頼区間

母集団分布から無作為標本に基づき，母集団分布の未知の母数を区間で推定する方法とその推定精度の評価法について学ぶ．

1 信頼区間の定義

定義 8.1 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ を実数値母数 $\theta (\theta \in \Theta)$ に依存する同時分布に従う確率変数ベクトルとし， $L(\mathbf{X}) < U(\mathbf{X})$ を統計量とする．ただし， Θ を母数空間とする．このとき，区間 $[L(\mathbf{X}), U(\mathbf{X})]$ が θ に対する信頼係数 $100p\%$ ($0 < p < 1$) の信頼区間であるとは，すべての $\theta \in \Theta$ に対し

$$\mathbb{P}_\theta[L(\mathbf{X}) \leq \theta \leq U(\mathbf{X})] \geq p$$

が成立することである．ただし，すくなくともひとつの θ に対し等号が成立しななければならない．

p は信頼水準，信頼限界または被覆確率という．

定義 8.2 すべての $\theta \in \Theta$ に対し

$$\mathbb{P}_\theta\{\theta \geq L(\mathbf{X})\} \geq p$$

となる統計量 $L(\mathbf{X})$ を θ に対する信頼係数 $100p\%$ 信頼下限といい，

$$\mathbb{P}_\theta\{U(\mathbf{X}) \geq \theta\} \geq p$$

となる統計量 $U(\mathbf{X})$ を θ に対する信頼係数 $100p\%$ 信頼上限という．

$L(\mathbf{X})$ を θ に対する $100p_1\%$ 信頼下限とし， $U(\mathbf{X})$ を θ に対する $100p_2\%$ 信頼上限で $L(\mathbf{X}) < U(\mathbf{X})$ したとき，区間 $[L(\mathbf{X}), U(\mathbf{X})]$ は θ に対する信頼係数 $100p\%$ の信頼区間となる．ただし， $p = p_1 + p_2 - 1$ である．

例 8.1 X_1, X_2, \dots, X_n を独立同一に正規分布 $N(\mu, 1)$ に従う確率変数とする．このとき， $\sqrt{n}(\bar{X}_n - \mu)$ は標準正規分布 $N(0, 1)$ に従い，

$$\mathbb{P}_\mu\{-1.96 \leq \sqrt{n}(\bar{X}_n - \mu) \leq 1.96\} = 0.95$$

となる．ただし， $\bar{X}_n = (1/n) \sum_{i=1}^n X_i$ である．事象 $\{-1.96 \leq \sqrt{n}(\bar{X}_n - \mu) \leq 1.96\}$ は $\{\bar{X}_n - 1.96/\sqrt{n} \leq \mu \leq \bar{X}_n + 1.96/\sqrt{n}\}$ に等しいので，

$$\mathbb{P}_\mu \left[\bar{X}_n - \frac{1.96}{\sqrt{n}} \leq \mu \leq \bar{X}_n + \frac{1.96}{\sqrt{n}} \right] = 0.95$$

となる．したがって， $[\bar{X}_n - 1.96/\sqrt{n}, \bar{X}_n + 1.96/\sqrt{n}]$ は μ に対する信頼係数 95% の信頼区間となる．

X_1, X_2, \dots, X_n を独立同一に平均 μ ，分散 1 の分布に従う確率変数とする．このとき，中心極限定理から $\sqrt{n}(\bar{X}_n - \mu)$ は標準正規分布 $N(0, 1)$ に漸的に従うことがわかるので， n が十分大きいとき

$$\mathbb{P}_\mu\{-1.96 \leq \sqrt{n}(\bar{X}_n - \mu) \leq 1.96\} \approx 0.95$$

となる． $[\bar{X}_n - 1.96/\sqrt{n}, \bar{X}_n + 1.96/\sqrt{n}]$ は μ に対する信頼係数 95% の近似信頼区間となる．

2 ピボット法

$\mathbf{X} = (X_1, X_2, \dots, X_n)$ を実数値母数 $\theta (\theta \in \Theta)$ に依存する同時分布に従う確率変数ベクトルとし， $g(\mathbf{X}, \theta)$ を確率変数とし，その分布は θ には依存しないとす．すなわち，

$$\mathbb{P}_\theta\{g(\mathbf{X}, \theta) \leq x\} = G(x), \quad \forall x \in \mathbb{R}$$

は分布関数で θ に依存しないものである．以後では，簡単なために $G(x)$ は $\{x \in \mathbb{R} : G(x) > 0\}$ 上で狭義単調増加の分布関数と仮定する．したがって，うまく実数 a, b をとり

$$\mathbb{P}_\theta\{a \leq g(\mathbf{X}, \theta) \leq b\} = p$$

とできる．さらに，事象 $\{a \leq g(\mathbf{X}, \theta) \leq b\}$ を変形して，

$$\mathbb{P}_\theta[L(\mathbf{X}) \leq \theta \leq U(\mathbf{X})] = p$$

なる $L(X)$, $U(X)$ を見つけることができることが期待できる．このように，信頼区間を構成するにあたり利用する確率変数 $g(X, \theta)$ を θ に対するピボットという．要点は $g(X, \theta)$ の分布が未知母数 θ に依存しないことである．

例 8.2 X_1, X_2, \dots, X_n ($n \geq 2$) は独立同一に $[0, \theta]$ 上の一様分布に従うとする．ただし， $\theta > 0$ である．このとき， θ の最尤推定量は X_1, X_2, \dots, X_n ($n \geq 2$) の最大値 $X_{(n)}$ である．また， $X_{(n)}/\theta$ の分布関数は

$$G(x) \equiv \mathbb{P}_\theta\{X_{(n)}/\theta \leq x\} = x^n \mathbb{I}_{[0, 1]}(x) + \mathbb{1}_{(1, \infty)}(x)$$

である．したがって， $X_{(n)}/\theta$ は θ に対するピボットとなる．信頼係数 100p% の信頼区間を見つめるためには

$$\mathbb{P}_\theta \left[a \leq \frac{X_{(n)}}{\theta} \leq b \right] = p$$

なる a と b を見つければよい． $X_{(n)}/\theta$ の確率密度関数は単調増加なので， $a = (1-p)^{1/n}$ と $b = 1$ とすれば，区間 $[a, b]$ は最短¹になる．さらに，信頼係数が 100p% の信頼区間 $[X_{(n)}/a, X_{(n)}/b]$ の中で $[X_{(n)}, X_{(n)}/(1-p)^n]$ が最短となること²もわかる．

例 8.3 X_1, X_2, \dots, X_{10} が独立同一に母数 θ の指数分布

$$f_X(x) = \theta e^{-\theta x} \mathbb{1}_{(0, \infty)}(x)$$

¹ $[(1-p)^{1/n}, 1]$ が最短区間になることは以下のことからわかる． $X_{(n)}/\theta$ の確率密度関数を $f(x)$ とする． $b < 1$ と仮定すれば，

$$\int_a^b f(x) dx = \int_a^{a+1-b} f(x) dx + \int_{a+1-b}^b f(x) dx - \int_b^1 f(x) dx + \int_b^1 f(x) dx \leq \int_{a+1-b}^b f(x) dx$$

となる．なぜならば， $f(x)$ の単調性から

$$\int_a^{a+1-b} f(x) dx - \int_b^1 f(x) dx \leq 0$$

からわかる．あとは，

$$\mathbb{P}_\theta \left[a \leq \frac{X_{(n)}}{\theta} \leq 1 \right] = G(1) - G(a) = 1 - a^n = p$$

