確率統計学 I

杉浦　誠

2022 年 7 月 6 日

目次

確率統計学 I では下の文献表の [Y], [AEL] におおよそ沿って、

1 確率空間, **2** 確率変数, **3** 確率変数の変換, **4** 期待値

について解説します。

講義ノートは WebClass にアップロードしていきます。全体像は

`http://www.math.u-ryukyu.ac.jp/~sugiura/`

に置きますが、最新版の WebClass のものを参照ください。

授業は複合棟 412 室で毎週水曜日 10:20 – に行います。出席を希望する人は感染対策をして出席ください。

ただし、**発熱等体調に不良がある人、新型コロナウイルスに感染した若しくは感染したおそれ等がある人は大学に来てはいけません**。特に、下記の中間テスト、期末テストの日にそうなった場合は連絡ください。後日追試を行います。

また、WebClass に講義ノートを掲載するので、必ずしも講義に出席する必要はありません。毎回、授業終了後に次回分以降の講義ノートを少しづつ WebClass にアップロードしていく予定ですが、その資料の説明としてその日の授業でどこまで進んだかを記載して行く予定です。WebClass でも質問を受け付けます。

成績について　中間テスト (6 月 29 日) と期末テスト (8 月 3 日) の合計点数で判定します。6 割以上の得点で合格です。テストの範囲は遅くとも二週間前には WebClass を通じて連絡します。確率統計学 I の受講者は中間テストと期末テストを必ず受けて下さい。教育実習の日程と重なって中間テストが受けられない場合は 4 月中に、インターンシップなどで期末テストが受けられない場合は 6 月中に、杉浦に連絡ください。

参考文献

[AEL] 浅野, 江島, 李 共著 基本統計学 森北出版

[Y] 柳川 尭 著 統計数学 近代科学社

[F] 藤田岳彦 著 弱点克服大学生の確率・統計 東京図書

[Ku] 黒田耕嗣 著 生保年金数理 I 増補版 培風館

[Ko] 小寺平治 著 明解演習 数理統計 共立出版

[TS] 統計と社会の教科書「新確率統計」

教科書を購入する必要はありませんが、もし購入する場合は問題集形式で詳しく解答が記述されている [F] か [Ko] をお勧めします。

合わせて上記の URL で昨年度までに作成したアクチュアリー試験数学対策用のノートもアップロードしておきます。興味ある人はご覧ください。

1　確率空間

1.1　確率の定義

定義 1.1 集合 $\Omega(\neq \emptyset)$ の部分集合からなる集合を Ω 上の集合族という。Ω 上の集合族 $\mathcal{F}$ が次の条件

(i)　$\Omega \in \mathcal{F}$

(ii)　$A \in \mathcal{F} \Longrightarrow A^{\mathrm{c}} \in \mathcal{F}$　(A^{c} は A の補集合を表す。)

(iii)　$A_1, A_2, \ldots, A_n, \ldots \in \mathcal{F} \Longrightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

を満たすとき、この $\mathcal{F}$ を Ω 上の σ 集合族という。また、$(\Omega, \mathcal{F})$ を可測空間という。

確率論では、σ 集合族 $\mathcal{F}$ の元を事象という。特に、Ω を全事象と、$\emptyset = \Omega^{\mathrm{c}}$ を空事象という。

定理 1.1 $(\Omega, \mathcal{F})$ を可測空間、$\{B_n\} \subset \mathcal{F}$ とすると次が成り立つ。

(0)　$\emptyset \in \mathcal{F}$,　(1)　$\displaystyle\bigcup_{i=1}^{n} B_i \in \mathcal{F}$,　(2)　$\displaystyle\bigcap_{i=1}^{n} B_i \in \mathcal{F}$,　(3)　$\displaystyle\bigcap_{n=1}^{\infty} B_n \in \mathcal{F}$,　(4)　$B_1 \backslash B_2 := B_1 \cap B_2^{\mathrm{c}} \in \mathcal{F}$

証明: (0) 定義 1.1 (i), (ii) より $\Omega^{\mathrm{c}} \in \mathcal{F}$. よって、$\Omega^{\mathrm{c}} = \emptyset$ より主張を得る。

(1) $A_i = B_i$ $(1 \leq i \leq n)$, $A_i = \emptyset$ $(i \geq n+1)$ とおくと、仮定と (0) より $\forall i$ に対し $A_i \in \mathcal{F}$ であるから、定義 1.1 (iii) より $\bigcup_{i=1}^{n} B_i = \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ を得る。

(2) 定義 1.1 (ii) より $B_i^{\mathrm{c}} \in \mathcal{F}$ であるから、(1) と de Morgan の法則により $\bigcap_{i=1}^{n} B_i = \left(\bigcup_{i=1}^{n} B_i^{\mathrm{c}}\right)^{\mathrm{c}} \in \mathcal{F}$.

(3) 定義 1.1 (iii) と de Morgan の法則により (2) と全く同様に証明できる。

(4) $A_1 = B_1$, $A_2 = B_2^{\mathrm{c}}$, $n = 2$ として (2) を用いればよい。　□

定義 1.2 $(\Omega, \mathcal{F})$ を可測空間とする。次の条件を満たす $P : \mathcal{F} \to [0, 1]$ を $(\Omega, \mathcal{F})$ 上の確率測度という。[*1]

(i)　$P(\Omega) = 1$

(ii)　$A_n \in \mathcal{F}$, $n = 1, 2, \ldots,$ が任意の $m \neq n$ なる組に対し $A_m \cap A_n = \emptyset$ を満たせば (このとき $A_1, A_2, \ldots$ は互いに排反であるという)、$P\Big(\displaystyle\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} P(A_n)$.

このとき、$(\Omega, \mathcal{F}, P)$ を確率空間という。また、$\mathcal{F}$ の元を事象、$A \in \mathcal{F}$ に対し $P(A)$ を事象 A の確率という。

定理 1.2 $(\Omega, \mathcal{F}, P)$ を確率空間とする。

(1)　$P(\emptyset) = 0$

(2)　$B_i \in \mathcal{F}$, $i = 1, \ldots, n,$ が互いに排反であれば、$P\Big(\displaystyle\bigcup_{i=1}^{n} B_i\Big) = \sum_{i=1}^{n} P(B_i)$.

(3)　$A, B \in \mathcal{F}, A \subset B \Longrightarrow P(A) \leq P(B)$

(4)　$B \in \mathcal{F} \Longrightarrow P(B^{\mathrm{c}}) = 1 - P(B)$

(5)　$A, B \in \mathcal{F}$ に対し $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

証明: (1) 定理 1.1 に注意して、$a = P(\emptyset)$ とおくと、$0 \leq a \leq 1$. ここで、$A_n = \emptyset$ $(\forall n)$ とおくと、$\bigcup_{n=1}^{\infty} A_n = \emptyset$ で、$A_m \cap A_n = \emptyset$ $(m \neq n)$ であるから定義 1.2 (ii) より $P(\emptyset) = \sum_{n=1}^{\infty} P(\emptyset)$ を得る。もし $a > 0$ であれば、これは $a = \infty$ を意味し、$0 \leq a \leq 1$ に矛盾する。従って、$P(\emptyset) = 0$ となる。

(2) $A_i = B_i$ $(1 \leq i \leq n)$, $A_i = \emptyset$ $(i \geq n+1)$ とおくと、$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{n} B_i$ で、$A_i \cap A_j = \emptyset$ $(i \neq j)$ であるから、定義 1.2 (ii) と (1) より、$P(\bigcup_{i=1}^{n} B_i) = \sum_{i=1}^{n} P(B_i) + \sum_{i=n+1}^{\infty} P(\emptyset) = \sum_{i=1}^{n} P(B_i)$.

(3) $B_1 = A$, $B_2 = B \backslash A$ とすると $B = B_1 \cup B_2$, $B_1 \cap B_2 = \emptyset$ より (2) から、$P(B) = P(B_1) + P(B_2) \geq P(A)$.

[*1] この授業では「測度」は存在するものと仮定する。測度の構成方法は関数解析学 I,II を受講して勉強ください。

(4) $B \cup B^c = \Omega$, $B \cap B^c = \emptyset$ より、定義 1.2 (i) と (2) により、$P(B) + P(B^c) = P(\Omega) = 1$ となり従う。
(5) $B_1 = A \cap B, B_2 = A \cap B^c, B_3 = A^c \cap B$ とおくと、B_1, B_2, B_3 は互いに排反であるから、(2) により $A = B_1 \cup B_2$ から $P(A) = P(B_1) + P(B_2)$, $B = B_1 \cup B_3$ から $P(B) = P(B_1) + P(B_3)$, $A \cup B = B_1 \cup B_2 \cup B_3$ から $P(A \cup B) = P(B_1) + P(B_2) + P(B_3)$. よって、与式は従う。 □

定理 1.3 $A_1, A_2, \ldots \in \mathcal{F}$ に対し $P(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} P(A_n)$.

証明: $B_1 = A_1$, $B_n = A_n \backslash (\bigcup_{k=1}^{n-1} A_k)$, $n = 2, 3, \ldots$, とおくと、$B_1, B_2, \ldots$ は互いに排反で $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. よって、定義 1.2 (ii) により、

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

ここで、最後の不等式は $A_n \subset B_n$ と定理 1.2(3) を用いた。 □

系 1.1 $A_1, A_2, \ldots \in \mathcal{F}$ に対して、すべての $n \in \mathbf{N}$ に対し $P(A_n) = 1$ ならば $P(\bigcap_{n=1}^{\infty} A_n) = 1$.

問 1.1 系 1.1 を定理 1.3 を用いて証明せよ。

定理 1.4 $(\Omega, \mathcal{F}, P)$ を確率空間、$\{B_n\} \subset \mathcal{F}$ とする。

(1) $B_1 \subset B_2 \subset \cdots \subset B_n \subset \cdots \implies P(\bigcup_{n=1}^{\infty} B_n) = \lim_{n \to \infty} P(B_n)$.
(2) $B_1 \supset B_2 \supset \cdots \supset B_n \supset \cdots \implies P(\bigcap_{n=1}^{\infty} B_n) = \lim_{n \to \infty} P(B_n)$.

証明: (1) $A_1 = B_1$, $A_n = B_n \backslash \bigcup_{k=1}^{n-1} B_k$ $(n \geq 2)$ とおくと、$A_1, A_2, \ldots$ は互いに排反で $\bigcup_{k=1}^{n} A_k = B_n$ $(n \in \mathbf{N})$, $\bigcup_{k=1}^{\infty} A_k = \bigcup_{n=1}^{\infty} B_n$. よって、定義 1.2 (ii), 定理 1.2 (2) により、

$$P(\bigcup_{n=1}^{\infty} B_n) = P(\bigcup_{k=1}^{\infty} A_k) = \sum_{k=1}^{\infty} P(A_k) = \lim_{n \to \infty} \sum_{k=1}^{n} P(A_k) = \lim_{n \to \infty} P(\bigcup_{k=1}^{n} A_k) = \lim_{n \to \infty} P(B_n).$$

(2) $B_1^c \subset \cdots \subset B_n^c \subset \cdots$ であるから、定理 1.2 (4), de Morgan の法則と前半より

$$P(\bigcap_{n=1}^{\infty} B_n) = 1 - P((\bigcap_{n=1}^{\infty} B_n)^c) = 1 - P(\bigcup_{n=1}^{\infty} B_n^c) = 1 - \lim_{n \to \infty} P(B_n^c) = \lim_{n \to \infty} (1 - P(B_n^c)) = \lim_{n \to \infty} P(B_n).$$ □

定理 1.5 Ω を集合とし $\mathcal{A}$ をその部分集合族とする。$\mathcal{A}$ を含む Ω 上の σ 集合族はいくつもある。これに index を付け、$\mathcal{F}_\lambda$ $(\lambda \in \Lambda)$ と表す。このとき、次が成り立つ。
(1) $\bigcap_{\lambda \in \Lambda} \mathcal{F}_\lambda$ は σ 集合族である。
(2) $\bigcap_{\lambda \in \Lambda} \mathcal{F}_\lambda$ は $\mathcal{A}$ を含む σ 集合族の中で最小の σ 集合族である。

証明: (1) は演習問題とする。(2) まず、$\mathcal{A}$ を含む Ω 上の σ 集合族として、Ω のべき集合があることに注意する (よって Λ は空ではない)。(1) より、$\mathcal{F} := \bigcap_{\lambda \in \Lambda} \mathcal{F}_\lambda$ が $\mathcal{A}$ を含む σ 集合族の中で最小であることを示せばよい。今、$\mathcal{F}$ より小さい (真部分集合となる)、$\mathcal{A}$ を含む σ 集合族 $\mathcal{F}_0$ があったとする。$\mathcal{F}_0$ は $\mathcal{A}$ を含む σ 集合族なので、$\mathcal{F}_\lambda$ のいずれかである。従って、$\mathcal{F}_0 \supset \bigcap_{\lambda \in \Lambda} \mathcal{F}_\lambda = \mathcal{F}$ であるが、これは $\mathcal{F}_0$ が $\mathcal{F}$ の真部分集合であることに矛盾する。 □

問 1.2 定理 1.5 (1) を証明せよ。

問 1.3 $A_1, \ldots, A_n$ を事象とする。$p_r = \displaystyle\sum_{1 \leq i_1 < \cdots < i_r \leq n} P(\bigcap_{k=1}^{r} A_{i_k})$ とするとき、$P(\bigcup_{i=1}^{n} A_i) = \displaystyle\sum_{r=1}^{n} (-1)^{r-1} p_r$ を示せ。(一般の場合が難しければ、$n = 3$ と $n = 4$ の場合を示せ。)

問 1.4 $X_1, X_2, X_3, X_4, X_5, X_6$ を、6 個の数字 $1, 2, 3, 4, 5, 6$ の順列とするとき、(1) $X_1 \neq 1$ かつ $X_2 \neq 2$ かつ $X_3 \neq 3$ となる確率と (2) $X_1 \neq 1$ かつ $X_2 \neq 2$ かつ $X_3 \neq 3$ かつ $X_4 \neq 4$ となる確率を求めよ。

1.2 条件つき確率と事象の独立

確率空間 $(\Omega, \mathcal{F}, P)$ において $C \in \mathcal{F}$, $P(C) > 0$ のとき、次のように定める。

$$P(A|C) := \frac{P(A \cap C)}{P(C)} \qquad (A \in \mathcal{F})$$

定理 1.6 $P(\,\cdot\,|C)$ は可測空間 $(\Omega, \mathcal{F})$ 上の確率測度である。

証明: 定理 1.2 (3) より $0 \leq P(A \cap C) \leq P(C)$ となり、$0 \leq P(A|C) \leq 1$. 定義 1.2 (i) は明らか。(ii) は $\{A_n\} \subset \mathcal{F}$ が互いに排反であれば、$B_n = A_n \cap C$ とすると $B_n \in \mathcal{F}$ で $B_m \cap B_n = A_m \cap A_n \cap C = \emptyset$ $(m \neq n)$ より、

$$P(\bigcup_{n=1}^{\infty} A_n|C) = \frac{P((\bigcup_{n=1}^{\infty} A_n) \cap C)}{P(C)} = \frac{P(\bigcup_{n=1}^{\infty} B_n)}{P(C)} = \frac{\sum_{n=1}^{\infty} P(B_n)}{P(C)} = \sum_{n=1}^{\infty} \frac{P(A_n \cap C)}{P(C)} = \sum_{n=1}^{\infty} P(A_n|C). \square$$

定義 1.3 $P(A|C)$ を事象 C が与えられたときの $A \in \mathcal{F}$ の条件つき確率という。また、$P(\,\cdot\,|C)$ を C が与えられたときの条件つき確率測度という。

定理 1.7 次が成り立つ。

(1) (**乗法公式**) $P(A_1 \cap A_2 \cap \cdots \cap A_n) > 0$, $A_i \in \mathcal{F}$ $(i = 1, 2, \ldots, n)$ のとき

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1) \cdots P(A_n|A_1 \cap \cdots \cap A_{n-1})$$

(2) (**全確率の公式**) $\{A_n\} \subset \mathcal{F}$ を Ω の分割、即ち、$\{A_n\}$ は互いに排反で $P(A_n) > 0$ $(n \in \mathbf{N})$ かつ $\Omega = \bigcup_{n=1}^{\infty} A_n$ を満たすとする。このとき、

$$P(A) = \sum_{i=1}^{\infty} P(A_i)P(A|A_i) \qquad (A \in \mathcal{F})$$

(3) **(Bayes の定理)** $\{A_n\} \subset \mathcal{F}$ を Ω の分割とする。このとき、$B \in \mathcal{F}$ で $P(B) > 0$ なるものに対して次が成立する。

$$P(A_k|B) = \frac{P(A_k)P(B|A_k)}{\sum_{n=1}^{\infty} P(A_n)P(B|A_n)}.$$

注意 1.1 定理 1.7 (2), (3) は Ω が有限個の $A_1, \ldots, A_N \in \mathcal{F}$ によって分割されている場合にも成立する。証明は、定義 1.2 (ii) のかわりに定理 1.2 (2) を用いることで、全く同様にできる。

証明: (1) $P(A \cap C) = P(A|C)P(C)$ に注意する。これを繰り返し用いて、
(右辺)$= P(A_1 \cap \cdots \cap A_{n-1})P(A_n|A_1 \cap \cdots A_{n-2} \cap A_{n-1}) = P(A_1 \cap \cdots \cap A_{n-2})P(A_{n-1}|A_1 \cap \cdots \cap A_{n-2})P(A_n|A_1 \cap \cdots \cap A_{n-1}) = \cdots =$(左辺).
(2) (右辺)$= \displaystyle\sum_{n=1}^{\infty} P(A_n)\frac{P(B \cap A_n)}{P(A_n)} = P(\bigcup_{n=1}^{\infty}(B \cap A_n)) = P(B \cap \bigcup_{n=1}^{\infty} A_n) = P(B \cap \Omega) = P(B)$.
ここで、第 2 の等号は $m \neq n$ であれば $(B \cap A_m) \cap (B \cap A_n) = B \cap A_m \cap A_n = \emptyset$ より $\{B \cap A_n\}$ は互いに排反であることに注意し定義 1.2 (ii) を適用した。
(3) (右辺)$= \dfrac{P(A_k)\frac{P(B \cap A_k)}{P(A_k)}}{P(B)} = \dfrac{P(B \cap A_k)}{P(B)} = P(A_k|B)$. ここで、第 1 の等号は (2) を用いた。 $\square$

定義 1.4 事象の列 $\{B_n\} \subset \mathcal{F}$ (有限個でも無限個でもよい) に対し、

$$P(B_{n_1} \cap B_{n_2} \cap \cdots \cap B_{n_l}) = P(B_{n_1})P(B_{n_2}) \cdots P(B_{n_l}) \qquad \text{for all } n_1 < n_2 < \cdots < n_l \qquad (1.1)$$

を満たすとき、事象の列 $\{B_n\}$ は独立であるという。

問 1.5 (1) $\Omega=\{1,2,3,4\}$, $P(\{k\})=1/4$ $(k=1,2,3,4)$ において、$A=\{1,2\}$, $B=\{1,3\}$, $C=\{2,3\}$ とおくと、事象 A,B,C に対しどの 2 つの事象の組も独立であるが、A,B,C が独立とならないことを示せ。
(2) $\Omega=\{1,2,3,4\}$, $P(\{1\})=\dfrac{1}{\sqrt{2}}-\dfrac{1}{4}$, $P(\{2\})=\dfrac{3}{4}-\dfrac{1}{\sqrt{2}}$, $P(\{3\})=P(\{4\})=1/4$ において、$A=\{1,2\}$, $B=\{2,3\}$, $C=\{2,4\}$ とおくと、$P(A\cap B\cap C)=P(A)P(B)P(C)$ となるが、事象 A,B,C に対しどの 2 つの事象の組も独立とならないことを示せ。

定理 1.8 事象の列 $B_1,B_2,\ldots,B_n$ が独立であれば、任意の $k=1,\ldots,n$ に対して、$B_1^{\mathrm{c}},\ldots,B_k^{\mathrm{c}},B_{k+1},\ldots,B_n$ も独立となる。

証明: $B_1,B_2,\ldots,B_n$ が独立ならば、$B_1^{\mathrm{c}},B_2,\ldots,B_n$ も独立となることを示せばよい。実際、これより $B_2,B_3,\ldots,B_n,B_1^{\mathrm{c}}$ が独立となるから、$B_2^{\mathrm{c}},B_3,\ldots,B_n,B_1^{\mathrm{c}}$ は独立となり、これを繰り返すことで主張を得る。

このため任意の $1\le n_1<n_2<\cdots<n_l\le n$ に対し (1.1) に相当する式を示せばよが、$n_1\ge 2$ の場合は B_1^{c} が関与しないため (1.1) は明らかに成り立つ。$n_1=1$ の場合:

$$P(B_1^{\mathrm{c}}\cap B_{n_2}\cap\cdots\cap B_{n_l})=P(B_1^{\mathrm{c}})P(B_{n_2})\cdots P(B_{n_l}) \tag{1.2}$$

を示せばよい。まず、

$$P(B_1\cap B_{n_2}\cap\cdots\cap B_{n_l})+P(B_1^{\mathrm{c}}\cap B_{n_2}\cap\cdots\cap B_{n_l})=P(B_{n_2}\cap\cdots\cap B_{n_l})$$

に注意する。ここで、$B_1,B_2,\cdots,B_n$ の独立性を用いると

$$P(B_1\cap B_{n_2}\cap\cdots\cap B_{n_l})=P(B_1)P(B_{n_2})\cdots P(B_{n_l}),\quad P(B_{n_2}\cap\cdots\cap B_{n_l})=P(B_{n_2})\cdots P(B_{n_l})$$

となる。よって、

$$P(B_1^{\mathrm{c}}\cap B_{n_2}\cap\cdots\cap B_{n_l})=\{1-P(B_1)\}P(B_{n_2})\cdots P(B_{n_l})=P(B_1^{\mathrm{c}})P(B_{n_2})\cdots P(B_{n_l}).\qquad\square$$

問 1.6 $\Omega=\{1,2,\ldots,210\}$ から一つの数字をランダムに選び、その数が k の倍数であるか考える。
$A_k=\{km\in\Omega; m\in\mathbb{Z}\}$ とする。
(1) k が 210 の約数ならば、$P(A_k)=1/k$ となることを確認せよ。また、$P(A_4)$ を求めよ。
(2) A_3 と A_5 が独立であることを示せ。また、A_3 と A_4 は独立か調べよ。
(3) $P(A_6|A_4)$, $P(A_4|A_6)$ を求めよ。
(4) $P(A_2\cup A_3\cup A_7)$, $P(A_2\cup A_3\cup A_5\cup A_7)$ を求めよ。

2 確率変数

2.1 確率変数の定義

可測空間 $(\Omega,\mathcal{F})$ において、Ω 上の実数値関数 $X:\Omega\to\mathbf{R}$ が $\forall a\in\mathbf{R}$ に対して

$$\{\omega; X(\omega)\le a\}\in\mathcal{F}$$

を満たすとき X を $(\Omega,\mathcal{F})$ 上の確率変数という。

例 2.1 1 回のコイン投げでは、$\Omega=\{H,T\}$, $\mathcal{F}=2^\Omega=\{\emptyset,\{H\},\{T\},\{H,T\}\}$ (<u>H</u>ead: 表, <u>T</u>ail: 裏) となる。
このとき、$X(\omega)=\begin{cases}1 & (\omega=H)\\ 0 & (\omega=T)\end{cases}$ とすると、$\{\omega;X(\omega)\le a\}=\begin{cases}\Omega & (1\le a)\\ \{T\} & (0\le a<1)\\ \emptyset & (a<0)\end{cases}$ となるので、
$\forall a\in\mathbf{R}$ に対して $\{\omega;X(\omega)\le a\}\in\mathcal{F}$ がわかる。従って、X は確率変数である。

定理 2.1 X は $(\Omega, \mathcal{F})$ 上の確率変数 $\iff$ $\forall a \in \mathbf{R}$ に対して $\{\omega; X(\omega) < a\} \in \mathcal{F}$.

証明: $(\Longrightarrow)$ $\{\omega; X(\omega) < a\} = \displaystyle\bigcup_{n=1}^{\infty} \{\omega; X(\omega) \leq a - \frac{1}{n}\}$ より従う。
$(\Longleftarrow)$ $\{\omega; X(\omega) \leq a\} = \displaystyle\bigcap_{n=1}^{\infty} \{\omega; X(\omega) < a + \frac{1}{n}\}$ と定理 1.1 (2) より従う。 □

ここで、σ 集合族の定義から

$$\{\omega; X(\omega) \leq a\} \in \mathcal{F} \iff \{\omega; X(\omega) > a\} \in \mathcal{F},$$
$$\{\omega; X(\omega) < a\} \in \mathcal{F} \iff \{\omega; X(\omega) \geq a\} \in \mathcal{F}.$$

これと定理 2.1 より、一般に次が成立する。

定理 2.2 X は $(\Omega, \mathcal{F})$ 上の確率変数 $\iff$ 任意の区間 I に対して $\{\omega; X(\omega) \in I\} \in \mathcal{F}$.

証明: $(\Longleftarrow)$ は明らか。$(\Longrightarrow)$ について、$I = (-\infty, b), (-\infty, b], (a, \infty), [a, \infty)$ の場合はすでに示した。$I = (a, b), [a, b], (a, b], [a, b)$ の場合に $\{\omega; X(\omega) \in I\} \in \mathcal{F}$ を示せばよい。例えば、$I = (a, b)$ のときは、

$$\{\omega; X(\omega) \in (a, b)\} = \{\omega; X(\omega) < b\} \cap \{\omega; X(\omega) > a\}$$

より従う。他は演習問題とする。 □

2.2 可測関数と確率変数

定理 1.5 より、$\mathbf{R}$ 上の区間 $(a, b]$ 全体からなる族を含む最小の σ 集合族が存在する。この集合族を $\mathbf{R}$ の Borel 集合族といい、$\mathcal{B}(\mathbf{R})$ と表す。Borel 集合族に属する集合を Borel 集合という。

補題 2.1 高々可算個の元からなる集合 E は Borel 集合である。また、開集合 G も Borel 集合である。

証明: 一点からなる集合 $\{x\}$ は $\{x\} = \bigcap_{n=1}^{\infty} \left(x - \dfrac{1}{n}, x\right]$ となるから Borel 集合である。よって、定義 1.1 (iii) により E は Borel 集合となる。次に開集合 G の各点 $x \in G$ に対して、$a_x, b_x \in \mathbf{Q}$ があって、$x \in (a_x, b_x] \subset G$ とできる。このとき、$\bigcup_{x \in G} (a_x, b_x] = G$ であるが、有理数体 $\mathbf{Q}$ は可算集合なので、$\{(a_x, b_x); x \in G\} = \{(a'_n, b'_n); n \in \mathbf{N}\}$ とでき、$G = \bigcup_{n \in \mathbf{N}} (a'_n, b'_n]$ となる。よって、G は Borel 集合。 □

補題 2.2 可測空間 $(\Omega, \mathcal{F})$ において、関数 $X : \Omega \to \mathbf{R}$ によって与えられる $\mathbf{R}$ の部分集合族 $\{A \subset \mathbf{R}; X^{-1}(A) \in \mathcal{F}\}$ は σ 集合族である。ただし、$X^{-1}(A) = \{\omega \in \Omega; X(\omega) \in A\}$ とする。

証明: $X^{-1}(\mathbf{R}) = \Omega$, $X^{-1}(A^{\mathrm{c}}) = (X^{-1}(A))^{\mathrm{c}}$, $X^{-1}(\bigcup_{n=1}^{\infty} A_n) = \bigcup_{n=1}^{\infty} X^{-1}(A_n)$ となることを示せば、容易に証明できるので演習問題とする。 □

定理 2.3 X が $(\Omega, \mathcal{F})$ 上の確率変数 $\iff$ 任意の Borel 集合 $B \in \mathcal{B}(\mathbf{R})$ に対して $\{\omega; X(\omega) \in B\} \in \mathcal{F}$.

証明: $(\Longleftarrow)$ $B = (-\infty, a]$ ととれば明らか。
$(\Longrightarrow)$ X を確率変数とし、$\mathcal{A} = \{A \subset \mathbf{R}; X^{-1}(A) \in \mathcal{F}\}$ とおく。定理 2.2 より $(a, b] \in \mathcal{A}$ であり、補題 2.2 より $\mathcal{A}$ は σ 集合族である。従って、$\mathcal{B}(\mathbf{R})$ は区間 $(a, b]$ 全体からなる集合族を含む最小の σ 集合族であるから $\mathcal{A} \supset \mathcal{B}(\mathbf{R})$ となる。よって $B \in \mathcal{B}(\mathbf{R})$ なら $B \in \mathcal{A}$ となり、$X^{-1}(B) \in \mathcal{F}$ を得る。 □

定義 2.1 可測空間 $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ において、関数 $f : \mathbf{R} \to \mathbf{R}$ が任意の $a \in \mathbf{R}$ に対して $\{x; f(x) \leq a\} \in \mathcal{B}(\mathbf{R})$ を満たすとき、f は $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上の可測関数、あるいは Borel 可測関数であるという。

注意 2.1 f が $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上で可測であることは、f が $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上の確率変数であることは同値である。

定理 2.4 高々可算個の点を除いて連続な関数 $f : \mathbf{R} \to \mathbf{R}$ は Borel 可測である。特に、連続関数は Borel 可測関数である。

証明: D_f は高々可算個の点からなる集合で、f は $\mathbf{R}\backslash D_f$ で連続とする。$\forall a \in \mathbf{R}$ に対して $\{x; f(x) > a\}$ が Borel 集合であることを示せばよい。$C_{a,1} := \{x \in D_f; f(x) > a\}$ は高々可算なのでこれは Borel 集合。また、$C_{a,2} := \{x \in \mathbf{R}\backslash D_f; f(x) > a\}$ は開集合 G を用いて $C_{a,2} = G\backslash D_f$ と表せる。実際、$\forall x \in C_{a,2}$ に対して、$\delta'_x > 0$ を $y \in (x - \delta'_x, x + \delta'_x)\backslash D_f$ ならば $|f(y) - f(x)| < f(x) - a$ となるように選べば、このとき $f(y) > a$ なので $(x - \delta'_x, x + \delta'_x)\backslash D_f \subset C_{a,2}$. よって、$G = \bigcup_{x \in C_{a,2}}(x - \delta'_x, x + \delta'_x)$ とおくと、G は $\mathbf{R}$ の開集合で、$C_{a,2} = G\backslash D_f$ と表せるので補題 2.1 より $C_{a,2} \in \mathcal{B}(\mathbf{R})$. よって、$\{x; f(x) < a\} = C_{a,1} \cup C_{a,2} \in \mathcal{B}(\mathbf{R})$. □

定理 2.5 X が $(\Omega, \mathcal{F})$ 上の確率変数、$f : \mathbf{R} \to \mathbf{R}$ が Borel 可測関数のとき、$f(X(\omega))$ は $(\Omega, \mathcal{F})$ 上の確率変数である。

証明: $\forall a \in \mathbf{R}$ とし、$B = \{x; f(x) \leq a\}$ とおくと、f は可測だから $B \in \mathcal{B}(\mathbf{R})$. よって、定理 2.3 により $\{\omega; f(X(\omega)) \leq a\} = \{\omega; X(\omega) \in B\} \in \mathfrak{B}$ を得る。 □

例 2.2 定理 2.4、定理 2.5 より、X が確率変数であるとき $aX + b$ (a, b は定数), X^a, $\sqrt{X}$, $\log X$, e^X, $\cos X$ などは、すべて確率変数である。

2.3 分布関数

定義 2.2 (1) 確率空間 $(\Omega, \mathcal{F}, P)$ 上の確率変数 X に対して、$F_X(x) = P(X \leq x) = P(\{\omega; X(\omega) \leq x\})$ で定義される関数 $F_X : \mathbf{R} \to [0, 1]$ を確率変数 X の分布関数という。
(2) 確率変数 X, Y に対して、その分布関数が一致するとき、即ち $F_X(x) = F_Y(x)$ $(\forall x \in \mathbf{R})$ となるとき、X と Y は同じ分布に従うという。

$\{\omega; X(\omega) \leq a\}$ を単に $\{X \leq a\}$ と書くことが多い。以下このような略した書き方を用いる。

定理 2.6 $P(a < X \leq b) = F_X(b) - F_X(a)$, $a, b \in \mathbf{R}$ $(a < b)$

証明: $A = \{X \leq a\}$, $B = \{a < X \leq b\}$ とすると、$A \cap B = \emptyset$, $A \cup B = \{X \leq b\}$ より、
$F_X(b) = P(A \cup B) = P(A) + P(B) = F_X(a) + P(a < X \leq b)$. □

問 2.1 確率変数 X の分布関数を $F_X(x) = 0$ $(x < 0)$, $= \dfrac{1}{3}x + \dfrac{1}{6}$ $(0 \leq x < 2)$, $= 1$ $(2 \leq x)$ とする。
(1) $P(0 < X \leq 1)$, $P(0 \leq X \leq 1)$, $P(1 \leq X \leq 2)$, $P(1 \leq X < 2)$ を求めよ。
(2) $Y := X^2$ の分布関数 $F_Y(x)$ を求めよ。

定理 2.7 $F_X(x)$ を確率変数 X の分布関数とする。このとき、次が成立する。

(1) F_X は単調非減少である、即ち、$x < y$ ならば $F_X(x) \leq F_X(y)$.
(2) 右連続関数である、即ち、$F_X(a + 0) := \lim_{x \to a+0} F_X(x) = F_X(a)$ $(\forall a \in \mathbf{R})$.
(3) $F_X(-\infty) := \lim_{x \to -\infty} F_X(x) = 0$, $F_X(\infty) := \lim_{x \to \infty} F_X(x) = 1$.