に注意すればよい．

² $ab \leq 1$ から $1/b - 1/a = (b-a)/(ab) > 1 - (1-p)^p$ となることよりわかる．

に従うとする．ただし， $\theta > 0$ である．このとき， $2\theta X_1$ は自由度 2 のカイ 2 乗分布³ に従うので， $2\theta \sum_{i=1}^{10} X_i$ は自由度 20 のカイ 2 乗分布に従う．したがって，たとえば，信頼係数 90% の信頼区間は

$$\mathbb{P}_\theta \left[a \leq 2\theta \sum_{i=1}^{10} X_i \leq b \right] = 0.90$$

なる a, b を見つけることになる．これは

$$\mathbb{P}_\theta \left[2\theta \sum_{i=1}^{10} X_i \leq a \right] = 0.05, \quad \mathbb{P}_\theta \left[2\theta \sum_{i=1}^{10} X_i \leq b \right] = 0.95$$

から $a = 10.85, b = 31.41$ となる．したがって，

$$\left[\frac{5.425}{\sum_{i=1}^{10} X_i}, \frac{15.705}{\sum_{i=1}^{10} X_i} \right]$$

となる．しかし，カイ 2 乗分布の確率密度関数はモードに関して非対称なので，これは最短の区間ではない．実際，

$$\left[\frac{4.893}{\sum_{i=1}^{10} X_i}, \frac{14.938}{\sum_{i=1}^{10} X_i} \right]$$

が最短となることが知られている．

未知母数を $\theta = (\theta_1, \theta_2, \dots, \theta_p)$ とし， θ_1 に対する信頼区間を構成することを考える． $g(\mathbf{X}, \theta_1)$ は θ_1 のみに依存する確率変数で $\theta_2, \theta_3, \dots, \theta_p$ には依存しないとする．さらに， $g(\mathbf{X}, \theta_1)$ の分布は θ には依存しないとする．すなわち

$$\mathbb{P}_\theta \{g(\mathbf{X}, \theta_1) \leq x\} = G(x)$$

とする．このような場合，ピボット $g(\mathbf{X}, \theta_1)$ を用いて θ_1 の信頼区間を構成することができる．

³自由度 k のカイ 2 乗分布の確率密度関数は

$$\frac{1}{\Gamma(k/2)} \left(\frac{1}{2}\right)^{k/2} x^{(k/2)-1} e^{-(1/2)x} \mathbb{I}_{(0, \infty)}(x)$$

であるので，自由度 2 のカイ 2 乗分布の密度関数は $(1/2)e^{-(1/2)x}$ となることに注意．

例 8.4 X_1, X_2, \dots, X_n を独立同一に正規分布 $N(\mu, \sigma^2)$ に従う確率変数とする．ここで, μ, σ^2 は未知とする． μ に対する信頼区間を求めよう．そのために,

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{S}$$

とおく．ただし

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i, \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

である． $\sqrt{n}(\bar{X}_n - \mu)/S$ は自由度 $(n-1)$ の t 分布⁴に従いので, この分布は μ, σ^2 に依存しないので, μ のピボットとなる．

また, σ^2 の信頼区間を構成するために

$$\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^n \left(\frac{X_i - \bar{X}_n}{\sigma} \right)^2 \sim \chi^2(n-1)$$

となるので, σ^2 のピボットとなる．

多くの場合には, ピボットの分布を厳密に求めることが難しい．しかし, ピボットの分布の近似(漸近分布)を求めることができる．すなわち, ピボット $g(\mathbf{X}, \theta)$ に対し

$$\mathbb{P}_\theta\{g(\mathbf{X}, \theta) \leq x\} \approx G(x)$$

なる θ に依存しない分布関数 $G(x)$ を求めることができることがおおい．以下は近似がうまくいくために, 標本数 n は大きいと仮定し, 大きさ n のランダム標本 (i.i.d. 標本) が得られる場合を考える． $\hat{\theta}_n$ は n 個の標本に基づく θ の推定量とし, その漸近分布が平均 θ , 分散 $\sigma_n^2(\theta)$ の正規分布と仮定する．すなわち,

$$\frac{\hat{\theta}_n - \theta}{\sigma_n(\theta)} \xrightarrow{L} N(0, 1)$$

⁴自由度 k の t 分布の確率密度関数は

$$f(x) = \frac{\Gamma((k+1)/2)}{\Gamma(k/2)\sqrt{\pi k} \left(1 + \frac{x^2}{n}\right)^{(k+1)/2}}$$

で与えられる．

とする⁵．さらに， $\sigma(\theta)$ に $\hat{\theta}_n$ を代入したもの

$$\frac{\hat{\theta}_n - \theta}{\sigma_n(\hat{\theta}_n)}$$

も漸近的に標準正規分布に従うことが多くの場合保証されるので，これをピボットとして用いればよい．これらのことより

$$\mathbb{P}_\theta \left\{ -z_{p/2} \leq \frac{\hat{\theta}_n - \theta}{\sigma_n(\hat{\theta}_n)} \leq z_{p/2} \right\} \approx 1 - p$$

を得る．ただし， $z_{p/2}$ は $1 - \Phi(z_{p/2}) = p/2$ をみたす点である．したがって， θ に対する信頼係数 $100(1-p)\%$ の近似信頼区間は $[\hat{\theta}_n - z_{p/2}\sigma_n(\hat{\theta}_n), \hat{\theta}_n + z_{p/2}\sigma_n(\hat{\theta}_n)]$ で与えられる．

例 8.5 X を二項分布 $Bin(n, \theta)$ に従う確率変数とする．ただし， θ ($0 < \theta < 1$) は未知とする．このとき， θ の最尤推定量は $\hat{\theta}_n = X/n$ である． n が十分大きければ， $\sqrt{n}(\hat{\theta}_n - \theta)/\sqrt{\theta(1-\theta)}$ は近似的に標準正規分布に従うことが中心極限定理よりわかる．したがって， θ に対する信頼係数 95% の近似信頼区間は

$$\mathbb{P}_\theta \left[-1.96 \leq \frac{\sqrt{n}(\hat{\theta}_n - \theta)}{\sqrt{\theta(1-\theta)}} \leq 1.96 \right] = \mathbb{P}_\theta \left[\frac{n(\hat{\theta}_n - \theta)^2}{\theta(1-\theta)} \leq (1.96)^2 \right] \approx 0.95$$

より構成できる．したがって， θ に対する信頼係数 95% の近似信頼区間は

$$\{g(t) := n(\hat{\theta}_n - t)^2 - (1.96)^2 t(1-t) \leq 0\}$$

となる． $g(t) = 0$ の解は

$$t = \frac{\hat{\theta}_n + (1.96)^2/(2n) \pm 1.96\{\hat{\theta}_n(1-\hat{\theta}_n)/n + (1.96)^2/(4n^2)\}^{1/2}}{1 + (1.96)^2/n}$$

となる．したがって，信頼区間はこの 2 点を両端にもつ区間となる．または，ピボットとして

$$\frac{\sqrt{n}(\hat{\theta}_n - \theta)}{\sqrt{\hat{\theta}_n(1-\hat{\theta}_n)}}$$

⁵これはすべての $x \in \mathbb{R}$ に対し， $n \rightarrow \infty$ のとき

$$\mathbb{P}_\theta \left\{ (\hat{\theta}_n - \theta)/\sigma_n(\theta) \right\} \rightarrow \Phi(x)$$

となることである．ただし， $\Phi(\cdot)$ は標準正規分布の累積分布関数である．

を利用すれば, θ に対する信頼係数 95% の近似信頼区間は

$$\left[\hat{\theta}_n - 1.96 \sqrt{\frac{\hat{\theta}_n(1 - \hat{\theta}_n)}{n}}, \hat{\theta}_n + 1.96 \sqrt{\frac{\hat{\theta}_n(1 - \hat{\theta}_n)}{n}} \right]$$

となる. n が大きいときには, このふたつの信頼区間は違いは小さい.