証明: (1) $x < y$ より $\{X \leq x\} \subset \{X \leq y\}$. 従って、定理 1.2 (3) により $F_X(x) = P(X \leq x) \leq P(X \leq y) = F_X(y)$ を得る。
(2) $\{a_n\}$ を単調減少で a に収束する任意の数列とし、$C_n = \{X \leq a_n\}$ とおく。このとき、$\bigcap_{n=1}^{\infty} C_n = \{X \leq a\}$ となる。実際、「$\supset$」は $\forall n$ で $C_n \supset \{X \leq a\}$ だから明らか。$\omega \in \{X > a\}$ とすると $X(\omega) > a$ で $a_n \to a + 0$ $(n \to \infty)$ より、ある N があって $n \geq N \Longrightarrow a_n - a < X(\omega) - a$, これより $X(\omega) > a_N$,

即ち $\omega \notin C_N \supset \bigcap_{n=1}^{\infty} C_n$ となり「$\subset$」が従う。よって、$\{C_n\}$ は単調減少あったらから、定理 1.4 (2) より $\lim_{n\to\infty} F_X(a_n) = \lim_{n\to\infty} P(C_n) = P(X \le a) = F_X(a)$ となる。
(3) $\{x_n\}$ を単調減少で $-\infty$ に発散する任意の数列とする。このとき $B_n = \{X \le x_n\}$ とおくと、$\{B_n\}$ は単調減少な事象の列で、また X は実数値なので、$\bigcap_{n=1}^{\infty}\{X \le x_n\} = \emptyset$. 従って、定理 1.4 (2) と命題 1.2 (1) により、$\lim_{n\to\infty} F_X(x_n) = \lim_{n\to\infty} P(B_n) = P(\emptyset) = 0$ となり、前半の主張を得る。後半は、$\{x_n\}$ を単調増加で ∞ に発散する任意の数列とし、定理 1.4 (1) と $P(\Omega) = 1$ となることを用いれば、同様に証明できる (演習問題とする)。 □

注意 2.2 関数 $F : \mathbf{R} \to \mathbf{R}$ が定理 2.7 の条件 (1), (2), (3) を満たせば、$F(x)$ を分布関数としてもつ確率変数 X が (無限個) 存在することが知られている。(*cf*. 舟木著 確率論 朝倉書店 p.47, p.83.)

問 2.2 確率変数 X に対し $P(X = a) > 0$ となる a の個数は高々可算個であることを示せ。

定義 2.3 確率空間 $(\Omega, \mathcal{F}, P)$ 上の確率変数 X に対して、$P(X \in E) = 1$ となる高々可算個の元からなる集合 $E \subset \mathbf{R}$ が存在するとき、X を離散型確率変数という。(補題 2.1 より E は Borel 集合となることに注意。)

以下、確率変数 X がとりうるのは可算個の値 $x_1, x_2, \ldots$ であるとして説明する。(有限個の場合も同様である。) このとき、$p(x_i) := P(X = x_i) > 0$ $(i \in \mathbf{N})$ となる。もちろん、$\sum_{i=1}^{\infty} p(x_i) = 1$ である。この $p(x)$ を確率変数 X の確率関数という。この場合、分布関数は

$$F_X(x) = P(X \le x) = \sum_{i=1}^{\infty} p(x_i) 1_{[x_i,\infty)}(x), \qquad x \in \mathbf{R}$$

で与えられる。ここで、$1_A(x) = 1$ $(x \in A)$, $1_A(x) = 0$ $(x \notin A)$ と定義する (定義関数という)。これは、不連続点だけで増加する関数で、$p(x_i) = F_X(x_i) - F_X(x_i - 0)$ となる。

離散型分布の例

(0) 定数

$a \in \mathbf{R}$ とし、$X = a$ (定数関数) とする。このとき、$\{X \le x\} = \emptyset$ $(x < a)$, で $\{X \le x\} = \Omega$ $(x \ge a)$, となるので、これも確率変数となり、分布関数は $F_X(x) = P(X \le x) = 1_{[a,\infty)}(x)$ となる。

(a) Bernoulli 過程

歪んだコイン投げの反復試行のように、結果 S (success) の起こる確率が p, 結果 F (false) が起こる確率が $1 - p$ となる試行 (Bernoulli 試行という) を繰り返し行う。このとき、i 番目に S が起こる事象を A_i とすると、$A_1, A_2 \cdots$ は独立な事象の列となる。このとき、$X_i(\omega) = 1_{A_i}(\omega)$ (A_i が起こったとき 1, 起こらなかったとき 0 とする) とすると、$X_1, X_2, \cdots$ は同じ分布に従う (独立な) 確率変数となる (確率変数の独立の定義は 2.6 節)。このとき各 X_i の分布関数はつぎのようになり、同じ分布に従うことがわかる。

$$F_{X_i}(x) = P(X_i \le x) = (1 - p)1_{[0,\infty)}(x) + p1_{[1,\infty)}(x), \qquad x \in \mathbf{R}.$$

(b) 二項分布 $\mathrm{B}(n, p)$

(a) と同様の結果 S の起こる確率が p となる Bernoulli 試行を n 回反復するとき、結果 S が起こる回数を Y とする。このとき、Y のとり得る値は $0, 1, \ldots, n$ で

$$P(Y = k) = \binom{n}{k} p^k (1 - p)^{n-k} \qquad (k = 0, 1, \ldots, n)$$

となる。この分布を二項分布 $\mathrm{B}(n, p)$ という。また、確率変数 Y がこのような分布を持つとき Y は $\mathrm{B}(n, p)$ に従うといい $Y \sim \mathrm{B}(n, p)$ とかく。ここで実数 β と自然数 k に対して $\binom{\beta}{k} = \dfrac{\beta(\beta - 1)\cdots(\beta - k + 1)}{k!}$ と定義する。また $\binom{\beta}{0} = 1$ と定める。β が k 以上の自然数のとき $\binom{\beta}{k} = \dfrac{\beta!}{k!(\beta - k)!}$ である。

(a) で定義した Bernoulli 過程 $X_1, \cdots, X_n$ に対して、$Y = \sum_{i=1}^{n} X_i$ とすると、$Y \sim \mathrm{B}(n, p)$ となる。

問 2.3 公正なサイコロを $6n$ 回投げたとき、1 の目がちょうど n 回出る確率 p_n とする。Stirling の公式 ($cf.$ 定理 A.4) を用いて $\lim_{n\to\infty} n^\gamma p_n$ が正の数に収束するように定数 γ を定め、その極限を求めよ。

問 2.4 確率変数 X_n が二項分布 $\mathrm{B}(2n, \frac{1}{2})$ に従うとする。数列 $\{k_n\} \subset \mathbf{Z}$ が $|k_n| \le n$, $\lim_{n\to\infty} k_n/\sqrt{n} = t \in \mathbf{R}$ を満たすとき、Stirling の公式を用いて $\lim_{n\to\infty} \sqrt{n}P(X_n = n + k_n) = \frac{1}{\sqrt{\pi}}e^{-t^2}$ を示せ。

問 2.5 二項分布 $\mathrm{B}\Big(n, \dfrac{1}{7}\Big)$ について、$p(k) = \dbinom{n}{k}\Big(\dfrac{1}{7}\Big)^k\Big(1 - \dfrac{1}{7}\Big)^{n-k}$ を最大にする k を決定せよ。ヒント: $\dfrac{p(k+1)}{p(k)}$ と 1 の大小を調べよ。n についての場合分けが必要です。

(c) 幾何分布 $\mathrm{Ge}(p)$

(a) と同様の結果 S の起こる確率が p となる Bernoulli 試行の反復においてを、S が初めて出現するまでの F の出現回数を X とすると、X のとり得る値は $0, 1, 2, \ldots$ であり、この分布は幾何分布といい、

$$P(X = k) = (1-p)^k p \quad (k = 0, 1, 2, \ldots)$$

である。容易にわかるように $P(X \ge k) = \sum_{l=k}^\infty P(X = l) = (1-p)^k$ となる。また、

$$P(X = k + s | X \ge k) = P(X = s) \qquad (s, k = 0, 1, 2, \ldots)$$

となる。実際、

$$(\text{左辺}) = \frac{P(\{X = k+s\} \cap \{X \ge k\})}{P(X \ge k)} = \frac{P(X = k+s)}{P(X \ge k)} = \frac{(1-p)^{k+s}p}{(1-p)^k} = (1-p)^s p = P(X = s)$$

となるからである。

問 2.6 X が幾何分布 $\mathrm{Ge}(p)$ に従うとき、確率 $P(X$ は奇数$)$ と $Y := \sin^2 \dfrac{\pi}{6}X$ の確率分布表を求めよ。

(d) 負の二項分布 $\mathrm{NB}(\alpha, p)$, $\alpha > 0$,

(a) と同様の結果 S の起こる確率が p となる Bernoulli 試行を、S が α 回出現するまで反復するとき、F が出現する回数を Y とする。このとき、Y のとり得る値は $0, 1, \ldots$ で、$Y = k$ となるとき、$\alpha + k$ 回の試行で結果 S は最後を除いて $\alpha - 1$ 回、F は k 回出現するので、

$$P(Y = k) = \binom{\alpha + k - 1}{k} p^\alpha (1-p)^k \qquad (k = 0, 1, \ldots) \tag{2.1}$$

となる。以下自然数とは限らず一般に $\alpha > 0$ とする。ここで、

$$\begin{aligned}\binom{\alpha + k - 1}{k} &= \frac{(\alpha + k - 1)(\alpha + k - 2)\cdots(\alpha + 1)\alpha}{k!} \\ &= (-1)^k \frac{(-\alpha)(-\alpha - 1)\cdots(-\alpha - k + 1)}{k!} = (-1)^k \binom{-\alpha}{k}\end{aligned}$$

との Taylor 展開式 $(1+x)^\beta = \sum_{k=0}^\infty \dbinom{\beta}{k} x^k$ $(-1 < x < 1)$ を $\beta = -\alpha$ と用いて ($cf.$ 例 A.1 (5))

$$\sum_{k=0}^\infty \binom{-\alpha}{k} p^\alpha \{-(1-p)\}^k = p^\alpha \{1 - (1-p)\}^{-\alpha} = 1$$

となる。この分布を負の二項分布 $\mathrm{NB}(\alpha, p)$ という。この分布は α が自然数でないときも定義できる。ただし、表が出る確率が p である硬貨を n 回投げて k 回以上出ることと、k 回表が出るまでに裏が $n - k$ 回以下しか出ないことは同値なので、$X \sim \mathrm{B}(n, k)$, $Y \sim \mathrm{NB}(k, p)$ とすると

$$\sum_{l=k}^n \binom{n}{l} p^l (1-p)^{n-l} = P(X \ge k) = P(Y \le n - k) = \sum_{l=0}^{n-k} \binom{k + l - 1}{l} p^k (1-p)^l$$

この両端の式が等しいことを直接証明するのはかなり面倒である。

(e) Poisson 分布 Po(λ)

$\lambda > 0$ とする。確率変数 X が非負整数値で、その確率関数が

$$P(X=k) = \frac{\lambda^k}{k!}e^{-\lambda} \qquad (k=0,1,\ldots) \tag{2.2}$$

で与えられるとき、この確率変数 X は母数 λ の Poisson 分布 Po(λ) に従うという。$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ $(x \in \mathbf{R})$ に注意すれば、$\sum_{k=0}^{\infty} P(X=k) = 1$ がわかる。Poisson 分布は次の命題で見るように、一定時間間隔の事故の件数などを表すと考ええられる。

例題 2.1 確率変数 X_n は二項分布 B(n, p_n) に従うとする。ここで、p_n は $0 < p_n < 1$ および $\lim_{n\to\infty} np_n = \lambda > 0$ を満たすとする。このとき、$\{X_n\}$ は Poisson 分布 Po(λ) を近似している、即ち、次が成立する。

$$\lim_{n\to\infty} P(X_n = k) = \frac{\lambda^k}{k!}e^{-\lambda} \qquad (k=0,1,\ldots)$$

証明: $P(X=k) = 1\Big(1-\dfrac{1}{n}\Big)\cdots\Big(1-\dfrac{k-1}{n}\Big)\dfrac{(np_n)^k}{k!}\Big\{\Big(1-\dfrac{np_n}{n}\Big)^n\Big\}^{1-\frac{k}{n}} \to \dfrac{\lambda^k}{k!}e^{-\lambda}$ $(n\to\infty)$ □

定義 2.4 $F_X(x)$ を、確率空間 $(\Omega, \mathcal{F}, P)$ 上の確率変数 X の分布関数とする。$(\mathbf{R}, \mathcal{B}(\mathbf{R}))$ 上の非負可測関数 $f(x)$ が存在して

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt \qquad (\forall x \in \mathbf{R})$$

と表せるとき、F を絶対連続な分布関数, X を絶対連続型確率変数という。また、この $f(x)$ を X の密度関数 (density function) という。

注意 2.3 (1) X が絶対連続型なら $F_X(x)$ は連続で、x が密度関数 $f(x)$ の連続点であれば $\dfrac{dF_X}{dx}(x) = f(x)$ となる。また、$f(x)$ は

$$f(x) \geq 0, \qquad \int_{-\infty}^{\infty} f(x)\,dx = 1$$

を満たし、$P(X=a) = 0$ $(\forall a \in \mathbf{R})$ であり、

$$P(a < X \leq b) = P(a \leq X \leq b) = \int_a^b f(t)\,dt \qquad (a < b)$$

となる。

(2) $F_X(x)$ が連続であっても、密度関数 $f(x)$ が存在するとは限らない ($cf.$ A4 Cantor 関数)。

絶対連続型分布の例

(a) 一様分布 U(a,b)

確率変数 X の密度関数 f_X あるいは分布関数 F_X が

$$f_X(x) = \begin{cases} \frac{1}{b-a} & (a < x < b) \\ 0 & (x \leq a \text{ または } b \leq x) \end{cases} \qquad F_X(x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & (x \leq a) \\ \frac{x-a}{b-a} & (a < x \leq b) \\ 1 & (b < x) \end{cases}$$

で与えられるとき、確率変数 X は区間 (a,b) 上の一様分布 U(a,b) に従うという。

(b) 指数分布 Ex(λ) $(\lambda > 0)$

確率変数 X の密度関数 $f_X(x)$ あるいは分布関数 $F_X(x)$ が次で与えられるとき、この分布を指数分布 Ex(λ) という。

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases} \qquad F_X(x) = \begin{cases} 1 - e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

指数分布は事故などの Poisson 事象が生起する時間間隔の分布として広く用いられている。

(c) 正規分布 $\mathrm{N}(\mu,\sigma^2)$ ($\mu\in\mathbf{R}$, $\sigma>0$)
次の密度関数をもつ分布を正規分布 (normal distribution) という。

$$f_X(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{2.3}$$

この正規分布を記号 $\mathrm{N}(\mu,\sigma^2)$ で表す。

正規分布 $\mathrm{N}(\mu,\sigma^2)$ において、特に $\mu=0,\sigma=1$ のとき、この正規分布 $\mathrm{N}(0,1)$ を標準正規分布という。この密度関数と分布関数をそれぞれ $\phi(x)$, $\Phi(x)$ とすると、

$$\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}},\qquad \Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\,dt.$$

例題 2.2 $X\sim\mathrm{N}(\mu,\sigma^2)$ のとき、定数 $a\neq0,b$ に対して $aX+b\sim\mathrm{N}(a\mu+b,a^2\sigma^2)$. 特に、$Z=\frac{X-\mu}{\sigma}\sim\mathrm{N}(0,1)$.

証明: $F_{aX+b}(y)=P(aX+b\le y)=P(X\le\frac{y-b}{a})=F_X(\frac{y-b}{a})$ より
$f_{aX+b}(y)=\frac{d}{dy}\{F_X(\frac{y-b}{a})\}=f_X(\frac{y-b}{a})\frac{1}{a}=\frac{1}{\sqrt{2\pi\sigma^2a^2}}\exp\{-\frac{(y-b-a\mu)^2}{2a^2}\}$. □

(d) ガンマ分布 $\Gamma(\alpha,\beta)$ ($\alpha>0$, $\beta>0$)
確率変数 X の密度関数が次の $f_X(x)$ のとき、X はガンマ分布 $\Gamma(\alpha,\beta)$ に従うといい、$X\sim\Gamma(\alpha,\beta)$ と表す。

$$f_X(x)=\begin{cases}\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x} & (x>0)\\ 0 & (x\le 0)\end{cases} \tag{2.4}$$

ここで、$\Gamma(\alpha)=\displaystyle\int_0^\infty x^{\alpha-1}e^{-x}\,dx$ はガンマ関数を表す。また、ガンマ分布 $\Gamma(1,\beta)$ と指数分布 $\mathrm{Ex}(\beta)$ は同じ分布である。

(e) ベータ分布 $\mathrm{B{\small ETA}}(a,b)$ ($a>0$,$b>0$)
確率変数 X の密度関数が次のとき、X はベータ分布 $\mathrm{B{\small ETA}}(a,b)$ に従うといい、$X\sim\mathrm{B{\small ETA}}(a,b)$ と表す。

$$f_X(x)=\begin{cases}\frac{1}{\mathrm{B}(a,b)}x^{a-1}(1-x)^{b-1} & (0<x<1)\\ 0 & (\text{その他})\end{cases} \tag{2.5}$$

ここで、$\mathrm{B}(a,b)=\displaystyle\int_0^1 x^{a-1}(1-x)^{b-1}\,dx$ はベータ関数を表す。また、ベータ分布 $\mathrm{B{\small ETA}}(1,1)$ と一様分布 $\mathrm{U}(0,1)$ は同じ分布である。

補題 2.3 (ベータ関数 $\mathrm{B}(p,q)$、ガンマ関数 $\Gamma(p)$ に関する公式の復習) 以下が成り立つ。

(1) $\Gamma(1)=1$, $\Gamma(\frac{1}{2})=\sqrt{\pi}$.
(2) $\Gamma(p+1)=p\Gamma(p)$ ($p>0$) 特に、自然数 n に対して $\Gamma(n)=(n-1)!$ となる。
(3) $\mathrm{B}(p,q)=\dfrac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$ ($p,q>0$).