3 演習問題

問題 8.1 $X_1, X_2, \dots, X_n (n \geq 2)$ を確率密度関数

$$f_X(x|\mu) = e^{-(x-\mu)} \mathbb{1}_{[\mu, \infty)}(x) \quad \mu > 0$$

からの大きさ n のランダム標本とする. ただし, $I_A(x)$ は指示関数とする. すなわち, 集合 A に対し, $x \in A$ ならば, $I_A(x) = 1$, $x \notin A$ ならば, $\mathbb{1}_A(x) = 0$ である.

- (1) X_1 の累積分布関数を求めよ.
- (2) $Y = \min\{X_1, X_2, \dots, X_n\}$ としたとき, Y の確率密度関数は

$$f_Y(y|\mu) = ne^{-n(y-\mu)} \mathbb{1}_{[\mu, \infty)}(y)$$

で与えられることを示せ.

- (3) 信頼係数 $100p\%$ ($0 < p < 1$) の μ の信頼区間を構成せよ. ただし

$$\mathbb{P}_\mu(Y - \mu \leq a) = \frac{1-p}{2}, \quad \mathbb{P}_\mu(Y - \mu \geq b) = \frac{1-p}{2}$$

をみたす点 a, b ($a < b$) を利用する.

問題 8.2 $X_1, X_2, \dots, X_n (n \geq 2)$ を $(0, \theta)$ 上の一様分布からのランダム標本とする. ただし, $\theta > 0$ である. また, $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ を X_1, X_2, \dots, X_n の順序統計量とし, r を $1 \leq r \leq n-1$ なる任意の自然数とする. このとき, $X_{(r)}/\theta$ をピボットとして, 信頼係数 $100p\%$ ($0 < p < 1$) の θ の信頼区間を構成せよ.

ヒント 1: $Y_1, Y_2, \dots, Y_n (n \geq 2)$ は独立同一分布に従い, 連続な分布関数 $F_Y(y)$ (確率密度関数を $f_Y(y)$ とする) を持つとき, Y_1, Y_2, \dots, Y_n の r 番目の順序統計量 $Y_{(r)}$ の確率密度関数は

$$f_Y(y) = \frac{n!}{(r-1)!(n-r)!} f_Y(y) [F_Y(y)]^{r-1} [1 - F_Y(y)]^{n-r}$$

となる.

ヒント 2: 確率変数 Z はパラメータ (α, β) のベータ分布に従うとする: すなわち, Z の確率密度関数は

$$f_Z(z) = \frac{(\alpha + \beta - 1)!}{(\alpha - 1)!(\beta - 1)!} z^{\alpha-1} (1-z)^{\beta-1} \mathbb{1}_{(0,1)}(z)$$

である. ただし, $\alpha > 0, \beta > 0$ である. このとき, q ($0 < q < 1$) に対して, $u_q(\alpha, \beta)$ と $l_q(\alpha, \beta)$ を

$$\int_{u_q(\alpha, \beta)}^1 f_Z(z) dz = q, \quad \int_0^{l_q(\alpha, \beta)} f_Z(z) dz = q$$

をみたすものとする.

付録A 補遺

1 二項定理と多項定理

二項定理

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$$

ただし,

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}, \quad 0! = 1$$

である.

多項定理

$$(a_1 + a_2 + \cdots + a_r)^n = \sum_{k_1+k_2+\cdots+k_r=n} \frac{n!}{k_1!k_2!\cdots k_r!} a_1^{k_1} \cdot a_2^{k_2} \times \cdots \times a_r^{k_r}$$

2 数と数列

2.1 数の集合

- N 自然数全体の集合
- Z 整数全体の集合
- Q 有理数全体の集合
- R 実数全体の集合
- C 複素数全体の集合

注意: $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$.

2.2 指数法則

$a > 0, b > 0$ で m, n は実数のとき

$$(1) \quad a^m \times a^n = a^{m+n},$$

$$(2) \quad (a^m)^n = a^{mn},$$

$$(3) \quad (ab)^n = a^n b^n,$$

$$(4) \quad \frac{a^m}{a^n} = a^{m-n}.$$

2.3 数列

実数を無限個

$$a_1, a_2, \dots, a_n, \dots$$

のようになればたものを数列といい，簡単に $\{a_n\}_{n=1}^{\infty}$ または $\{a_n\}$ と書く．数列の項の番号 n を限りなく大きくすることを記号で

$$n \rightarrow \infty$$

と書く．

数列 $\{a_n\}$ と実数 a に関して n を限りなく大きくしていきとき a_n の値が限りなく近づくならば，数列 $\{a_n\}$ は a に収束するといい，このことを

$$\lim_{n \rightarrow \infty} a_n = a$$

と書く． a を数列 $\{a_n\}$ の極限值という．

収束列の演算

ふたつの数列 $\{a_n\}_{n=1}^{\infty}, \{b_n\}_{n=1}^{\infty}$ がともに収束するならば， $\{a_n + b_n\}_{n=1}^{\infty}$ も $\{a_n b_n\}_{n=1}^{\infty}$ も収束列であって

$$(1) \quad \lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n,$$

$$(2) \quad \lim_{n \rightarrow \infty} a_n b_n = \lim_{n \rightarrow \infty} a_n \lim_{n \rightarrow \infty} b_n.$$

さらに， $b_n \neq 0, n = 1, 2, \dots$ ，かつ $\lim_{n \rightarrow \infty} b_n \neq 0$ ならば， $\{a_n/b_n\}_{n=1}^{\infty}$ も収束列であって，

$$(3) \quad \lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n}.$$

2.4 有限級数

n 個の数 a_1, a_2, \dots, a_n の和を \sum (シグマ) の記号を用いて簡易に表現する:

$$a_1 + a_2 + \cdots + a_n = \sum_{k=1}^n a_k.$$

このとき,

$$(1) \quad \sum_{k=1}^n (a_k + b_k) = \sum_{k=1}^n a_k + \sum_{k=1}^n b_k,$$

$$(2) \quad \sum_{k=1}^n ca_k = c \sum_{k=1}^n a_k,$$

$$(3) \quad \sum_{k=1}^n c = nc.$$

ただし, c は定数である.

注意: $\sum_{k=1}^n a_k^2$ と $(\sum_{k=1}^n a_k)^2$ は異なることに注意せよ.

2.5 無限級数

数列 $\{a_n\}$ の各項を $+$ の記号で結んだ式

$$a_1 + a_2 + \cdots + a_n + \cdots$$

を級数という. また,

$$S_n = a_1 + a_2 + \cdots + a_n$$

をこの級数の部分和という. 部分和の数列 $\{S_n\}$ について

$$\lim_{n \rightarrow \infty} S_n = S \quad \text{ただし, } S \text{ は有限な確定値}$$

が成り立つとき, この級数は収束するといひ,

$$a_1 + a_2 + \cdots + a_n + \cdots = S$$

と書く. 収束しない級数は発散するといひ.