証明 (1) $\Gamma(1)=\displaystyle\int_0^\infty e^{-x}\,dx=\left[-e^{-x}\right]_0^\infty=1$,

$$\Gamma\Big(\frac{1}{2}\Big)=\int_0^\infty x^{-\frac{1}{2}}e^{-x}\,dx=\int_0^\infty\Big(\frac{t^2}{2}\Big)^{-\frac{1}{2}}e^{-\frac{1}{2}t^2}t\,dt=\sqrt{2}\int_0^\infty e^{-\frac{1}{2}t^2}\,dt=\sqrt{2}\frac{\sqrt{2\pi}}{2}=\sqrt{\pi}.\ (x=\frac{1}{2}t^2)$$

(2) $\Gamma(p+1)=\displaystyle\int_0^\infty x^p(-e^{-x})'\,dx=\left[-x^pe^{-x}\right]_0^\infty+\int_0^\infty px^{p-1}(-e^{-x})'\,dx=p\Gamma(p)$.

(3) まず、D : $x>0,y>0$ とすると、$\Gamma(p)\Gamma(q)=\displaystyle\iint_D e^{-x-y}x^{p-1}y^{q-1}\,dxdy$ となる。ここで、$n\ge2$ に対して $K_n:\dfrac{1}{n}\le x+y\le n,\dfrac{1}{n-1}x\le y\le(n-1)x$ とすると、$\{K_n\}$ は D に収束する増大列である。このと

き、$x = uv,\ y = u(1-v)$ と変換すると $u = x+y,\ v = \dfrac{x}{x+y}$ より、K_n は $F_n : \dfrac{1}{n} \le u \le n, \dfrac{1}{n} \le v \le 1 - \dfrac{1}{n}$ に対応し、ヤコビアンは $\dfrac{\partial(x,y)}{\partial(u,v)} = -u$ となる。よって、

$$\begin{aligned}\iint_{K_n} e^{-x-y} x^{p-1} y^{q-1}\, dxdy &= \iint_{F_n} e^{-uv-u(1-v)} (uv)^{p-1} \{u(1-v)\}^{q-1} |-u|\, dudv \\ &= \int_{\frac{1}{n}}^{n} u^{p+q-1} e^{-u}\, du \int_{\frac{1}{n}}^{1-\frac{1}{n}} v^{p-1}(1-v)^{q-1}\, dv\end{aligned}$$

となり、$n \to \infty$ として $\Gamma(p)\Gamma(q) = \Gamma(p+q)\mathrm{B}(p,q)$ を得る。 □

問 2.7 (1) $X \sim \mathrm{N}(0,1)$ のとき $Y = X^2$ の密度関数 $f_Y(y)$ を求め $Y \sim \Gamma(1/2, 1/2)$ を示せ。
(2) $X \sim \mathrm{N}(\mu, \sigma^2)$ のとき、$Y = e^X$ の密度関数 $f_Y(y)$ を求めよ。Y の分布を対数正規分布という。
(3) $U \sim \mathrm{U}(0,1)$ $X = -\lambda \log U$ $(\lambda > 0)$ の密度関数 $f_X(x)$ を求め、$X \sim \mathrm{Ex}(1/\lambda)$ を示せ。
(4) $U \sim \mathrm{U}(0,1)$ のとき $X = \cos \pi U$ と $Y = \sin \pi U$ の密度関数 $f_X(x), f_Y(y)$ を求めよ。
(5) $X \sim \textsc{Beta}(a,b)$ のとき、$Y = \dfrac{X}{1-X}$ の密度関数 $f_Y(y)$ を求めよ。Y の分布を第 2 種ベータ分布という。

問 2.8 $\lambda > 0,\ n \in \mathbf{N}$ とする。X がガンマ分布 $\Gamma(n,1)$ に Y が Poisson 分布 $\mathrm{Po}(\mu)$ に従うとき、$P(X \le \mu) = P(Y \ge n)$ を示せ。(この結果は Poisson 母集団の母平均 μ の区間推定や検定に応用される。)

2.4 多次元確率変数

定義 2.5 $X_1(\omega), \ldots, X_n(\omega)$ を確率空間 $(\Omega, \mathcal{F}, P)$ 上の確率変数とするとき、

$$\boldsymbol{X}(\omega) = (X_1(\omega), \ldots, X_n(\omega)) \tag{2.6}$$

を n 次元確率変数または n 次元確率ベクトルという。ここに、$\boldsymbol{X} = (X_1, \ldots, X_n)$ とも略記される。

定義 2.6 任意の $(x_1, \ldots, x_n) \in \mathbf{R}^n$ に関し、(2.6) の多次元確率変数 $\boldsymbol{X} = (X_1, \ldots, X_n)$ の同時分布関数を次式で定義する。

$$F_{\boldsymbol{X}}(x_1, x_2, \ldots, x_n) = P(X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n).$$

注意 2.4 $P(X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n)$ は $P(\{X_1 \le x_1\} \cap \{X_2 \le x_2\} \cap \cdots \cap \{X_n \le x_n\})$ を表す。今後もこのように “$\cap$” を略した記号を用いる。

n 次元 Borel 集合族を問 1.5 を用いて次のように定義する。

定義 2.7 $\mathbf{R}^n$ のすべての半開区間 $(a_1, b_1] \times (a_2, b_2] \times \cdots \times (a_n, b_n],\ a_i < b_i,\ i = 1, 2, \cdots, n,$ を含む最小の σ 集合族を、n 次元 Borel 集合族といい、$\mathcal{B}(\mathbf{R}^n)$ と表す。$\mathcal{B}(\mathbf{R}^n)$ の元を Borel 集合という。また、関数 $\varphi : \mathbf{R}^n \to \mathbf{R}$ が任意の $B \in \mathcal{B}(\mathbf{R})$ に対して、$\varphi^{-1}(B) \in \mathcal{B}(\mathbf{R}^n)$ を満たすとき、φ は $(\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n))$ 上の可測関数、あるいは、単に Borel 可測であるという。

原理的には、任意の $B \in \mathcal{B}(\mathbf{R})$ に対して $P((X_1, \ldots, X_n) \in B)$ の値は $F_{\boldsymbol{X}}(x_1, x_2, \ldots, x_n)$ から定まる。例えば、2 次元の場合、$B = (a_1, b_1] \times (a_2, b_2]$ の場合、

$$\begin{aligned}&P(a_1 < X_1 \le b_1, a_2 < X_2 \le b_2) = P(a_1 < X_1 \le b_1, X_2 \le b_2) - P(a_1 < X_1 \le b_1, X_2 \le a_2) \\ &\quad = P(X_1 \le b_1, X_2 \le b_2) - P(X_1 \le a_1, X_2 \le b_2) - \{P(X_1 \le b_1, X_2 \le a_2) - P(X_1 \le a_1, X_2 \le a_2)\} \\ &\quad = F_{\boldsymbol{X}}(b_1, b_2) - F_{\boldsymbol{X}}(a_1, b_2) - F_{\boldsymbol{X}}(b_1, a_2) + F_{\boldsymbol{X}}(a_1, a_2)\end{aligned}$$

を得る。

次の定理は定理 2.5 の拡張である。証明は省略する。

定理 2.8 $X_1, \ldots, X_n$ が $(\Omega, \mathcal{F})$ 上の確率変数、$f : \mathbf{R}^n \to \mathbf{R}$ が $(\mathbf{R}^n, \mathcal{B}(\mathbf{R}^n))$ 上の可測関数のとき、$f(X_1(\omega), \cdots, X_n(\omega))$ は $(\Omega, \mathcal{F})$ 上の確率変数である。

定理 2.7 と同様に次が成り立つ。証明は省略する。

命題 2.1 (1) 各変数 x_j ごとに分布関数 $F_{\boldsymbol{X}}(x_1, x_2, \ldots, x_n)$ は非減少な右連続関数である。
(2) $F_{\boldsymbol{X}}(x_1, \ldots, -\infty, \ldots, x_n) := \lim_{x_j \to -\infty} F_{\boldsymbol{X}}(x_1, \ldots, x_j, \ldots, x_n) = 0.$
また、$F_{\boldsymbol{X}}(\infty, \ldots, \infty) := \lim_{x_1, \cdots, x_n \to \infty} F_{\boldsymbol{X}}(x_1, \ldots, x_n) = 1.$

多次元確率変数 $\boldsymbol{X} = (X_1, \ldots, X_n)$ において、

$$P(X_j \leq x_j) = F_{\boldsymbol{X}}(\infty, \cdots, \infty, x_j, \infty, \cdots, \infty)$$

を X_j の 1 次元周辺分布という。同様に、$1 < m < n$ に対して m 次元周辺分布も定義される。

一次元の場合と同様に離散分布, 絶対連続分布を以下のように定義する。

(1) 高々可算個の点からなる集合 $D \subset \mathbf{R}^n$ があって、$P((X_1, \ldots, X_n) \in D) = 1$ となる場合、この多次元確率変数を離散型という。

(2) $\mathbf{R}^n$ 上の可測関数 $f(x_1, \ldots, x_n)$ が存在して

$$F_{\boldsymbol{X}}(x_1, \ldots, x_n) = \int_{-\infty}^{x_n}\int_{-\infty}^{x_{n-1}} \cdots \int_{-\infty}^{x_1} f(t_1, t_2, \ldots, t_n)\, dt_1 \cdots dt_{n-1} dt_n$$

と表せるとき、F を絶対連続な分布関数, $\boldsymbol{X} = (X_1, \ldots, X_n)$ を絶対連続型確率変数という。また、この $f(x_1, \ldots, x_n)$ を $\boldsymbol{X}$ の同時密度関数という。

絶対連続型確率変数 $(X_1, \ldots, X_n)$ に対し、X_i の周辺分布は絶対連続型となり、その密度関数は他の座標をすべて $\mathbf{R}$ で積分してしまうことで得られる。実際、$j = 1$ のとき証明すると

$$P(X_1 \leq x) = P(X \leq x_1, X_2 \in \mathbf{R}, \ldots, X_n \in \mathbf{R}) = \int_{-\infty}^{x_1}\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(t_1, t_2, \ldots, t_n)\, dt_n \cdots dt_2 dt_1$$

より、X_1 の周辺密度関数 f_{X_1} は

$$f_{X_1}(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x, t_2, \ldots, t_n)\, dt_n \cdots dt_2 \tag{2.7}$$

となることがわかる。また、$B \subset \mathbf{R}^n$ が Borel 集合のとき、

$$P((X_1, \ldots, X_n) \in B) = \int\int \cdots \int_B f(t_1, t_2, \ldots, t_n)\, dt_1 dt_2 \cdots dt_n \tag{2.8}$$

となることが、(後期に学ぶ Dynkin 族定理を用いて) 証明できる。

例題 2.3 (X, Y) の同時密度関数を $f(x, y) = \begin{cases} Kxy & (0 \leq x, 0 \leq y, x^2 + y^2 \leq 1) \\ 0 & (\text{その他}) \end{cases}$ とする。このとき定数 K を定め、X の周辺密度関数 $f_X(x)$ を求めよ。また、$T = X + Y$ の分布関数 $F_T(t)$ と密度関数 $f_T(t)$ を求めよ。

解 $1 = \iint_{\mathbf{R}^2} f(x, y)\, dxdy = \int_0^1 dx \int_0^{\sqrt{1-x^2}} Kxy\, dy = \cdots = \dfrac{K}{8}$ より $K = 8$.
X の周辺密度関数は $f_X(x) = \displaystyle\int_{-\infty}^{\infty} f(x, y)\, dy$ より、$x \notin (0, 1)$ のとき明らかに $f_X(x) = 0$. $0 < x < 1$ のとき、

$$f_X(x) = \int_0^{\sqrt{1-x^2}} 8xy\, dy = 4x(1 - x^2) \quad \text{より} \quad f_X(x) = \begin{cases} 4x(1 - x^2) & (0 < x < 1) \\ 0 & (\text{その他}) \end{cases} \quad \text{となる。}$$

T の分布について、$t<0$ のとき $F_T(t)=0$, $\sqrt{2}\le t$ のとき $F_T(t)=1$ は明らか。$0\le t\le 1$ のとき、

$$F_T(t)=P(X+Y\le t)=\int_0^t dx\int_0^{x-t}8xy\,dy=\cdots=\frac{1}{3}t^4.$$

$1\le t<\sqrt{2}$ のとき、$x+y=t$, $x^2+y^2=1$ の交点の x 座標を $\alpha,\beta\ (\alpha<\beta)$ とすると、$\alpha+\beta=t$, $\alpha\beta=\dfrac{t^2-1}{2}$.
よって、$(\beta-\alpha)^2=2-t^2$ となることに注意する。よって、

$$\begin{aligned}1-F_T(t)=P(X+Y>t)&=\iint_{t-x\le y\le\sqrt{1-x^2}}8xy\,dxdy=\int_\alpha^\beta dx\int_{t-x}^{\sqrt{1-x^2}}8xy\,dy\\&=\int_\alpha^\beta\left\{4x(1-x^2)-4x(t-x)^2\right\}dx=\left[\frac{8}{3}tx^3-2x^4-2(t^2-1)x^2\right]_\alpha^\beta\\&=(\beta-\alpha)\left\{\frac{8}{3}t\left((\alpha+\beta)^2-\alpha\beta\right)-2(\alpha+\beta)(\alpha^2+\beta^2)-2(t^2-1)(\alpha+\beta)\right\}\\&=\sqrt{2-t^2}\left\{\frac{4}{3}t\left(t^2+1\right)-2t(t^2-t^2+1)-2(t^2-1)t\right\}=\cdots=\frac{2}{3}t(2-t^2)^{3/2}.\end{aligned}$$

以上をまとめて、

$$F_T(t)=\begin{cases}0&(t<0)\\\dfrac{1}{3}t^4&(0\le t<1)\\1-\dfrac{2}{3}t(2-t^2)^{3/2}&(1\le t<\sqrt{2})\\1&(\sqrt{2}\le t)\end{cases},\quad f_T(t)=\frac{d}{dt}F_T(t)=\begin{cases}0&(t<0)\\\dfrac{4}{3}t^3&(0\le t<1)\\\dfrac{4}{3}(2t^2-1)\sqrt{2-t^2}&(1\le t<\sqrt{2})\\0&(\sqrt{2}\le t)\end{cases}\quad\square$$

問 2.9 (1) (X,Y) の同時密度関数を $f(x,y)=\begin{cases}\dfrac{16}{\pi}y^2&(0\le x,0\le y,x^2+y^2\le1)\\0&(\text{その他})\end{cases}$ とする。X の周辺密度関数 $f_X(x)$ を求めよ。また、$P(Y\le tX)$, $t>0$, を求め、$T=Y/X$ の周辺密度関数 $f_T(t)$ を求めよ。
(2) (X,Y) の同時密度関数を $f(x,y)=\begin{cases}K(x+y)&(0<y\le x\le2)\\0&(\text{その他})\end{cases}$ とする。定数 K を定め、X の周辺密度関数 $f_X(x)$ を求めよ。また、$P(XY\ge s)$, $0<s<4$, を求め、$S=XY$ の密度関数 $f_S(s)$ を求めよ。
(3) (X,Y) の同時密度関数を $f(x,y)=\begin{cases}K(x+y)&(0\le x,0\le y,x+y\le1)\\0&(\text{その他})\end{cases}$ とする。定数 K を定め、Y の周辺密度関数 $f_Y(y)$ を求めよ。また、$P(Y\le 2X^2)$ を求めよ。

問 2.10 $X\sim\mathrm{Ex}(\lambda)$ とし、$Y=\lfloor X\rfloor$, $Z=X-\lfloor X\rfloor$ とする。ただし、$\lfloor a\rfloor$ は a 以下の最大の整数とする。このとき、$P(Y=k)$, $k=0,1,\ldots$, および Z の分布関数 $F_Z(z)$ と密度関数 $f_Z(z)$ を求めよ。