収束する級数の演算

ふたつの数列 $\sum_{n=1}^{\infty} a_n$ と $\sum_{n=1}^{\infty} b_n$ がともに収束するならば, $\sum_{n=1}^{\infty} (a_n + b_n)$ も $\sum_{n=1}^{\infty} ca_n$ も収束して

$$(1) \quad \sum_{n=1}^{\infty} (a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n,$$

$$(2) \quad \sum_{n=1}^{\infty} ca_n = c \sum_{n=1}^{\infty} a_n$$

である。ただし、 c は定数。

2.6 上極限と下極限

$\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ を補完数直線といい、順序は $-\infty < x < \infty$, $x \in \mathbb{R}$ で定める。 $\{a_n\}_{n=1}^{\infty}$ を数列とする。このとき、 $b_n = \sup_{k \geq n} a_k$ と¹おけば、 $\{b_n\}_{n=1}^{\infty}$ は単調減少列 (ただし、上限が存在しない場合は、 $b_n = \infty$ とする) となる。 $\lim_{n \rightarrow \infty} b_n = \inf_{n \geq 1} b_n$ の極限が存在する ($\overline{\mathbb{R}}$ において) とき、

$$\limsup_n a_n = \lim_{n \rightarrow \infty} b_n = \inf_{n \geq 1} \sup_{k \geq n} a_k$$

と定める。同様に、 $c_n = \inf_{k \geq n} a_k$ とおけば、 $\{c_n\}_{n=1}^{\infty}$ は単調増加列 (ただし、下限が存在しない場合は、 $c_n = -\infty$ とする) となる。 $\lim_{n \rightarrow \infty} c_n = \sup_{n \geq 1} c_n$ の極限が存在する ($\overline{\mathbb{R}}$ において) とき、

$$\liminf_n a_n = \lim_{n \rightarrow \infty} c_n = \sup_{n \geq 1} \inf_{k \geq n} a_k$$

と定める。

数列 $\{a_n\}_{n=1}^{\infty}$ の極限が存在すること ($\overline{\mathbb{R}}$ において) と

$$\limsup_n a_n = \liminf_n a_n$$

は同値である。証明は杉浦 (「解析入門 I」 p. 365 等を参照のこと)

3 微積分学の復習

3.1 関数の極限

関数の極限値の定義

$f(x)$ が定義されている区間で、 x が a と異なる値をとりながら a に近づくととき、 $f(x)$ の値が一定の値 b に限りなく近づくなれば、これを

$$\lim_{x \rightarrow a} f(x) = b$$

¹正確には、 $A_n = \{a_k : k \geq n\}$ とし、 $b_n = \sup A_n$ の意味である。

と表し、値 b を x が a に限りなく近づくときの $f(x)$ の極限值という。

右極限と左極限の定義

- x が a より大きな値をとりながら、限りなく a に近づくことを x は右から a に近づくといい、

$$x \rightarrow a + 0$$

と表す。

- x が a より小さな値をとりながら、限りなく a に近づくことを x は左から a に近づくといい、

$$x \rightarrow a - 0$$

と表す。

- $x \rightarrow a + 0$, $x \rightarrow a - 0$ のときの $f(x)$ の極限を、それぞれ x が a に近づくときの右極限、左極限といい、

$$\lim_{x \rightarrow a + 0} f(x), \quad \lim_{x \rightarrow a - 0} f(x)$$

とかく。

- $\lim_{x \rightarrow a} f(x)$ が存在するためには、 $\lim_{x \rightarrow a + 0} f(x)$, $\lim_{x \rightarrow a - 0} f(x)$ がともに存在し、それらが等しくならなければならない。

極限値の性質

$\lim_{x \rightarrow a} f(x) = \alpha$, $\lim_{x \rightarrow a} g(x) = \beta$ (α, β は定数) のとき、

$$(1) \quad \lim_{x \rightarrow a} kf(x) = k\alpha \quad (k \text{ は定数})$$

$$(2) \quad \lim_{x \rightarrow a} (f(x) \pm g(x)) = \alpha \pm \beta$$

$$(3) \quad \lim_{x \rightarrow a} f(x)g(x) = \alpha\beta, \quad \lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{\alpha}{\beta}, \quad \text{ただし, } g(x) \neq 0, \beta \neq 0$$

3.2 連続関数の定義

- 関数 $f(x)$ が $x = a$ において連続であるとは

$$\lim_{x \rightarrow a} f(x) = f(a)$$

がなりたつことをいう。

- f が区間 I の各点で連続のとき, f は区間 I 上で連続という。
- ただし, f が閉区間 $I = [a, b]$ 上で連続とは, 开区間 (a, b) 上で連続であり, 区間の端 a, b においては

$$\lim_{x \rightarrow a+0} f(x) = f(a), \quad \lim_{x \rightarrow b-0} f(x) = f(b)$$

がなりたつことをいう。

3.3 合成関数の極限値の性質

$f(x)$ は x_0 で連続とし, $g(y)$ が $y_0 = f(x_0)$ で連続ならば, 合成関数 $g(f(x))$ も x_0 で連続. すなわち

$$\lim_{x \rightarrow x_0} g(f(x)) = g(f(x_0))$$

また, $\lim_{x \rightarrow x_0} f(x) = f(x_0)$ を代入すれば,

$$\lim_{x \rightarrow x_0} g(f(x)) = g(\lim_{x \rightarrow x_0} f(x))$$

となる。

3.4 導関数の定義

- 関数 $f(x)$ と t に対して, 極限値

$$\lim_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h} = c$$

が存在するとき, $f(x)$ は $x = t$ で微分可能であるという。

- c を f の t における微分係数という. このとき,

$$c = f'(t) = \frac{df}{dt}(t)$$

等と記す。

- 開区間 (a, b) の各点で f が微分可能であるとき, $x \in (a, b)$ に対して関数 $x \mapsto f'(x)$ を f の導関数という.

微分可能性と連続性

関数 $f(x)$ が $x = a$ で微分可能であれば, 関数 $f(x)$ は $x = a$ で連続.

この命題の逆は成立しないことに注意せよ.

3.5 合成関数の微分

関数 $z = f(y)$ と $y = g(x)$ が微分可能ならば, 合成関数 $z = f(g(x))$ も x について微分可能で, その微分は

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx} = f'(g(x))g'(x)$$

で与えられる.

3.6 逆関数の微分法

$f(x)$ が微分可能で

$$f'(x) \neq 0$$

かつ逆関数が存在するならば, $f^{-1}(x)$ も微分可能でその微分は

$$(f^{-1}(x))' = \frac{1}{f'(y)}$$

で与えられる.

これは

$$y = f^{-1}(x) \iff x = f(y)$$

から右側の式を x で微分して

$$1 = \frac{d}{dx} f(y) = (f(y))' \cdot y'$$

よりわかる.

3.7 微積分学の基本定理

$f(x)$ が連続ならば, x の関数 $\int_a^x f(t) dt$ は微分可能であって

$$\frac{d}{dx} \int_a^x f(t) dt = f(x)$$

3.8 指数関数の級数展開

任意の実数 x に対して

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \cdots$$

これはマクローリンの定理から得られる.

3.9 積分の計算

以下の議論での積分は可積とする.

- (i) $\int_a^b c f(x) dx = c \int_a^b f(x) dx$ ただし, c は定数
- (ii) $\int_a^b (f(x) \pm g(x)) dx = \int_a^b f(x) dx \pm \int_a^b g(x) dx$
- (iii) $a \leq b$ のとき, $|\int_a^b f(x) dx| \leq \int_a^b |f(x)| dx$
- (iv) $f(x)$ が偶関数²のとき, $\int_{-a}^a f(x) dx = 2 \int_0^a f(x) dx$
- (v) $f(x)$ が奇関数³のとき, $\int_{-a}^a f(x) dx = 0$
- (vi) (部分積分法) $\int_a^b f'(x)g(x) dx = [f(x)g(x)]_a^b - \int_a^b f(x)g'(x) dx$ ただし, f と g は微分可能とする.

3.10 広義積分

\mathbb{R} の半開区間 $I = [a, b)$ で⁴定義された実数値関数 f がつぎをみたすものとする.

- (1) 任意の $u \in I$ に対し, 有界閉区間 $[a, u]$ で f は有界で可積分⁵
- (2) $\lim_{u \rightarrow b-0} \int_a^u f(x) dx = J \in \mathbb{R}$ が存在する.

このとき, J を I における f の広義積分といい,

$$J = \int_a^b f(x) dx$$

²すなわち, $f(x) = f(-x)$ が成立.

³すなわち, $f(x) = -f(-x)$ が成立.

⁴ $b = \infty$ でもよい

⁵すなわち, $\int_a^u |f(x)| dx < \infty$

と書く．

同様にして， \mathbb{R} の半開区間 $I = (a, b]$ で⁶定義された実数値関数 f が

- (1) 任意の $u \in I$ に対し，有界閉区間 $[u, b]$ で f は有界で可積分⁷
 (2) $\lim_{u \rightarrow a+0} \int_u^b f(x) dx = J \in \mathbb{R}$ が存在する．

このとき， f は I 上で広義可積分といい， J を I における f の広義積分といい，

$$J = \int_a^b f(x) dx$$

と書く．広義可積分であることを，広義積分が収束するともいう．

また， $f: (a, b) \rightarrow \mathbb{R}$ に対して，任意の $c \in (a, b)$ をとり⁸，

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$$

の右辺の二つの積分が存在するとき，左辺の (a, b) 上の f の広義積分を右辺で定義⁹する．すなわち，

$$\int_{v \rightarrow a+0, u \rightarrow b-0} \int_v^u f(x) dx = \int_a^b f(x) dx$$

代表的な広義積分

- (i) $\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$
 (ii) $\alpha > 0$ ならば， $\Gamma(\alpha) = \int_0^\infty e^{-x} x^{\alpha-1} dx$ は広義可積分である．特に， α が正の整数のとき，
 (iia) $\Gamma(\alpha + 1) = \alpha \Gamma(\alpha)$
 (iib) $\Gamma(1) = 1$
 (iic) $\Gamma(\alpha + 1) = \alpha!$

⁶ $a = -\infty$ でもよい

⁷ すなわち， $\int_u^b |f(x)| dx < \infty$

⁸ $a = -\infty, b = \infty$ でもよい．

⁹ 内点 c には $\int_a^b f(x) dx$ の値はよらないことに注意せよ．

である。

- (iii) $\alpha > 0, \beta > 0$ のとき, $B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$ は広義可積分である。また,

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

が成立する。

4 偏微分とその計算法

4.1 偏導関数

関数 $f(x, y)$ が点 (a, b) の近くで定義されていて

$$\lim_{(x, y) \rightarrow (a, b)} f(x, y) = f(a, b)$$

がなりたつとき, $f(x, y)$ は (a, b) で連続であるという. また, $f(x, y)$ がある領域 D のすべての点で連続であるとき, $f(x, y)$ は D において連続であるという.

関数 $z = f(x, y)$ が点 (a, b) の近くで定義されているとする. y の値を b に固定して x だけの関数として $f(x, b)$ を考えるとき, x に関して微分可能ならば, $f(x, y)$ は (a, b) で x について偏微分可能であるといい, その微分係数を $f(x, y)$ の x についての偏微分係数とよび,

$$f_x(a, b) \quad \text{または} \quad \frac{\partial}{\partial x} f(a, b)$$

と書く. また, $f(x, y)$ がある領域 D のすべての点で x について偏微分可能であるとき, $f(x, y)$ は D において x について偏微分可能であるといい, D 上で定義された関数 $f_x(x, y)$ を $f(x, y)$ の x についての偏導関数といい,

$$f_x, \quad \frac{\partial}{\partial x} f(x, y), \quad \frac{\partial z}{\partial x}$$

などと書く.

y を固定する代わりに x を固定し, y の関数として微分を考ええると y についての偏微分可能性や偏導関数 f_y ($\frac{\partial}{\partial y} f(x, y)$, または $\frac{\partial z}{\partial y}$) が同様に定義される.

3変数以上の関数についても同様に偏導関数が定義される.

関数 $z = f(x, y)$ が領域 D において x と y について偏微分可能とする. さらに, 偏導関数 $f_x(x, y)$ と $f_y(x, y)$ が領域 D において x と y について偏微分可能ならば, 4種類の2次偏導関数

$$f_{xx} = \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial x} \right), \quad f_{xy} = \frac{\partial}{\partial y} \left(\frac{\partial z}{\partial x} \right), \quad f_{yx} = \frac{\partial}{\partial x} \left(\frac{\partial z}{\partial y} \right), \quad f_{yy} = \frac{\partial}{\partial y} \left(\frac{\partial z}{\partial y} \right),$$

が得られる¹⁰.

¹⁰先に偏微分した方が添え字の先に現われることに注意せよ.

f_{xy} と f_{yx} が一致すること

f_{xy} と f_{yx} がともに (a, b) で連続ならば, $f_{xy}(a, b) = f_{yx}(a, b)$ となる. したがって, f_{xy} と f_{yx} がともに領域 D 上で連続ならば, D 上で $f_{xy} = f_{yx}$ となる.

4.2 合成関数の偏微分

以下では, 必要なところは微分可能もしくは偏微分可能と仮定する.

(1) 関数 $z = f(x, y)$ において, x, y がともに t の関数のとき, 合成関数 $z = f(x(t), y(t))$ の導関数は

$$\frac{dz}{dt} = f_x(x, y) \frac{dx}{dt} + f_y(x, y) \frac{dy}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}$$

となる.

(2) 関数 $z = f(x, y)$ において, x, y がともに u, v の関数のとき, 合成関数 $z = f(x(u, v), y(u, v))$ の u と v についての偏導関数は

$$\begin{aligned} \frac{\partial z}{\partial u} &= \frac{\partial z}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial u} \\ \frac{\partial z}{\partial v} &= \frac{\partial z}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial v} \end{aligned}$$

となる.

4.3 2変数関数のテイラー展開

$f(x, y)$ が点 (a, b) のまわりで 2 階連続偏微分可能ならば,

$$\begin{aligned} f(a+h, b+k) &= f(a, b) + \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right) f(a, b) + \frac{1}{2!} \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right)^2 f(a+\theta h, b+\theta k) \\ &= f(a, b) + hf_x(a, b) + kf_y(a, b) + \frac{h^2}{2} f_{xx}(a+\theta h, b+\theta k) \\ &\quad + hkf_{xy}(a+\theta h, b+\theta k) + \frac{k^2}{2} f_{yy}(a+\theta h, b+\theta k) \end{aligned}$$

となる. ただし, $0 < \theta < 1$ である.

4.4 2 変数関数の極大・極小

$f(x, y)$ が点 (a, b) で極大 (または極小) であるとは, 点 (a, b) の近くでは, 点 (a, b) を除いたすべての点 (x, y) に対して, $f(x, y) < f(a, b)$ (または $f(x, y) > f(a, b)$) が成立することである. また, $f(a, b)$ を極大値 (または極小値) といひ, 極大値と極小値のことを極値という.