例 2.3 よく知られている多次元確率変数とその分布を例示する。

(i) 多項分布
$\{A_1,A_2,\ldots,A_L\}\subset\mathcal{F}$ を Ω の分割: $\Omega=\bigcup_{j=1}^L A_j$, $A_i\cap A_j=\emptyset\ (i\ne j)$ とする。また、1 回の試行で事象 A_j が出現する確率を $p_j>0$, $p_1+p_2+\ldots+p_L=1$, とし、この試行を n 回反復して各事象 A_j の出現回数を X_j $(j=1,\ldots,L)$ とすれば、$(X_1,\ldots,X_L)$ の同時分布は次式で示される。

$$P(X_1=k_1,X_2=k_2,\cdots,X_L=k_L)=\frac{n!}{k_1!k_2!\cdots k_L!}p_1^{k_1}p_2^{k_2}\cdots p_L^{k_L}$$

ただし、$k_1+k_2+\cdots+k_L=n$ である。このような分布を多項分布という。
この多項分布に対し X_1 の 1 次元周辺分布が二項分布 $\mathrm{B}(n,p_1)$ であることを示そう。
簡単のため $L=4$ として示す。一般の場合はそれを繰り返せばよい。$X_1=k_1,X_2=k_2$ とすると $X_3+X_4=$

$n_3 := n - k_1 - k_2$ であるから、

$$P(X_1 = k_1, X_2 = k_2) = \sum_{k_3=0}^{n_3} P(X_1 = k_1, X_2 = k_2, X_3 = k_3, X_4 = n_3 - k_3)$$
$$= \frac{n!}{k_1!k_2!n_3!} p_1^{k_1} p_2^{k_2} \sum_{k_3=0}^{n_3} \frac{n_3!}{k_3!(n_3-k_3)!} p_3^{k_3} p_4^{n_3-k_3} = \frac{n!}{k_1!k_2!n_3!} p_1^{k_1} p_2^{k_2} (p_3+p_4)^{n_3}$$

最後の等式は二項定理を用いた。特に $(X_1, X_2, X_3 + X_4)$ も多項分布となる。

$$P(X_1 = k_1) = \sum_{k_2=0}^{n-k_1} P(X_1 = k_1, X_2 = k_2) = \frac{n!}{k_1!(n-k_2)!} p_1^{k_1} \sum_{k_2=0}^{n-k_1} \frac{(n-k_1)!}{k_2!n_3!} p_2^{k_2} (p_3+p_4)^{n_3}$$
$$= \frac{n!}{k_1!(n-k_1)!} p_1^{k_1} (p_2+p_3+p_4)^{n-k_1} = \binom{n}{k_1} p_1^{k_1} (1-p_1)^{n-k_1}$$

3 つ目の等号は $n_3 = n - k_1 - k_2$, 最後の等号は $p_1 + p_2 + p_3 + p_4 = 1$ を用いた。 □

(ii) 多次元正規分布 $\mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)'$ と n 次正定値対称行列 $\boldsymbol{\Sigma} = (\sigma_{ij})$ に対して、n 次元確率変数 $\boldsymbol{X} = (X_1, \ldots, X_n)'$ が次の密度関数をもつとき、この分布を n 次元正規分布 $\mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ という。

$$f(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}(\det \boldsymbol{\Sigma})^{1/2}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}, \qquad \boldsymbol{x} = (x_1, \ldots, x_n)'$$

ここに、A' は行列 A の転置を表す。例えば $\boldsymbol{\mu}, \boldsymbol{x}$ は縦ベクトルとなる。

これが $\mathbf{R}^n$ 上で積分すれば 1 となることは次のように示すことが出来る。

$\boldsymbol{\Sigma}$ は正定値対称行列なので、線形代数の対角化に関する定理から、直交行列 P と対角成分がすべて正の対角行列 $D = (d_{ij})$ がとれて、$P'\boldsymbol{\Sigma}P = D$ とできる。ここで、$\boldsymbol{y} = (y_1, \ldots, y_n)' = P(\boldsymbol{x} - \boldsymbol{\mu})$ とすると、$D^{-1} = (P'\boldsymbol{\Sigma}P)^{-1} = P'\boldsymbol{\Sigma}^{-1}P$ より

$$(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = \boldsymbol{y}'P'\boldsymbol{\Sigma}^{-1}P\boldsymbol{y} = \boldsymbol{y}'D^{-1}\boldsymbol{y} = \sum_{j=1}^{n} \frac{1}{d_{jj}} y_j^2$$

で、ヤコビアンは $\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}} = \det\left(\left(\frac{\partial y_i}{\partial x_j}\right)_{ij}\right) = \det P = \pm 1$, $\det \boldsymbol{\Sigma} = \det D = \prod_{j=1}^{n} d_{jj}$ なので、

$$\int \cdots \int_{\mathbf{R}^n} f(x_1, \ldots, x_n)\, dx_1 \cdots dx_n = \prod_{j=1}^{n} \int_{-\infty}^{\infty} \frac{1}{(2\pi d_{jj})^{1/2}} e^{-\frac{1}{2}\frac{y_j^2}{d_{jj}}}\, dy_j = 1$$

となる。 □

2.5 条件つき確率分布

この節では確率変数で条件付けをした分布について、離散型確率変数の場合と絶対連続型確率変数の場合に定義する。一般の確率変数について条件付け確率分布を定義するためには、Lebesgue 積分論の Radon-Nikodym の定理を必要とするのでこの授業では取り扱わない。

ここでは、簡単のため (X, Y) を 2 次元確率変数として、条件つき分布 $P(Y \in A | X = x)$ を定義する。

(i) X の周辺分布が離散型の場合

X のとり得る値が $\{x_1, x_2, \ldots\}$ とすると、$P(X = x_j) > 0\ (\forall j)$ となることに注意する。この場合、

$$P(Y \in A | X = x_j) = \frac{P(X = x_j, Y \in A)}{P(X = x_j)}$$

と定め、これを $X = x_j$ の条件下での Y の条件つき確率分布という。ただし、$P(X = x) = 0$ となる x では定義しないものとする。

(ii) X の周辺分布が絶対連続型の場合
この場合 $P(X = x) = 0$ $(\forall x)$ であるから工夫が必要となる。

X の周辺密度関数を $f_X(x)$ とする。ここでは、簡単のため x は $f_X(x)$ の連続点で $f_X(x) > 0$ となるときに考える。$P(Y < y|X = x)$ を次のように定義する。

$$P(Y \in A|X = x) = \lim_{\delta \to +0} P(Y \in A|x \leq X < x + \delta)$$

と定める。このとき、$\xi_k = a + \frac{k}{n}(b - a)$ とすることで、

$$\begin{aligned}
\int_a^b P(Y \in A|X = x) f_X(x)\, dx &= \lim_{n\to\infty} \sum_{k=1}^{n} \int_{\xi_{k-1}}^{\xi_k} P(Y \in A|X = x) f_X(x)\, dx \\
= \lim_{n\to\infty} \sum_{k=1}^{n} P(Y \in A|\xi_{k-1} \leq X < \xi_k) &\int_{\xi_{k-1}}^{\xi_k} f_X(x)\, dx = P(Y \in A, a \leq X < b)
\end{aligned} \tag{2.9}$$

となることに注意する。次に、(X, Y) は絶対連続型で $f(x, y)$ をその同時密度関数とする。このとき、

$$P(Y \leq y|x \leq X < x + \delta) = \frac{P(Y \leq y, x \leq X < x + \delta)}{P(x \leq X < x + \delta)} = \frac{\int_x^{x+\delta}\int_{-\infty}^{y} f(s, t)\, dtds}{\int_x^{x+\delta} f_X(s)ds}$$

ここで、分母分子を δ で割り $\delta \to +0$ とすることで、微分積分学の基本定理により

$$P(Y \leq y|X = x) = \frac{\int_{-\infty}^{y} f(x, t)\, dt}{f_X(x)}$$

を得る。これより、条件 $X = x$ の下での Y の分布は絶対連続型で、その密度関数 $f_{Y|X}(y|x)$ が

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} \tag{2.10}$$

となることがわかった。$f_{Y|X}(y|x)$ を条件 $X = x$ の下での Y の条件つき密度関数という。 □

例題 2.4 (X, Y) を例題 2.3 のそれとする。$0 < x < 1$ のとき $f_{Y|X}(y|x)$ を求め、$P(Y > \frac{1}{2}|X = \frac{1}{2})$ を計算せよ。

解 例題 2.3 より $0 < x < 1$ のとき $f_X(x) = 4x(1 - x^2)$ であったから、

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} = \begin{cases} \dfrac{2y}{1 - x^2} & (0 \leq y \leq \sqrt{1 - x^2}) \\ 0 & (\text{その他}) \end{cases}$$

また、$P\Big(Y > \dfrac{1}{2}\Big|X = \dfrac{1}{2}\Big) = \displaystyle\int_{\frac{1}{2}}^{\infty} f_{Y|X}\Big(y\Big|\frac{1}{2}\Big)\, dy = \int_{\frac{1}{2}}^{\sqrt{1-(\frac{1}{2})^2}} \frac{2y}{1 - (\frac{1}{2})^2}\, dy = \frac{2}{3}.$ □

問 2.11 (1) (X, Y) を問 2.9 (1) のそれとする。$0 < x < 1$ のとき $f_{Y|X}(y|x)$ を求め、$P\Big(Y > \dfrac{1}{2}\Big|X = \dfrac{1}{2}\Big)$ を計算せよ。
(2) (X, Y) を問 2.9 (2) のそれとする。$0 < x < 2$ のとき $f_{Y|X}(y|x)$ を求め、$P\Big(Y > \dfrac{1}{2}\Big|X = \dfrac{3}{2}\Big)$ を計算せよ。
(3) (X, Y) の同時密度関数を $f(x, y) = \begin{cases} Kxy & (x^2 \leq y \leq 2x) \\ 0 & (\text{その他}) \end{cases}$ とする。定数 K を定め、X の周辺密度関数 $f_X(x)$ を求めよ。また、$0 < x < 2$ に対して $f_{Y|X}(y|x)$ を求め、$P\Big(Y \leq \dfrac{3}{2}\Big|X = 1\Big)$ を計算せよ。

例 2.4 (X,Y) を 2 次元正規分布 $\mathrm{N}(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \boldsymbol{\Sigma}=\begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{12} & \sigma_{22}\end{pmatrix})$ に従うとする。ただし、$\boldsymbol{\Sigma}$ は正定値であったから、$\det(\lambda I-\boldsymbol{\Sigma})=\lambda^2-(\sigma_{11}+\sigma_{22})\lambda+\sigma_{11}\sigma_{22}-\sigma_{12}^2=0$ の根が 2 つとも正となるため、$\sigma_{11},\sigma_{22}>0$, $\sigma_{11}\sigma_{22}-\sigma_{12}^2>0$ を得る。これより、$\sigma_1,\sigma_2>0$, $|\rho|<1$ を用いて、$\boldsymbol{\Sigma}=\begin{pmatrix}\sigma_1^2 & \sigma_1\sigma_2\rho\\ \sigma_1\sigma_2\rho & \sigma_2^2\end{pmatrix}$ と表せる。このとき、$\det\boldsymbol{\Sigma}=\sigma_1^2\sigma_2^2(1-\rho^2)$ で $\boldsymbol{\Sigma}^{-1}=\dfrac{1}{1-\rho^2}\begin{pmatrix}1/\sigma_1^2 & -\rho/(\sigma_1\sigma_2)\\ -\rho/(\sigma_1\sigma_2) & 1/\sigma_2^2\end{pmatrix}$ であるから、(X,Y) の密度関数は

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\Big[-\frac{1}{2(1-\rho^2)}\Big\{\frac{(x-\mu_1)^2}{\sigma_1^2}-\frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\Big\}\Big] \tag{2.11}$$

となる。

これを踏まえ $f_{Y|X}(y|x)$ を求め、$X=x$ の条件のもと Y の分布を調べよう。まず、X の周辺密度関数 $f_X(x)$ を求める。(2.11) の右辺の $\{\cdots\}$ の中を y について平方完成すると

$$\frac{1}{\sigma_2^2}\Big[(y-\mu_2)-\frac{\rho\sigma_2}{\sigma_1}(x-\mu_1)\Big]^2+\frac{1-\rho^2}{\sigma_1^2}(x-\mu_1)^2$$

となるから、$t=\frac{1}{\sigma_2\sqrt{1-\rho^2}}\big[y-\mu_2-\frac{\rho\sigma_2}{\sigma_1}(x-\mu_1)\big]$ とおくと、

$$f_X(x)=\int_{\mathbf{R}}f(x,y)\,dy=\frac{1}{2\pi\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}\,dt=\frac{1}{\sqrt{2\pi}\sigma_1}e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}$$

となる。(これより X の周辺分布は $\mathrm{N}(\mu_1,\sigma_1^2)$ であることがわかった。) よって、

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\frac{1}{\sqrt{2\pi}\sigma_2\sqrt{1-\rho^2}}\exp\Big[-\frac{(y-\{\mu_2+\frac{\rho\sigma_2}{\sigma_1}(x-\mu_1)\})^2}{2\sigma_2^2(1-\rho^2)}\Big]$$

を得る。これは、$X=x$ の条件下の Y の分布が正規分布 $\mathrm{N}(\mu_2+\frac{\rho\sigma_2}{\sigma_1}(x-\mu_1),\sigma_2^2(1-\rho^2))$ となることを示している。 □

注意 2.5 (Bayes の定理) (2.10) は $f(x,y)=f_{X|Y}(x|y)f_Y(y)$ に注意して、次のように表せる。

$$f_{Y|X}(y|x)=\frac{f(x,y)}{f_X(x)}=\frac{f(x,y)}{\int_{-\infty}^{\infty}f(x,u)\,du}=\frac{f_{X|Y}(x|y)f_Y(y)}{\int_{-\infty}^{\infty}f_{X|Y}(x|u)f_Y(u)\,du}. \tag{2.12}$$

一方、X が離散型であれば $P(X=x)>0$ となる x に対して

$$P(Y\le y|X=x)=\frac{P(Y\le y,X=x)}{P(X=x)}=\frac{1}{P(X=x)}\int_{-\infty}^{y}P(X=x|Y=u)f_Y(u)\,du$$

より、$X=x$ のもとでの Y の条件つき密度関数は

$$f_{Y|X}(y|x)=\frac{d}{dy}P(Y\le y|X=x)=\frac{P(X=x|Y=y)f_Y(y)}{P(X=x)} \tag{2.13}$$

を得る。これらは Bayes の定理の一般化で機械学習のある分野で用いられる。

例題 2.5 (1) Λ は $\Gamma(\alpha,\beta)$ に従い、$\Lambda=\lambda$ の条件のもと X が $\mathrm{Po}(\lambda)$ に従うとき、$k=0,1,\ldots$ に対して $P(X=k)$ および $f_{\Lambda|X}(\lambda|k)$ を求めよ。
(2) Λ は $\Gamma(\alpha,\beta)$ に従い、$\Lambda=\lambda$ の条件のもと Y が $\mathrm{N}(0,1/\lambda)$ に従うとき、$f_Y(y)$ 及び $f_{\Lambda|Y}(\lambda|y)$ を求めよ。

解 (1) (2.9) を $A=\{k\}$, $(a,b)=(-\infty,\infty)$ として用いると

$$P(X=k)=\int_{-\infty}^{\infty}P(X=k|\Lambda=\lambda)f_\Lambda(\lambda)\,d\lambda=\int_0^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}\frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}\,d\lambda$$

$$
\begin{aligned}
&= \frac{\beta^\alpha}{k!\Gamma(\alpha)} \int_0^\infty \lambda^{k+\alpha-1} e^{-(\beta+1)\lambda}\, d\lambda = \frac{\beta^\alpha}{k!\Gamma(\alpha)} \frac{\Gamma(\alpha+k)}{(\beta+1)^{\alpha+k}} \\
&= \frac{(\alpha+k-1)(\alpha+k-2)\cdots\alpha\Gamma(\alpha)}{k!\Gamma(\alpha)} \frac{\beta^\alpha}{(\beta+1)^{\alpha+k}} = \binom{\alpha+k-1}{k} \Big(\frac{\beta}{\beta+1}\Big)^\alpha \Big(\frac{1}{\beta+1}\Big)^k.
\end{aligned}
$$

3 行目の最初の等号は補題 2.3(2) を用いた。よって、$X \sim \mathrm{NB}\Big(\alpha, \dfrac{\beta}{\beta+1}\Big)$ となる。次に (2.13) より

$$
f_{\Lambda|X}(\lambda|k) = \frac{P(X=k|\Lambda=\lambda) f_\Lambda(\lambda)}{P(X=k)} = \frac{\dfrac{\lambda^k}{k!} e^{-\lambda} \dfrac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}}{\dfrac{\beta^\alpha}{k!\Gamma(\alpha)} \dfrac{\Gamma(\alpha+k)}{(\beta+1)^{\alpha+k}}} = \frac{(\beta+1)^{\alpha+k}}{\Gamma(\alpha+k)} \lambda^{\alpha+k-1} e^{-(\beta+1)\lambda}, \quad \lambda > 0.
$$