極値をとるために十分条件

$f_x(a, b) = 0, f_y(a, b) = 0$ とする. $f_{xx}(a, b) = A, f_{xy}(a, b) = H, f_{yy}(a, b) = B$ とおくととき,

(1) $H^2 - AB < 0$ ならば,

(1-a) $A > 0$ のとき, $f(x, y)$ は (a, b) で極小,

(1-b) $A < 0$ のとき, $f(x, y)$ は (a, b) で極大.

(2) $H^2 - AB > 0$ ならば, $f(x, y)$ は (a, b) で極値を取らない.

(3) $H^2 - AB = 0$ ならば, $f(x, y)$ は (a, b) で極値を取ることも取らないこともある.

条件付き極値をとるために十分条件

条件 $g(x, y) = 0$ のもとで $f(x, y)$ の極値を求めるためには, λ を定数として,

$$F(x, y) = f(x, y) - \lambda g(x, y)$$

とおいたとき,

$$g(x, y) = 0, \quad F_x(x, y) = f_x(x, y) - \lambda g_x(x, y) = 0, \quad F_y(x, y) = f_y(x, y) - \lambda g_y(x, y) = 0$$

をみたす (x, y) の中からさがせばよい. このようにして極値を求める方法をラグランジェの乗数法という.

5 重積分

5.1 重積分の定義: リーマン和と積分可能性

関数 $z = f(x, y)$ は有界閉領域 $D = [a, b] \times [c, d]$ 上の有界とする: すなわち, ある正数 $M > 0$ が存在し,

$$|f(x, y)| \leq M, \quad (x, y) \in D$$

である .

区間 $[a, b]$ と $[c, d]$ を分割する :

$$a = x_0 < x_1 < \cdots < x_m = b$$

$$c = y_0 < y_1 < \cdots < y_n = d$$

領域 D を $\delta_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$, $i = 1, 2, \dots, m, j = 1, 2, \dots, n$, という mn 個の長方形に分割する . この分割を Δ と書き , 分割の幅を

$$|\Delta| = \max_{i,j} (x_i - x_{i-1}, y_j - y_{j-1})$$

で定める .

ひとつの分割 Δ に対し , 体積の近似和

$$\sum_{i=1}^m \sum_{j=1}^n f(\xi_{ij})(x_i - x_{i-1})(y_j - y_{j-1})$$

を f のひとつのリーマン和という . ただし , $\xi_{ij} \in [x_{i-1}, x_i] \times [y_{j-1}, y_j]$ は任意とする .

有界関数 $f(x, y)$ が領域 D 上で積分可能であるとは , ある値 V があって , 任意の正の数 ϵ に対し , 正の数 δ を適当にとると , $|\Delta| < \delta$ なるすべての分割 Δ について

$$\left| \sum_{i=1}^m \sum_{j=1}^n f(\xi_{ij})(x_i - x_{i-1})(y_j - y_{j-1}) - V \right| < \epsilon$$

が成立することである . このとき , V を f の D 上の積分値といい

$$V = \int \int_D f(x, y) dx dy$$

と書くことにする .

5.2 面積確定な有界集合

\mathbb{R}^2 の部分集合 A の定義関数を

$$I_A(x, y) = \begin{cases} 1, & (x, y) \in A, \\ 0, & (x, y) \notin A, \end{cases}$$

で定める． \mathbb{R}^2 の有界集合 A が面積確定であるとは， A の定義関数が積分可能であることである．そして， A の面積を

$$\mu(A) = \iint_A dx dy$$

で定める．

\mathbb{R}^2 の部分集合 A が面積 0 であるとは，任意の正数 ϵ に対して， $[x_{i-1}, x_i] \times [y_{i-1}, y_i]$, $i = 1, 2, \dots, n$, の形の有限個の長方形 $\delta_1, \delta_2, \dots, \delta_n$ を作って

$$A \subset \bigcup_{i=1}^n \delta_i, \quad \sum_{i=1}^n (x_i - x_{i-1})(y_i - y_{i-1}) < \epsilon$$

とできることである．

5.3 重積分の性質

(1) $f(x, y)$ が D 上で連続ならば， $f(x, y)$ は D 上で積分可能．

(2) $D_1 = [a, x_0] \times [c, d]$, $D_2 = [x_0, b] \times [c, d]$ としたとき，

$$\iint_D f(x, y) dx dy = \iint_{D_1} f(x, y) dx dy + \iint_{D_2} f(x, y) dx dy.$$

(3) 関数 f, g が D 上で積分可能ならば， $f + g$ も積分可能で

$$\iint_D (f(x, y) + g(x, y)) dx dy = \iint_D f(x, y) dx dy + \iint_D g(x, y) dx dy.$$

(4) D 上で $f(x, y) \geq 0$ で f は積分可能ならば，

$$\iint_D f(x, y) dx dy \geq 0.$$

(5) $f(x, y)$ が有界閉集合 D 上で有界で可積分ならば， $|f(x, y)|$ も A 上で可積分であって，

$$\left| \iint_D f(x, y) dx dy \right| \leq \iint_D |f(x, y)| dx dy.$$

(6) 任意の有界関数 $f(x, y)$ は面積 0 の有界集合 A 上で可積分であり，

$$\iint_D f(x, y) dx dy = 0.$$

5.4 重積分の計算法：累次積分

領域 D は連続関数 $y = \varphi_1(x)$, $y = \varphi_2(x)$ ($\varphi_1(x) \leq \varphi_2(x)$) のグラフおよび直線 $x = a$, $x = b$ ($a < b$) によって囲まれているものとする。このとき, D 上の連続関数 $f(x, y)$ の積分は

$$\int \int_D f(x, y) dx dy = \int_a^b \left(\int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) dy \right) dx$$

5.5 重積分の変数変換

積分 $\int \int_D f(x, y) dx dy$ において x, y が変数変換

$$\begin{cases} x = g(u, v) \\ y = h(u, v) \end{cases} \quad (\text{A.1})$$

によって, u, v に変換されたとき, 積分の形はどう変わるかを調べよう。1変数の場合は

$$\int_a^b f(x) dx = \int_{\alpha}^{\beta} f(g(t))g'(t) dt$$

であった。ここで $x = g(t)$, $a = g(\alpha)$, $b = g(\beta)$ とした。

変換 (A.1) がふたつの有界領域 D, K 間で一対一の対応であって, 連続微分可能かつそのヤコビ行列式 (ヤコビアン) が K 上のいたるところ

$$J(u, v) = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} \\ \frac{\partial h}{\partial u} & \frac{\partial h}{\partial v} \end{vmatrix} \neq 0$$

であるとする。このとき, $\bar{D}_1 \subset D$ となる面積をもつ領域 D_1 に対し, 変換 (A.1) によって対応する K の部分領域 K_1 も面積確定で, D_1 上の任意の積分可能な関数 $f(x, y)$ に対し, $f(g(u, v), h(u, v))$ は K_1 上で積分可能であって,

$$\int \int_{D_1} f(x, y) dx dy = \int \int_{K_1} f(g(u, v), h(u, v)) |J| du dv$$

が成立する。

5.6 広義積分

D が有界でない場合を考える．有界関数 $f(x, y)$ は，境界¹¹ の面積が 0 である領域 D 内の面積確定な有界閉集合 D_1 上でつねに積分可能とする．¹²

$f(x, y) \geq 0$ ならば， $f(x, y)$ の D 上の広義積分を

$$\int \int_D f(x, y) dx dy = \sup_{K \subset D} \int \int_K f(x, y) dx dy$$

で定める．ただし，右辺の \sup は D 内のあらゆる面積確定な有界閉集合 K についての上限である．

次に， $f(x, y)$ の符号が一定でない場合を考える．そのために

$$\begin{aligned} f^+(x, y) &= \max(f(x, y), 0) \\ f^-(x, y) &= \max(-f(x, y), 0) \end{aligned}$$

とおく． $f^+(x, y) \geq 0$, $f^-(x, y) \geq 0$ であって

$$f(x, y) = f^+(x, y) - f^-(x, y)$$

となる． $\int \int_D f^+(x, y) dx dy < \infty$, $\int \int_D f^-(x, y) dx dy < \infty$ のとき， D 上の $f(x, y)$ の積分を

$$\int \int_D f(x, y) dx dy = \int \int_D f^+(x, y) dx dy - \int \int_D f^-(x, y) dx dy$$

で定める．

$f(x, y)$ の D 上の広義積分が存在¹³ するとする． D 内の面積確定な有界閉集合列 $\{K_n\}_{n=1}^{\infty}$ が次をみたすとする．

¹¹点 $x \in \mathbb{R}^2$ が D の内点でも外点でもないときに x を D の境界という．また， x が D の内点であるとは， x のある ϵ -近傍 $U_\epsilon(x)$ を取れば， $U_\epsilon(x) \subset D$ となることであり， x が D の外点であるとは， x のある ϵ -近傍 $U_\epsilon(x)$ を取れば， $U_\epsilon(x) \subset D^c$ となることである．したがって， \mathbb{R}^2 は境界がない．

¹² ∂D を D の境界としたとき，任意の正数 ϵ に対して， $[x_{i-1}, x_i] \times [y_{i-1}, y_i]$, $i = 1, 2, \dots, n$, の形の有限個の長方形 $\delta_1, \delta_2, \dots, \delta_n$ を作って

$$\partial D \subset \cup_{i=1}^n \delta_i, \quad \sum_{i=1}^n (x_i - x_{i-1})(y_i - y_{i-1}) < \epsilon$$

とできることである．

¹³すなわち， $\int \int_D |f(x, y)| dx dy < \infty$ である．

- (1) $K_1 \subset K_2 \subset \cdots \subset K_n \subset \cdots$.
- (2) 任意の有界閉集合 $K \subset D$ に対して, ある番号 n があって $K \subset K_n$ とできる.

このとき,

$$\iint_D f(x, y) dx dy = \lim_{n \rightarrow \infty} \iint_{K_n} f(x, y) dx dy.$$

最後に, 代表的な広義積分 (i) の証明を与えよう.

(i) $\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$ の証明

まず, 広義積分の定義および有界閉集合 $[0, R] \times [0, R]$ 上では $e^{-(x^2+y^2)}$ は連続であることに注意して,

$$\begin{aligned} I_\infty &= \int_0^\infty e^{-x^2} dx \int_0^\infty e^{-y^2} dy = \lim_{R \rightarrow \infty} \int_0^R e^{-x^2} dx \int_0^R e^{-y^2} dy \\ &= \lim_{R \rightarrow \infty} \iint_{[0, R] \times [0, R]} e^{-(x^2+y^2)} dx dy \end{aligned}$$

となること¹⁴に注意する. いま,

$$L_R = \lim_{R \rightarrow \infty} \int_0^R \int_0^R e^{-(x^2+y^2)} dx dy$$

と

$$I_R = \int \int_{D_R} e^{-(x^2+y^2)} dx dy$$

とおく. ただし, $D_R = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq R^2, x \geq 0, y \geq 0\}$ である. このとき, $L_{R/\sqrt{2}} \leq I_R \leq L_R$ となるので, $\lim_{R \rightarrow \infty} I_R = \lim_{R \rightarrow \infty} L_R$ となる.

いま

$$\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases}$$

とおき, $K_R = \{(r, \theta) : 0 \leq r \leq R, 0 \leq \theta \leq \frac{\pi}{2}\}$ としたとき, D_R と K_R には一対一対応がある. また,

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r$$

¹⁴二番目の等号は一次元の広義積分の定義, 三番目の等号は有界閉集合上の累次積分である.

となるので,

$$\begin{aligned} I_R &= \int \int_{K_R} r e^{-r^2} dr d\theta = \int_0^R \int_0^{\pi/2} r e^{-r^2} d\theta dr = \frac{\pi}{2} \int_0^R r e^{-r^2} dr \\ &= \frac{\pi}{2} \left[-\frac{1}{2} e^{-r^2} \right]_0^R = \frac{\pi}{2} \left[\frac{1}{2} - \frac{1}{2} e^{-R^2} \right] \end{aligned}$$

となる. したがって

$$I_\infty = \frac{\pi}{4} \lim_{R \rightarrow \infty} [1 - e^{-R^2}] = \frac{\pi}{4}$$

となる.

□

6 積分記号下の微分

$f(x|\theta)$ ($\theta \in \Theta$) を確率 (密度) 関数とし, $f(x|\theta)$ は θ の関数として微分可能とする:

$$\frac{\partial}{\partial \theta} f(x|\theta) = \lim_{\delta \rightarrow 0} \frac{f(x|\theta + \delta) - f(x|\theta)}{\delta}$$

これより

$$\int_{-\infty}^{\infty} \frac{\partial}{\partial \theta} f(x|\theta) dx = \int_{-\infty}^{\infty} \lim_{\delta \rightarrow 0} \frac{f(x|\theta + \delta) - f(x|\theta)}{\delta} dx$$

となる. 一方,

$$\frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} f(x|\theta) dx = \lim_{\delta \rightarrow 0} \int_{-\infty}^{\infty} \frac{f(x|\theta + \delta) - f(x|\theta)}{\delta} dx$$

である. もし, 積分記号と極限の交換が保障されるならば, 微分記号と微分記号の交換が保障され,

$$\int_{-\infty}^{\infty} \frac{\partial}{\partial \theta} f(x|\theta) dx = \frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} f(x|\theta) dx$$

が成立することが期待できる. これはどのようなときに保障されるかをみていこう.

定理 A.1 I を \mathbb{R} の开区間とし, $\mathbb{R} \times I$ 上の関数 $h(x, y)$ は y の関数として $y = y_0$ ($y_0 \in I$) において連続とし, \mathbb{R} 上の関数 $g(x)$ でつぎをみたすものが存在する:

(i) すべての $x \in \mathbb{R}$ と $y \in I$ に対して,

$$|h(x, y)| \leq g(x)$$

(ii) $\int_{-\infty}^{\infty} g(x) dx < \infty$ このとき,

$$\lim_{y \rightarrow y_0} \int_{-\infty}^{\infty} h(x, y) dx = \int_{-\infty}^{\infty} \lim_{y \rightarrow y_0} h(x, y) dx$$

が成立する.

上の定理からつぎの定理を導くことができる.

定理 A.2 母数空間 Θ を \mathbb{R} の開集合とし, 確率 (密度) 関数 $f(x|\theta)$ は θ の関数として $\theta = \theta_0$ ($\theta_0 \in \Theta$) で微分可能とする. すなわち, 各 x に対して,

$$\lim_{\delta \rightarrow 0} \frac{f(x|\theta_0 + \delta) - f(x|\theta_0)}{\delta} = \left. \frac{\partial}{\partial \theta} f(x|\theta) \right|_{\theta=\theta_0}$$

となる. さらに, ある関数 $g(x, \theta_0)$ とある定数 $\delta_0 > 0$ が存在して, つぎの条件をみたすとする:

(i) すべての x と $|\delta| \leq \delta_0$ なるすべての δ に対して,

$$\left| \frac{f(x|\theta_0 + \delta) - f(x|\theta_0)}{\delta} \right| \leq g(x, \theta_0)$$

(ii)

$$\int_{-\infty}^{\infty} g(x, \theta_0) dx < \infty$$

このとき,

$$\left[\frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} f(x|\theta) dx \right]_{\theta=\theta_0} = \int_{-\infty}^{\infty} \left[\frac{\partial}{\partial \theta} f(x|\theta) dx \right]_{\theta=\theta_0} dx$$

が成立する.