また、$\lambda \le 0$ のとき $f_\Lambda(\lambda) = 0$ より $f_{\Lambda|X}(\lambda|k) = 0$ である。これより、$X = k$ の条件のもと Λ はガンマ分布 $\Gamma\big(\alpha+k, \beta+1\big)$ に従うことがわかった。

(2) (Y, Λ) の同時密度関数 $f(y, \lambda)$ は $f(y,\lambda) = f_{Y|\Lambda}(y|\lambda) f_\Lambda(\lambda)$ であるから、

$$
\begin{aligned}
f_Y(y) &= \int_{-\infty}^\infty f(y,\lambda)\, d\lambda = \int_0^\infty \frac{1}{\sqrt{2\pi/\lambda}} e^{-\lambda y^2/2} \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}\, d\lambda = \frac{\beta^\alpha}{\sqrt{2\pi}\Gamma(\alpha)} \int_0^\infty \lambda^{\alpha+\frac{1}{2}-1} e^{-(\beta+\frac{y^2}{2})\lambda}\, d\lambda \\
&= \frac{\beta^\alpha}{\sqrt{2\pi}\Gamma(\alpha)} \frac{\Gamma(\alpha+\frac{1}{2})}{(\beta+\frac{1}{2}y^2)^{\alpha+\frac{1}{2}}} = \frac{\Gamma(\alpha+\frac{1}{2})}{\sqrt{\pi}\Gamma(\alpha)} \Big(\frac{1}{2\beta}\Big)^{1/2} \Big(1 + \frac{1}{2\beta} y^2\Big)^{-\alpha-\frac{1}{2}}.
\end{aligned}
$$

ここで、$\alpha = \beta$ であれば Y は自由度 2α の t 分布に従うことに注意する ($cf.$ 定理 3.6)。次に、$\lambda > 0$ のとき、

$$
f_{\Lambda|Y}(\lambda|y) = \frac{f_{Y|\Lambda}(y|\lambda)}{f_Y(y)} = \frac{\dfrac{\beta^\alpha}{\sqrt{2\pi}\Gamma(\alpha)} \lambda^{\alpha+\frac{1}{2}-1} e^{-(\beta+\frac{y^2}{2})\lambda}}{\dfrac{\Gamma(\alpha+\frac{1}{2})}{\sqrt{\pi}\Gamma(\alpha)} \Big(\dfrac{1}{2\beta}\Big)^{1/2} \Big(1+\dfrac{1}{2\beta}y^2\Big)^{-\alpha-\frac{1}{2}}} = \frac{\big(\beta+\frac{1}{2}y^2\big)^{\alpha+\frac{1}{2}}}{\Gamma(\alpha+\frac{1}{2})} \lambda^{\alpha+\frac{1}{2}-1} e^{-(\beta+\frac{1}{2}y^2)\lambda}.
$$

また、$\lambda \le 0$ のとき $f_{\Lambda|Y}(\lambda|y) = 0$ である。これより、$Y = y$ の条件のもと Λ はガンマ分布 $\Gamma\big(\alpha+\frac{1}{2}, \beta+\frac{1}{2}y^2\big)$ に従うことがわかった。 □

問 2.12 (1) X はガンマ分布 $\Gamma(\alpha, 1)$ に従い、$X = x$ の下で $Y \sim \mathrm{Ex}(x)$ とするとき、以下を求めよ。

(a) (X, Y) の同時密度関数 $f(x, y)$, (b) $f_Y(y)$, (c) $f_{X|Y}(x|y)$ $(y > 0)$.

(2) $X \sim \mathrm{B{\small ETA}}(a, b)$ で $X = x$ の下 $Y \sim \mathrm{B}(n, x)$ (二項分布) とするとき、以下を求めよ。

(a) $P(Y = k)$, $k = 0, 1, \ldots, n$ (b) $Y = k$ の下での X の条件つき密度関数 $f_{X|Y}(x|k)$.

(3) $X \sim \mathrm{Ge}(p)$ で $X = k$ の下 $k = 0$ のとき $Y = 0$ (定数), $k \ge 1$ のとき $Y \sim \Gamma(k, \beta)$ とするとき、Y の分布関数が $F_Y(y) = P(Y \le y) = p1_{[0,\infty)}(y) + (1-p)(1-e^{-p\beta y})1_{[0,\infty)}(y)$ となることを示せ。

Hint: $P(Y \le y) = \sum\limits_{k=0}^\infty P(Y \le y, X = k) = \sum\limits_{k=0}^\infty P(Y \le y | X = k) P(X = k)$ に帰着せよ。

2.6 確率変数の独立性

定義 2.8 (1) 確率変数の列 $X_1, X_2, \ldots, X_n$ が独立であるとは、任意の $x_1, x_2, \ldots, x_n \in \mathbf{R}$ に対して

$$
P(X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n) = P(X_1 \le x_1) P(X_2 \le x_2) \cdots P(X_n \le x_n) \tag{2.14}
$$

となるときにいう。

(2) 無限個の確率変数の族 $\{X_\lambda\}$ が独立であるとは、その任意の有限部分列 $X_{\lambda_1}, \ldots, X_{\lambda_q}$ が独立であるときにいう。

確率変数の列 $X_1, X_2, \ldots, X_n$ が独立であることと、任意の Borel 集合 $B_1, \ldots, B_n \subset \mathbf{R}$ に対して

$$
P(X_1 \in B_1, X_2 \in B_2, \cdots, X_n \in B_n) = P(X_1 \in B_1) P(X_2 \in B_2) \cdots P(X_n \in B_n) \tag{2.15}
$$

となることが同値である。(証明には後期に学ぶ Dynkin 族定理を用いる。) 特に、離散型確率変数であれば

$$P(X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n) = P(X_1 = x_1)P(X_2 = x_2)\cdots P(X_n = x_n) \qquad \forall x_1, x_2, \ldots, x_n \in \mathbf{R}$$

と同値となる。絶対連続型確率変数の場合は次の定理で見るように同時密度関数がそれぞれの一次元周辺密度関数の積に分解できればよい。

定理 2.9 絶対連続型確率変数の列 $X_1, \ldots, X_n$ が独立であることための必要十分条件は、$(X_1, \ldots, X_n)$ の同時密度関数を $f(x_1, \ldots, x_n)$, X_j の周辺密度関数を $f_{X_j}(x)$ $(j = 1, \ldots, n)$ とするとき

$$f(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2)\cdots f_{X_n}(x_n), \qquad \forall x_1, x_2, \ldots, x_n \in \mathbf{R} \tag{2.16}$$

が成立することである。

証明: 絶対連続型であるから (2.14) を密度関数で表すと

$$\int_{-\infty}^{x_n}\cdots\int_{-\infty}^{x_1} f(t_1, t_2, \ldots, t_n)\,dt_1\cdots dt_n = \Big(\int_{-\infty}^{x_1} f_{X_1}(t)\,dt\Big) \times \cdots \times \Big(\int_{-\infty}^{x_n} f_{X_n}(t)\,dt\Big). \tag{2.17}$$

これを $x_1, \cdots, x_n$ でそれぞれ 1 回ずつ偏微分すれば (2.16) を得る。逆に、(2.16) を各 x_j に関して $(-\infty, x_j]$ の範囲で定積分すれば (2.17) を得る。 □

例 2.5 $\boldsymbol{X} = (X_1, \ldots, X_n)$ を n 次元正規分布 $\mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ に従うとする。$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)'$, $\boldsymbol{\Sigma} = (\sigma_{ij})$ とかく。このとき、$X_1, \ldots, X_n$ が独立であるための必要十分条件が、$\boldsymbol{\Sigma}$ が対角行列であること、即ち、$i \neq j$ となる任意の i, j で $\sigma_{ij} = 0$ となることである。

証明: (必要性) $\boldsymbol{\Sigma}$ が対角行列のとき $\boldsymbol{\Sigma}^{-1}$ も対角行列でその jj 成分は $1/\sigma_{jj}$ であるから、明らかに

$$\frac{1}{(2\pi)^{n/2}(\det\boldsymbol{\Sigma})^{1/2}}e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})} = \prod_{j=1}^{n}\frac{1}{\sqrt{2\pi\sigma_{jj}}}e^{-\frac{(x_j-\mu_j)^2}{2\sigma_{jj}}}. \tag{2.18}$$

が成立する。よって、定理 2.9 より独立であることが従う。

(十分性) 共分散の節で証明する (定理 4.13 とその系)。 □

例 2.6 $X_1, \ldots, X_n$ を独立で同分布に従う (i.i.d. と書く) 確率変数とし、$F(x)$ で X_1 の分布関数を表す。このとき、次が成立する。

(1) $F_{(n)}(x)$ を $X_{(n)} := \max\{X_1, \cdots, X_n\}$ の分布関数とすると、$F_{(n)}(x) = F(x)^n$ となる。

(2) $F_{(1)}(x)$ を $X_{(1)} := \min\{X_1, \cdots, X_n\}$ の分布関数とすると、$1 - F_{(1)}(x) = (1 - F(x))^n$ となる。

証明: (1) $F_{X_{(n)}}(x) = P(X_1 \le x, \ldots, X_n \le x) = \prod_{j=1}^{n} P(X_j \le x) = F(x)^n$.

(2) $1 - F_{X_{(1)}}(x) = P(X_{(1)} > x) = P(X_1 > x, \ldots, X_n > x) = \prod_{j=1}^{n} P(X_j > x) = (1 - F(x))^n$. □

例 2.7 X_1, X_2 が独立で、各 X_i が Poisson 分布 $\mathrm{Po}(\lambda_i)$ $(\lambda_i > 0)$ に従うとき、$Y_1 = X_1 + X_2$ と $Y_2 = X_2$ の同時分布を求めよ。また、Y_1 の周辺分布を求めよ。

解: (Y_1, Y_2) のとりうる値は、$(y_1, y_2) \in \boldsymbol{Z}^2 : y_1 \ge y_2 \ge 0$ である。このとき、同時分布は

$$P(Y_1 = y_1, Y_2 = y_2) = P(X_1 = y_1 - y_2, X_2 = y_2) = P(X_1 = y_1 - y_2)P(X_2 = y_2) = \frac{\lambda_1^{y_1-y_2}\lambda_2^{y_2}}{(y_1 - y_2)!y_2!}e^{-\lambda_1-\lambda_2}$$

となる。また、Y_1 の周辺分布は $y_1 \in \{0\} \cup \boldsymbol{N}$ に対して

$$P(Y_1 = y_1) = \sum_{y_2=0}^{y_1}\frac{\lambda_1^{y_1-y_2}\lambda_2^{y_2}}{(y_1 - y_2)!y_2!}e^{-\lambda_1-\lambda_2} = \frac{1}{y_1!}\sum_{y_2=0}^{y_1}\binom{y_1}{y_2}\lambda_1^{y_1-y_2}\lambda_2^{y_2}e^{-\lambda_1-\lambda_2} = \frac{(\lambda_1 + \lambda_2)^{y_1}}{y_1!}e^{-(\lambda_1+\lambda_2)}$$

となり、これは Y_1 が Poisson 分布 $\mathrm{Po}(\lambda_1 + \lambda_2)$ に従うことを示している。 □

例題 2.6 (1) X, Y, Z が独立で X は Poisson 分布 $\mathrm{Po}(\lambda)$, Y は幾何分布 $\mathrm{Ge}(p)$, Z は指数分布 $\mathrm{Ex}(\alpha)$ に従うとき $P(2X = Y)$ と $P(2X \leq Z)$ を求めよ。

解

$$\begin{aligned}
P(2X=Y) &= \sum_{k=0}^{\infty} P(X=k, 2X=Y) = \sum_{k=0}^{\infty} P(X=k, Y=2k) = \sum_{k=0}^{\infty} P(X=k)P(Y=2k) \\
&= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} p(1-p)^{2k} = e^{-\lambda} p e^{\lambda(1-p)^2} = p e^{-\lambda(2p-p^2)}, \\
P(2X \leq Z) &= \sum_{k=0}^{\infty} P(X=k, 2X \leq Z) = \sum_{k=0}^{\infty} P(X=k, Z \geq 2k) = \sum_{k=0}^{\infty} P(X=k)P(Z \geq 2k) \\
&= \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \int_{2k}^{\infty} \alpha e^{-\alpha z}\, dz = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} e^{-2\alpha k} = e^{-\lambda} e^{\lambda e^{-2\alpha}} = e^{-\lambda(1-e^{-2\alpha})}. \qquad \square
\end{aligned}$$

問 2.13 (1) X, Y は独立で $X \sim \mathrm{Ge}(p)$, $Y \sim \mathrm{Ge}(q)$ のとき、

(a) $P(X = 3Y)$, $P(X > 3Y)$, $P(X = \lfloor Y/3 \rfloor)$ を求めよ。($\lfloor a \rfloor$ は a 以下の最大の整数)。

(b) $S = X + Y$ と Y の同時分布を求め、S の周辺分布を求めよ。特に $p = q$ なら $S \sim \mathrm{NB}(2, p)$ を確認せよ。

(2) X, Y は独立で $X \sim \mathrm{Ge}(p)$, $Y \sim \mathrm{Po}(\lambda)$ のとき、$P(X = 3Y + 1)$, $P(X \geq 3Y + 1)$, $P(X \geq 3Y - 1)$ を求めよ。

(3) X, Y は独立で $X \sim \mathrm{Ge}(p)$, $Y \sim \mathrm{Ex}(\lambda)$ のとき、$P(X \leq Y)$ と $P(X = \lfloor Y \rfloor)$ を求めよ

例題 2.7 U_1, U_2 は独立で $U_1 \sim \mathrm{U}(0, a)$, $U_2 \sim \mathrm{U}(0, 1)$ $(a > 0)$ とし、$X = U_1 U_2$ とおく。このとき X の密度関数 $f_X(x)$ を求めよ。

解 まず、X の分布関数 $F_X(x) = P(X \leq x)$ を考える。
$P(0 < X < a) = 1$ より $F_X(x) = 0$ $(x < 0)$, $F_X(x) = 1$ $(x \geq 1)$.
$0 \leq x < a$ のとき、(U_1, U_2) の同時密度関数は $f(u, v) = f_{U_1}(u) f_{U_2}(v) = \begin{cases} \dfrac{1}{a}, & (u,v) \in [0,a] \times [0,1], \\ 0 & \text{その他}. \end{cases}$ なので、$D_x = \{(u, v) \in [0, a] \times [0, 1]; uv \leq x\}$ (図示しましょう) とすると、

$$\begin{aligned}
F_X(x) &= P(U_1 U_2 \leq x) = \iint_{uv \leq x} f(u,v)\, dudv = \iint_{D_x} \frac{1}{a}\, dudv = \frac{1}{a} \times (D_x \text{ の面積}) \\
&= \frac{1}{a}\Big(x + \int_x^a \frac{x}{u}\, du\Big) = \frac{1}{a}(x + x\log a - x \log x).
\end{aligned}$$

よって $f_X(x) = \begin{cases} \dfrac{1}{a}(\log a - \log x), & 0 < x < a, \\ 0 & \text{その他} \end{cases}$ となる。 $\square$

問 2.14 U_1, U_2 は独立で $U_1 \sim \mathrm{U}(0, a)$, $U_2 \sim \mathrm{U}(0, 1)$ $(a > 1)$ とする。次の確率変数 X, Y, Z の分布関数と密度関数をそれぞれ求めよ。

(1) $X = U_1 + U_2$, (2) $Y = U_2 - U_1$, (3) $Z = U_2/U_1$.

注意 2.6 例題 2.7 の (U_1, U_2) は $[0, a] \times [0, 1]$ 上の一様分布に従う。例題 2.3 でも見たように (X, Y) の取り得る値が有界となる場合 $X + Y$ など四則演算で得られる確率変数の密度関数は、次節の定理 3.2 や問 3.1 の式を用いるより、上記のようにまず図を用いて分布関数を計算して求めるほうが容易となることが多い。

定理 2.10 m 次元確率変数 $\boldsymbol{X} = (X_1, \ldots, X_m)$ と 1 次元確率変数 $Y_1, \ldots, Y_n$ が独立であるとする: 即ち、

$$\begin{aligned}
&P(X_1 \leq x_1, \cdots, X_m \leq x_m, Y_1 \leq y_1, \cdots, Y_n \leq y_n) \\
&= P(X_1 \leq x_1, \cdots, X_m \leq x_m) P(Y_1 \leq y_1) \cdots P(Y_n \leq y_n), \qquad \forall x_1, \ldots, x_m, y_1, \ldots, y_n \in \mathbf{R}
\end{aligned}$$

とする。このとき、$f : \mathbf{R}^m \to \mathbf{R}$ が Borel 可測であれば、$f(X_1, \cdots, X_m), Y_1, \ldots, Y_n$ は独立である。

証明は難しいので省略する。(たとえば、後期に学ぶ Dynkin 族定理を用いて証明できる。)

3 確率変数の変換

3.1 絶対連続型確率変数の変換

離散型確率変数の変換は例 2.7 や演習問題にあるように比較的容易に考察できた。絶対連続型確率変数の場合は次の重積分の変数変換の公式を用いる。

定理 3.1 n 次元確率変数 $\boldsymbol{X} = (X_1, \dots, X_n)$ の密度関数を $f_{\boldsymbol{X}}(x_1, \cdots, x_n)$ とする。$y_i = u_i(x_1, \cdots, x_n)$ $(1 \le i \le n)$ が $f_{\boldsymbol{X}}(x_1, \cdots, x_n) > 0$ なる領域で次の仮定 (1)–(3) を満たす $\mathbf{R}^n$ から $\mathbf{R}^n$ への 1 対 1 変換とする。

(1) 変換 $y_i = u_i(x_1, \cdots, x_n)$ $(1 \le i \le n)$ およびその逆変換 $x_i = v_i(y_1, \cdots, y_n)$ $(1 \le i \le n)$ はともに連続である。

(2) 偏導関数 $\dfrac{\partial x_i}{\partial y_j}$ $(1 \le i, j \le n)$ が存在して連続である。

(3) ヤコビアン $\dfrac{\partial(x_1, \cdots, x_n)}{\partial(y_1, \cdots, y_n)}$ がゼロにならない。

このとき、$Y_i = u_i(X_1, \cdots, X_n)$ $(1 \le i \le n)$ で $(X_1, \cdots, X_n)$ を $(Y_1, \cdots, Y_n)$ に変換するとき、$\boldsymbol{Y} = (Y_1, \cdots, Y_n)$ の分布は絶対連続型でその密度関数 $f_{\boldsymbol{Y}}(y_1, \cdots, y_n)$ は次式で与えられる。

$$f_{\boldsymbol{Y}}(y_1, \cdots, y_n) = f_{\boldsymbol{X}}(v_1(y_1, \cdots, y_n), \cdots, v_n(y_1, \cdots, y_n))\left|\frac{\partial(x_1, \cdots, x_n)}{\partial(y_1, \cdots, y_n)}\right|.$$

証明: $(a_1, \cdots, a_n) \in \mathbf{R}^n$ に対して $E = \{(x_1, \cdots, x_n); u_1(x_1, \cdots, x_n) \le a_1, \dots, u_n(x_1, \cdots, x_n) \le a_n\}$ とおくと、この変換で E は $E' = \{(y_1, \cdots, y_n); y_1 \le a_1, \dots, y_n \le a_n\}$ に対応するから、多変数関数の積分の変数変換の公式により

$$\begin{aligned}
&P(Y_1 \le a_1, \cdots, Y_n \le a_n) = P(u_1(X_1, \cdots, X_n) \le a_1, \cdots, u_n(X_1, \cdots, X_n) \le a_n) \\
&= P((X_1, \cdots, X_n) \in E) = \int \cdots \int_E f_{\boldsymbol{X}}(x_1, \cdots, x_n)\, dx_1 \cdots dx_n = \int \cdots \int_{E'} f_{\boldsymbol{Y}}(y_1, \cdots, y_n)\, dy_1 \cdots dy_n \\
&= \int_{-\infty}^{a_n} \cdots \int_{-\infty}^{a_1} f_{\boldsymbol{Y}}(y_1, \cdots, y_n)\, dy_1 \cdots dy_n \qquad \square
\end{aligned}$$

定理 3.2 (X, Y) の密度関数を $f(x, y)$ とする。このとき、$Z = X + Y$ とすると Z の密度関数 $f_Z(z)$ は

$$f_Z(z) = \int_{-\infty}^{\infty} f(t, z - t)\, dt$$

となる。特に、$P(X \ge 0, Y \ge 0) = 1$ なら、$f_Z(z) = \displaystyle\int_0^z f(t, z - t)\, dt$ $(z > 0)$, $= 0$ $(z \le 0)$ となる。

証明: $Z = X + Y$, $T = X$ という変換を考えると、$X = T$, $Y = Z - T$ よりヤコビアンは

$$\frac{\partial(x, y)}{\partial(t, z)} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1.$$

よって、定理 3.1 より (T, Z) の密度関数は $f_{(T,Z)}(t, z) = f(t, z - t)|-1| = f(t, z - t)$ となるので、$f_Z(z)$ は Z の周辺密度関数となるから、t で積分して与式で与えられることがわかる。 $\square$

定理 3.3 確率変数 $X_1, \dots, X_n$ は独立で、各 X_i は正規分布 $\mathrm{N}(\mu_i, \sigma_i^2)$ に従うとする。このとき、$Y = X_1 + \cdots + X_n$ は正規分布 $\mathrm{N}(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)$ に従う。

証明: $n=2$ の場合に証明する。一般の場合は定理 2.10 より $X_1+\cdots+X_k$ と X_{k+1} が独立になることに注意して、これを繰り返し用いればよい。X_i の密度関数は $f_{X_i}(x_i)=\dfrac{1}{\sqrt{2\pi}\sigma_i}e^{-\frac{(x_i-\mu_i)^2}{2\sigma_i^2}}$ であり、(X_1,X_2) の同時密度関数は $f_{X_1}(x_1)f_{X_2}(x_2)$ であった。したがって、$Y=X_1+X_2$ の密度関数は定理 3.2 より、

$$f_Y(y)=\int_{-\infty}^{\infty}\frac{1}{2\pi\sigma_1\sigma_2}\exp\Big[-\frac{(t-\mu_1)^2}{2\sigma_1^2}-\frac{(y-t-\mu_2)^2}{2\sigma_2^2}\Big]\,dt.$$

ここで、$u=t-\mu_1$ とおき、exp の中を u について平方完成すると

$$-\frac{u^2}{2\sigma_1^2}-\frac{(y-u-\mu_1-\mu_2)^2}{2\sigma_2^2}=-\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}\Big[u-\frac{\sigma_1^2(y-\mu_1-\mu_2)}{\sigma_1^2+\sigma_2^2}\Big]^2-\frac{(y-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}$$

となり、$s=\dfrac{\sqrt{\sigma_1^2+\sigma_2^2}}{\sigma_1\sigma_2}\Big[u-\dfrac{\sigma_1^2(y-\mu_1-\mu_2)}{\sigma_1^2+\sigma_2^2}\Big]$ とおき、$\displaystyle\int_{-\infty}^{\infty}e^{-\frac{1}{2}s^2}\,ds=\sqrt{2\pi}$ となることを用いると、

$$f_Y(y)=\frac{1}{\sqrt{2\pi(\sigma_1^2+\sigma_2^2)}}\exp\Big[-\frac{(y-\mu_1-\mu_2)^2}{2(\sigma_1^2+\sigma_2^2)}\Big]. \qquad \square$$

定理 3.4 確率変数 $X_1,\cdots,X_n$ は独立で、各 X_i はガンマ分布 $\Gamma(\alpha_i,\beta)$ に従うとする。このとき、$Y=X_1+\cdots+X_n$ はガンマ分布 $\Gamma(\sum_{i=1}^n\alpha_i,\beta)$ に従う。

証明: $n=2$ の場合に証明する。一般の場合は、$X_1+\cdots+X_k$ と X_{k+1} が独立になることに注意してこれを繰り返し用いればよい。X_i の密度関数を f_{X_i} とすると、$Y=X_1+X_2$ の密度関数は定理 3.2 より、

$$f_Y(y)=\int_{-\infty}^{\infty}f_{X_1}(t)f_{X_2}(y-t)\,dt=\frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^{\max\{y,0\}}t^{\alpha_1-1}(y-t)^{\alpha_2-1}e^{-\beta(t+y-t)}\,dt \tag{3.1}$$

$y\le 0$ のとき $f_Y(y)=0$ は明らか。$y>0$ とき、$t=ys$ と置換することで、

$$\int_0^y t^{\alpha_1-1}(y-t)^{\alpha_2-1}\,dt=y^{\alpha_1+\alpha_2-1}\int_0^1 s^{\alpha_1-1}(1-s)^{\alpha_2-1}\,ds=y^{\alpha_1+\alpha_2-1}\mathrm{B}(\alpha_1,\alpha_2)$$

となり、補題 2.3(3) を用いて (3.1) に代入して、

$$f_Y(y)=\frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}y^{\alpha_1+\alpha_2-1}\frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}e^{-\beta y}=\frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}y^{\alpha_1+\alpha_2-1}e^{-\beta y}. \qquad \square$$

問 3.1 (X,Y) の同時密度関数が $f(x,y)$ で与えられるとし、$T=Y$, $U=X-Y$, $V=XY$, $W=X/Y$ とする。
(1) (U,T) の同時密度関数が $f(u+t,t)$ となることを示し、$f_U(u)=\displaystyle\int_{-\infty}^{\infty}f(u+t,t)\,dt$ を導け。
(2) (V,T) の同時密度関数が $f\Big(\dfrac{v}{t},t\Big)\dfrac{1}{|t|}$ となることを示し、$f_V(v)=\displaystyle\int_{-\infty}^{\infty}f\Big(\frac{v}{t},t\Big)\frac{1}{|t|}\,dt$ を導け。
(3) (W,T) の同時密度関数が $f(wt,t)|t|$ となることを示し、$f_V(v)=\displaystyle\int_{-\infty}^{\infty}f(vt,t)|t|\,dt$ を導け。

問 3.2 (1) 確率変数 X,Y が独立で $X\sim\mathrm{Ex}(\lambda_1)$, $Y\sim\mathrm{Ex}(\lambda_2)$ とし $Z=X+Y$, $W=\dfrac{X}{Y}$ とおく。$f_Z(z)$, $f_W(w)$ を求めよ。
(2) X,Y が独立でともに $\mathrm{N}(0,1)$ に従うとき、$W=\dfrac{X}{Y}$ について $f_W(w)$ を求めよ。
(3) X,Y が独立でともに $N(0,1)$ に従うとし $S=2X+3Y$, $T=X+Y$ とおく。(S,T) の同時密度関数 $g(s,t)$ を求めよ。また S の周辺分布を求めよ。
(4) X,Y が独立で $X\sim\Gamma(\alpha_1,\beta)$, $Y\sim\Gamma(\alpha_2,\beta)$ とし $S=X+Y$, $T=\dfrac{X}{X+Y}$ とおく。(S,T) の同時密度関数 $g(s,t)$ を求め、S,T は独立で T は $\textsc{Beta}(\alpha_{,1},\alpha_2)$ に従うことを示せ。
(5) X,Y が独立でともにその密度関数が $f(x)=\dfrac{1}{x^2}$ $(x>1)$, $=0$ $(x\le 1)$ であるとし、$S=\dfrac{X}{Y}$, $T=XY$ とおく。(S,T) の同時密度関数 $g(s,t)$ を求め、さらに、S と T の周辺密度関数をそれぞれ求めよ。

問 3.3 (1) Y, Z が独立で、$Y \sim \Gamma(n/2, 1/2)$, $Z \sim \mathrm{N}(0,1)$ のとき、$S = Y$, $T = \dfrac{Z}{\sqrt{Y/n}}$ の (S,T) の同時密度関数 $g(s,t)$ を求め、さらに $f_T(t)$ を求めよ。
(2) X, Y が独立で、$X \sim \Gamma(m/2, 1/2)$, $Y \sim \Gamma(n/2, 1/2)$ とするとき、$W = \dfrac{X/m}{Y/n}$ の密度関数 $f_W(w)$ を求めよ。

3.2 正規母集団における標本平均・不偏分散とその関数の分布

次の定理を踏まえて、ガンマ分布 $\Gamma(\frac{n}{2}, \frac{1}{2})$ を自由度 n のカイ二乗分布 χ_n^2 に従うという。

定理 3.5 確率変数 Z が標準正規分布 $\mathrm{N}(0,1)$ に従うとするとき、Z^2 は自由度 1 のカイ二乗分布 χ_1^2 に従う。特に、確率変数 $Z_1, \cdots, Z_n$ は独立で、各 Z_i は標準正規分布 $\mathrm{N}(0,1)$ に従うとする。このとき、$Y_n = X_1^2 + \cdots + X_n^2$ は自由度 n のカイ二乗分布 χ_n^2 に従う。

証明: 問 2.7 (1) により $Z^2 \sim \chi_1^2$. 後者の主張は、$X_1, \cdots, X_n$ は独立であれば、$X_1^2, \cdots, X_n^2$ も独立であるので、定理 3.4 からの帰結となる。 □

確率変数 $X_1, \ldots, X_n$ が独立で $\mathrm{N}(\mu, \sigma^2)$ に従うとき、

$$\text{標本平均: } \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i, \qquad \text{不偏分散: } U^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

の分布について考える。まず、次について考える。

定理 3.6 (1) 確率変数 Y, Z が独立で、それぞれ Y が自由度 n のカイ二乗分布 χ_n^2 に、Z は標準正規分布 $\mathrm{N}(0,1)$ に従うとき、$T = \frac{Z}{\sqrt{Y/n}}$ は次の密度関数 $f_T(t)$ をもつ。

$$f_T(t) = \frac{1}{\sqrt{n}B\left(\frac{n}{2}, \frac{1}{2}\right)}\left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}.$$

このとき T は自由度 n の t 分布 t_n に従うという。
(2) 確率変数 X, Y が独立で、それぞれが自由度 m, n のカイ二乗分布 χ_m^2, χ_n^2 に従うとき、$W = \frac{X/m}{Y/n}$ は次の密度関数 $f_W(w)$ をもつ。

$$f_W(w) = \frac{\left(\frac{m}{n}\right)^{m/2}}{B\left(\frac{m}{2}, \frac{n}{2}\right)} \frac{w^{\frac{m}{2}-1}}{\left(1 + \frac{m}{n}w\right)^{\frac{m+n}{2}}}. \quad (w > 0), \qquad = 0 \quad (w \le 0).$$

このとき W は自由度 (m,n) の F 分布 F_n^m に従うという。

証明は問 3.3 とした。 □

定理 3.7 $X_1, \ldots, X_n$ が独立で、それぞれ同一の正規分布 $\mathrm{N}(\mu, \sigma^2)$ に従うとするとき、次が成立する。
(1) $\overline{X}$ は正規分布 $\mathrm{N}\Big(\mu, \dfrac{\sigma^2}{n}\Big)$ に従う。
(2) $\overline{X}$ と $U^2 = \dfrac{1}{n-1}\sum\limits_{i=1}^{n}(X_i - \overline{X})^2$ は独立。
(3) $\dfrac{n-1}{\sigma^2}U^2 = \dfrac{1}{\sigma^2}\sum\limits_{i=1}^{n}(X_i - \overline{X})^2$ は自由度 $n-1$ のカイ二乗分布 χ_{n-1}^2 に従う。

証明: $Z_i = \dfrac{X_i - \mu}{\sigma}$, $i = 1, \ldots, n$ とおくと、$Z_1, \cdots, Z_n$ は独立で $\mathrm{N}(0,1)$ に従う。次に

$$\boldsymbol{p}_0 = \Big(\frac{1}{\sqrt{n}}, \cdots, \frac{1}{\sqrt{n}}\Big)', \quad \boldsymbol{p}_k = \Big(\underbrace{\frac{1}{\sqrt{k^2+k}}, \cdots, \frac{1}{\sqrt{k^2+k}}}_{k\text{ 個}}, \frac{-k}{\sqrt{k^2+k}}, 0 \cdots, 0\Big)', \ k = 1, \cdots, n-1$$

とし、$P=(\boldsymbol{p}_0\boldsymbol{p}_1\cdots\boldsymbol{p}_{n-1})$ とすると、P は直交行列となる。ここで $(T_1\cdots T_n)=(Z_1\cdots Z_n)P$ とすると、$T_1,\ldots,T_n$ は独立で $\mathrm{N}(0,1)$ に従う。実際、$(Z_1\cdots Z_n)$ の密度関数を $f(z_1,\cdots,z_n)$ とすると、

$$f(z_1,\cdots,z_n)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z_i^2}=\Big(\frac{1}{\sqrt{2\pi}}\Big)^n e^{-\frac{1}{2}\sum_{i=1}^{n}z_i^2}$$

であるが、$(T_1\cdots T_n)$ の密度関数 $g(t_1,\cdots,t_n)$ は $(t_1\cdots t_n)=(z_1\cdots z_n)P$ より $(z_1\cdots z_n)=(t_1\cdots t_n)P'$ となるから $\Big|\dfrac{\partial(z_1,\cdots,z_n)}{\partial(t_1,\cdots,t_n)}\Big|=|\det P'|=1$ で、$\boldsymbol{z}=(z_1\cdots z_n)'$, $\boldsymbol{t}=(t_1\cdots t_n)'$ と表すと、$\boldsymbol{z}=P'\boldsymbol{t}$ なので

$$\sum_{i=1}^{n}z_i^2=\boldsymbol{z}'\boldsymbol{z}=(P'\boldsymbol{t})'P'\boldsymbol{t}=\boldsymbol{t}'PP'\boldsymbol{t}=\boldsymbol{t}'\boldsymbol{t}=\sum_{i=1}^{n}t_i^2 \tag{3.2}$$

より、定理 3.1 より

$$g(t_1,\cdots,t_n)=f(P'(t_1\cdots t_n)')\Big|\frac{\partial(z_1,\cdots,z_n)}{\partial(t_1,\cdots,t_n)}\Big|=\Big(\frac{1}{\sqrt{2\pi}}\Big)^n e^{-\frac{1}{2}\sum_{i=1}^{n}t_i^2}=\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t_i^2}$$