さらに, $f(x|\theta)$ が θ の関数として Θ 上で微分可能であるとする. 中間値の定理からある定数 δ に対して,

$$\frac{f(x|\theta + \delta) - f(x|\theta)}{\delta} = \left. \frac{\partial}{\partial \theta} f(x|\theta') \right|_{\theta'=\theta+\delta^*(x)}$$

とできることがわかる. ただし, $\delta^*(x)$ は θ と $\theta + \delta$ の間にある.

いま, ある関数 $g(x, \theta)$ が存在して, すべての $\theta \in \Theta$ に対して,

$$\int_{-\infty}^{\infty} g(x, \theta) dx < \infty$$

をみだし, さらに, ある定数 $\delta_0(\theta)$ が存在して, $|\theta' - \theta| \leq \delta_0(\theta)$ なるすべての θ' に対して,

$$\left| \frac{\partial}{\partial \theta} f(x|y) \Big|_{y=\theta'} \right| \leq g(x, \theta) \quad (\text{A.2})$$

をみたらば、すべての $|\delta| \leq \delta_0(\theta)$ に対して、

$$\left| \frac{f(x|\theta + \delta) - f(x|\theta)}{\delta} \right| \leq g(x, \theta)$$

をみたらわかる。各 θ に対して、この議論を適用すれば、つぎの補題を得る：

系 A.1 $f(x|\theta)$ は θ の関数として微分可能で、ある関数 $g(x, \theta)$ が存在して、(A.2) とすべての θ に対して、 $\int_{-\infty}^{\infty} g(x, \theta) dx < \infty$ となるとする。このとき、

$$\frac{\partial}{\partial \theta} \int_{-\infty}^{\infty} f(x|\theta) dx = \int_{-\infty}^{\infty} \frac{\partial}{\partial \theta} f(x|\theta) dx$$

が成立する。

注意 A.1 十分条件 (A.2) をつぎの条件に置き換えることもできる：ある可積分関数 $g(x)$ が存在して、すべての θ に対して、

$$\left| \frac{\partial}{\partial y} f(x|\theta) \right| \leq g(x)$$

をみたら。

例 A.1 確率変数 X は母数 $1/\lambda$ の指数分布に従うとする：

$$f_X(x) = \frac{1}{\lambda} e^{-x/\lambda} I_{(0, \infty)}(x), \quad \lambda > 0$$

いま、

$$\frac{\partial}{\partial \lambda} \mathbb{E}[X^n] = \frac{\partial}{\partial \lambda} \int_0^{\infty} x^n \left(\frac{1}{\lambda} \right) e^{-x/\lambda} dx$$

を計算したいとする。ただし、 n は自然数とする。もし、積分記号と微分記号の交換が可能ならば、

$$\frac{d}{d\lambda} \mathbb{E}[X^n] = \int_0^{\infty} \frac{\partial}{\partial \lambda} x^n \left(\frac{1}{\lambda} \right) e^{-x/\lambda} dx = \int_0^{\infty} \frac{x^n}{\lambda} \left(\frac{x}{\lambda} - 1 \right) e^{-x/\lambda} dx = \frac{1}{\lambda^2} \mathbb{E}[X^{n+1}] - \frac{1}{\lambda}$$

となる。

積分記号と微分記号の交換を正当化するために, $x^n(1/\lambda)e^{-x/\lambda}$ の導関数を評価する:

$$\left| \frac{\partial}{\partial \lambda} \left(\frac{x^n e^{-x/\lambda}}{\lambda} \right) \right| = \frac{x^n e^{-x/\lambda}}{\lambda^2} \left| \frac{x}{\lambda} - 1 \right| \leq \frac{x^n e^{-x/\lambda}}{\lambda^2} \left(\frac{x}{\lambda} + 1 \right)$$

よって, $0 < \delta_0(\lambda) < \lambda$ なる定数 $\delta_0(\lambda)$ (これを簡単に δ_0 と書くことにする) に対して,

$$g(x, \lambda) = \frac{x^n e^{-x/(\lambda+\delta_0)}}{(\lambda-\delta_0)^2} \left(\frac{x}{\lambda-\delta_0} + 1 \right)$$

とおく. すると $|\lambda' - \lambda| \leq \delta_0$ なる λ' に対して,

$$\left| \frac{\partial}{\partial y} \left(\frac{x^n e^{-x/y}}{y} \right) \Big|_{y=\lambda'} \right| \leq g(x, \lambda)$$

となる. さらに, $\lambda - \delta_0 > 0$ であるかぎり, $\int_0^\infty g(x, \lambda) dx < \infty$ となる. したがって, 積分記号と微分記号の交換は正当化される.

例 A.2 確率変数 X は平均 μ , 分散 1 の正規分布に従うとする. X の積率母関数は

$$M_X(t) = \mathbb{E}[X^{tX}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-(x-\mu)^2/2} dx$$

である. X の積率を求めるためには, 積率母関数を微分すればよい:

$$\frac{d}{dt} M_X(t) = \frac{d}{dt} \mathbb{E}[X^{tX}] = \mathbb{E}[X e^{tX}]$$

となる. 積分記号と微分記号の交換が保障せれば, これは正当化される. 系 A.1 を用いて, これを確認しよう. そのために, ある可積分な関数 $g(x, t)$ とある δ_0 が存在して, $|t' - t| \leq \delta_0$ なるすべての t' に対して,

$$\left| \frac{\partial}{\partial y} e^{yx} e^{-(x-\mu)^2/2} \Big|_{y=t'} \right| \leq g(x, t) \quad (\text{A.3})$$

が成立することを示せばよい. まず,

$$\left| \frac{\partial}{\partial t} e^{tx} e^{-(x-\mu)^2/2} \right| \leq \left| x e^{tx} e^{-(x-\mu)^2/2} \right| \leq |x| e^{tx} e^{-(x-\mu)^2/2}$$

となる。つぎに, $g(x, t)$ を $x \geq 0$ と $x < 0$ の場合に分けて定義しよう:

$$g(x, t) = \begin{cases} |x|e^{(t-\delta_0)x}e^{-(x-\mu)^2/2} & (x < 0), \\ |x|e^{(t+\delta_0)x}e^{-(x-\mu)^2/2} & (x \geq 0). \end{cases}$$

$x \geq 0$ の場合には,

$$g(x, t) = xe^{-(x^2-2x(\mu+t+\delta_0)+\mu^2)}$$

であり, 指数関数の部分を平方完成すれば,

$$x^2 - 2x(\mu + t + \delta_0) + \mu^2 = \{x - (\mu + t + \delta_0)\}^2 + \mu^2 - (\mu + t + \delta_0)^2$$

となる。したがって,

$$g(x, t) = xe^{-[x-(\mu+t+\delta_0)]^2/2}e^{-[\mu^2-(\mu+t+\delta_0)^2]/2}$$

となる。したがって,

$$\int_0^\infty g(x, t) dx < e^{-[\mu^2-(\mu+t+\delta_0)^2]/2} \int_0^\infty xe^{-[x-(\mu+t+\delta_0)]^2/2} dx < \infty$$

となることがわかる。 $x < 0$ の場合も同様にすれば, $\int_{-\infty}^0 g(x, t) dx < \infty$ が確認できる。