となるからである。

ここで、$T_1=\dfrac{1}{\sqrt{n}}\sum\limits_{i=1}^{n}Z_i$ より

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i=\frac{1}{n}\sum_{i=1}^{n}(\sigma Z_i+\mu)=\frac{\sigma}{\sqrt{n}}T_1+\mu$$

であるから、$\overline{X}$ は $\mathrm{N}\Big(\mu,\dfrac{\sigma^2}{n}\Big)$ に従う。また、(3.2) より $\sum\limits_{i=1}^{n}Z_i^2=\sum\limits_{i=1}^{n}T_i^2$ であるが、$\overline{Z}=\dfrac{1}{n}\sum\limits_{i=1}^{n}Z_i$ とすると

$$\begin{aligned}\frac{n-1}{\sigma^2}U^2&=\frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{1}{\sigma^2}\sum_{i=1}^{n}(\sigma Z_i+\mu-\sigma\overline{Z}-\mu)^2=\sum_{i=1}^{n}(Z_i-\overline{Z})^2=\sum_{i=1}^{n}(Z_i^2-2\overline{Z}Z_i+\overline{Z}^2)\\&=\sum_{i=1}^{n}Z_i^2-2\overline{Z}\sum_{i=1}^{n}Z_i+n\overline{Z}^2=\sum_{i=1}^{n}Z_i^2-n\overline{Z}^2=\sum_{i=1}^{n}T_i^2-n\Big(\frac{1}{\sqrt{n}}T_1\Big)^2=\sum_{i=2}^{n}T_i^2.\end{aligned} \tag{3.3}$$

よって、定理 3.5 より (3) は従う。また、$\overline{X}$ は T_1 のみで表されることと (3.3) および $T_1,\cdots,T_n$ が独立であることから定理 2.10 より (2) は従う。 □

母分散の区間推定

$\chi_n^2(\alpha)$ で自由度 n の χ^2 分布 χ_n^2 の上側 α 点、即ち、$X\sim\chi_n^2$ のとき、$P(X>\chi_n^2(\alpha))=\alpha$ となる点とする。このとき、U^2 を不偏分散とすると、定理 3.7 より $\dfrac{n-1}{\sigma^2}U^2\sim\chi_{n-1}^2$ なので、

$$P\Big(\chi_{n-1}^2(1-\alpha/2)\le\frac{n-1}{\sigma^2}U^2\le\chi_{n-1}^2(\alpha/2)\Big)=1-\alpha$$

となる。これを変形して

$$\frac{(n-1)U^2}{\chi_{n-1}^2(\alpha/2)}\le\sigma^2\le\frac{(n-1)U^2}{\chi_{n-1}^2(1-\alpha/2)}$$

となるが、U^2 に実現値 u^2 を代入することで、σ^2 の $100(1-\alpha)\%$ 信頼区間を得る。

定理 3.8 $X_1,\ldots,X_n$ が独立で、それぞれ同一の正規分布 $\mathrm{N}(\mu,\sigma^2)$ に従うとするとき、

$$T=\frac{\overline{X}-\mu}{\sqrt{U^2/n}}$$

は自由度 $n-1$ の t 分布 t_{n-1} に従う。ここで、$\overline{X}$ は標本平均、U^2 は不偏分散を表す。

証明: 定理 3.7 より $Z = \dfrac{\overline{X}-\mu}{\sigma/\sqrt{n}}$, $Y = \dfrac{n-1}{\sigma^2}U^2$ とおくと、Z と Y は独立で Z は $\mathrm{N}(0,1)$ に、Y は χ^2_{n-1} に従う。ここで、

$$T = \frac{(\sigma/\sqrt{n})Z}{\sqrt{(\sigma^2 Y/(n-1))/n}} = \frac{Z}{\sqrt{Y/(n-1)}}$$

となるが、定理 3.6 (1) より T は t_{n-1} に従うことがわかる。 □

母分散が未知の場合の母平均の区間推定

$t_n(\alpha)$ で自由度 n の t 分布の (片側)α 点、即ち T が t_n 分布に従うとき、$P(T > t_n(\alpha)) = \alpha$ となる点とする。このとき、定理 3.8 より

$$P\Big(-t_{n-1}(\alpha/2) \leq \frac{\overline{X}-\mu}{\sqrt{U^2/n}} \leq t_{n-1}(\alpha/2)\Big) = 1-\alpha$$

となる。これを変形して

$$\overline{X} - t_{n-1}(\alpha/2)\sqrt{\frac{U^2}{n}} \leq \mu \leq \overline{X} + t_{n-1}(\alpha/2)\sqrt{\frac{U^2}{n}}$$

となるが、$\overline{X}$ と U^2 に実現値 $\overline{x}$ と u^2 を代入することで、μ の $100(1-\alpha)\%$ 信頼区間を得る。

例 3.1 [TS p.96] 次の数値はある複合肥料 6 袋の重量を測定した値である。[*2]

25.13, 25.32, 25.06, 24.98, 25.18, 25.17 (単位 kg)

この袋づめ複合肥料の重量は正規母集団 $\mathrm{N}(\mu,\sigma^2)$ に従い、上の 6 個の数値は、この母集団からの無作為標本と見て、(a) σ^2 と (b) μ の 95% 信頼区間を求めよ。

解: $\overline{x} = 25.14$, 標本分散 $s^2 = \overline{x^2} - \overline{x}^2 = \dfrac{0.067}{6}$ より $u^2 = \dfrac{6}{5}s^2$.

(a) $\dfrac{(6-1)u^2}{\chi^2_5(0.025)} = \dfrac{0.067}{12.83} \fallingdotseq 0.005222,\ \dfrac{(6-1)u^2}{\chi^2_5(0.975)} = \dfrac{0.067}{0.8312} \fallingdotseq 0.080606$ より
求める信頼区間は $0.0052 \leq \sigma^2 \leq 0.08061$.

(b) $\overline{x} \pm t_5(0.025)\sqrt{\dfrac{u^2}{6}} \fallingdotseq \begin{cases} 25.0185 \\ 25.2615 \end{cases}$ より求める信頼区間は $25.01 \leq \mu \leq 25.27$. □

問 3.4 次は、ある正規母集団から抽出した大きさ 10 の標本である:

5.3 5.0 5.8 5.1 5.9 5.2 6.1 5.4 5.6 5.6

この母分散 σ^2 と母平均 μ の 95% 信頼区間を求めよ。ただし、信頼区間の下限の端数は切り捨て上限の端数は切り上げることで母分散は小数点以下第 4 位まで、母平均は小数点以下第 2 位まで求めよ。

問 3.5 正規母集団 $\mathrm{N}(\mu,\sigma^2)$ において、μ,σ^2 ともに未知とする。μ の信頼係数 0.95 の信頼区間の幅が 2σ より小さくなる確率が 0.95 以上となるためには、標本数 n をどのくらい大きくとる必要があるか。

問 3.6 V がベータ分布 $\mathrm{B{\scriptsize ETA}}\Big(\dfrac{m}{2},\dfrac{n}{2}\Big)$ に、W が自由度 (m,n) の F 分布に従うとき、$P(V \leq x) = P\Big(W \leq \dfrac{n}{m}\dfrac{x}{1-x}\Big)$, $0 < x < 1$ を示せ。ヒント: 問 3.2(4) を用い、ガンマ分布 (カイ 2 乗分布) に帰着せよ。

[*2] [TS] 統計と社会の教科書「新確率統計」です。あわせて [TS] pp.99–101 の例題 1, 2, 3 と問 4, 5, 6, p.105 の練習問題 1-A 1, 2 を解いてみてください。また、[TS] pp.113–116 の例題 2, 3 と問 5, 6, p.123 の練習問題 2-A 2, 3, 5 も解いておいてください。

4　期待値

4.1　Lebesgue 積分

Lebesgue 積分論に基づいて $\mathbf{R}^n$ 上の Borel 可測関数 $\varphi(\boldsymbol{x})$ と n 次元確率変数 $\boldsymbol{X} = (X_1, \cdots, X_n)$ に対して、$\displaystyle\int_\Omega \varphi(\boldsymbol{X}(\omega))\, dP(\omega)$ の定義の概略を述べる。

<u>1st step</u>　$\varphi(\boldsymbol{x})$ が単関数のとき、即ち、$\varphi(\boldsymbol{x}) = \sum_{k=1}^N a_k 1_{A_k}(\boldsymbol{x})$ と表せるとき。ここで、$a_1, \ldots, a_N \in \mathbf{R}$, $A_1, \ldots, A_N \in \mathcal{B}(\mathbf{R}^n)$ とし、$1_A(\boldsymbol{x})$ は A の定義関数を表す。

$$1_A(\boldsymbol{x}) = \begin{cases} 1 & (\boldsymbol{x} \in A) \\ 0 & (\boldsymbol{x} \notin A) \end{cases}$$

このとき、次のように定義する。(本来は well-defined かどうか調べなければならないが省略する。)

$$\int_\Omega \varphi(\boldsymbol{X}(\omega))\, dP(\omega) = \sum_{k=1}^N a_k P(\boldsymbol{X} \in A_k). \tag{4.1}$$

ここで、φ, ψ がともに単関数で $\varphi(\boldsymbol{x}) \leq \psi(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbf{R}^n$, であれば $\displaystyle\int_\Omega \varphi(\boldsymbol{X}(\omega))\, dP(\omega) \leq \int_\Omega \psi(\boldsymbol{X}(\omega))\, dP(\omega)$ となることに注意する。(証明は略す。)

<u>2nd step</u>　$\varphi(\boldsymbol{x}) \geq 0$ のとき。各 $q \in \mathbf{N}$ に対して、

$$A_{q,i} = \left\{\boldsymbol{x} \in \mathbf{R}^n \,;\, \frac{i}{2^q} \leq \varphi(\boldsymbol{x}) < \frac{i+1}{2^q}\right\} \quad (i = 0, 1, \ldots, q2^q - 1), \qquad A_{q,q2^q} = \left\{\boldsymbol{x} \in \mathbf{R}^n \,;\, \varphi(\boldsymbol{x}) \geq q\right\}$$

とし、$\displaystyle\varphi_q(\boldsymbol{x}) = \sum_{i=0}^{q2^q} \frac{i}{2^q} 1_{A_{q,i}}(\boldsymbol{x})$ と定める。このとき、$\varphi_q(\boldsymbol{x})$ は単関数で (4.1) により $\displaystyle\int_\Omega \varphi_q(\boldsymbol{X}(\omega))\, dP(\omega)$ は定義される。ここで、$\varphi_q(\boldsymbol{x}) \leq \varphi_{q+1}(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbf{R}^n$, に注意して $\displaystyle\int_\Omega \varphi_q(\boldsymbol{X}(\omega))\, dP(\omega) \leq \int_\Omega \varphi_{q+1}(\boldsymbol{X}(\omega))\, dP(\omega)$ を得る。よって、$\displaystyle\left\{\int_\Omega \varphi_q(\boldsymbol{X}(\omega))\, dP(\omega)\right\}$ は単調増加列なので、その $q \to \infty$ とした極限は無限大かもしれないが定義される。これを用いて、また $\displaystyle\lim_{q\to\infty}(\boldsymbol{x}) = \varphi(\boldsymbol{x})$, に注意して、次のように定める。

$$\int_\Omega \varphi(\boldsymbol{X}(\omega))\, dP(\omega) = \lim_{q\to\infty} \int_\Omega \varphi_q(\boldsymbol{X}(\omega))\, dP(\omega). \tag{4.2}$$

(本来はここでも、積分値が列 $\{\varphi_q\}$ のとり方によらないか調べなければならないが省略する。)

<u>3rd step</u>　一般の $\varphi(\boldsymbol{x})$ について。

$$\varphi^+(\boldsymbol{x}) = \max\{\varphi(\boldsymbol{x}), 0\}, \qquad \varphi^-(\boldsymbol{x}) = \max\{-\varphi(\boldsymbol{x}), 0\} \tag{4.3}$$

とおく。このとき、$\varphi^+(\boldsymbol{x}) \geq 0$, $\varphi^-(\boldsymbol{x}) \geq 0$ および, $\varphi(\boldsymbol{x}) = \varphi^+(\boldsymbol{x}) - \varphi^-(\boldsymbol{x})$, $|\varphi(\boldsymbol{x})| = \varphi^+(\boldsymbol{x}) + \varphi^-(\boldsymbol{x})$ に注意する。これを用いて、$\displaystyle\int_\Omega \varphi^+(\boldsymbol{X}(\omega))\, dP(\omega) < \infty$ または $\displaystyle\int_\Omega \varphi^-(\boldsymbol{X}(\omega))\, dP(\omega) < \infty$ のとき、

$$\int_\Omega \varphi(\boldsymbol{X}(\omega))\, dP(\omega) = \int_\Omega \varphi^+(\boldsymbol{X}(\omega))\, dP(\omega) - \int_\Omega \varphi^-(\boldsymbol{X}(\omega))\, dP(\omega) \tag{4.4}$$

と定める。

注意 4.1 上記の定義は、$\mu(B) = P(\boldsymbol{X} \in B)$ $(B \in \mathcal{B}(\mathbf{R}^n))$ によって定義される確率測度 μ ($\boldsymbol{X}$ の分布という) に対して、$\displaystyle\int_{\mathbf{R}^n} \varphi(\boldsymbol{x})\, d\mu(\boldsymbol{x})$ を定義したことに相当する。特に、$n = 1$ のときこの積分を確率変数 X の分布関数 $F(x)$ を用いて、$\displaystyle\int_{-\infty}^{\infty} \varphi(x)\, dF(x)$ と表すことがある。

以下の Lebesgue 積分論の定理 4.1, 4.2 が成立することに注意する。(証明は関数解析学で勉強のこと。)

定理 4.1 (1) 積分は線形性や単調性を有する。(期待値の形で定理 4.3 で述べる。)
(2) $\varphi(\boldsymbol{x}) \geq 0,\ \int_\Omega \varphi(\boldsymbol{X}(\omega))\, dP(\omega) = 0 \Longrightarrow P(\varphi(\boldsymbol{X}) = 0) = 1.$

注意 4.2 $P(\varphi(\boldsymbol{X}) = 0) = 1$ のとき、ほとんどいたるところ (almost everywhere) $\varphi(\boldsymbol{X}) = 0$ といい、「$\varphi(\boldsymbol{X}) = 0$ a.e.」と表す。(ほとんど確かに (almost surely) といい、「$\varphi(\boldsymbol{X}) = 0$ a.s.」と表すこともある。)

定理 4.2 (1) **(単調収束定理)** $(R, \mathfrak{A}, \mu)$ を測度空間とする。可測関数の列 $\{f_q(x)\}$ が非負値で単調増加であるとする, $0 \leq f_1(x) \leq \cdots \leq f_q(x) \leq \cdots$ a.e. このとき、次が成立する。

$$\lim_{q\to\infty} \int_R f_q(x)\, d\mu(x) = \int_R \Big(\lim_{q\to\infty} f_q(x) \Big)\, d\mu(x).$$

(2) **(Lebesgue の収束定理)** $(R, \mathfrak{A}, \mu)$ を測度空間とする。可測関数の列 $\{f_q(x)\}$ に対して、ある可測関数 $M(x)$ で $\int_R M(x)\, d\mu(x) < \infty$ を満たすもの (可積分という) が存在して、$|f_q(x)| \leq M(x)$ a.e. $(q = 1, 2, \ldots)$ を満たすとるとする。このとき、$f(x) = \lim_{q\to\infty} f_q(x)$ a.e. が存在して可測であれば、$f(x)$ も可積分であり、次が成立する。

$$\lim_{q\to\infty} \int_R f_q(x)\, d\mu(x) = \int_R f(x)\, d\mu(x).$$

4.2 期待値の定義

§4.1 で準備した積分を用いて、$\varphi \geq 0$ あるいは $\int_\Omega \varphi^+(\boldsymbol{X}(\omega))\, dP(\omega) < \infty$ または $\int_\Omega \varphi^-(\boldsymbol{X}(\omega))\, dP(\omega) < \infty$ のとき、期待値 $E[\varphi(\boldsymbol{X})]$ を

$$E[\varphi(\boldsymbol{X})] = \int_\Omega \varphi(\boldsymbol{X}(\omega))\, dP(\omega)$$

で定義する。次の定理が成立する。証明は省略する。

定理 4.3 $\varphi, \psi : \mathbf{R}^n \to \mathbf{R}$ を Borel 可測関数、a, b, c を定数とする。このとき、次が成立する。
(1) $\varphi(\boldsymbol{X}) = c$ a.e. のとき、$E[\varphi(\boldsymbol{X})] = c.$
(2) $\varphi(\boldsymbol{X}) \leq \psi(\boldsymbol{X})$ a.e. のとき、$E[\varphi(\boldsymbol{X})] \leq E[\psi(\boldsymbol{X})].$
(3) $|E[\varphi(\boldsymbol{X})]| \leq E[|\varphi(\boldsymbol{X})|].$
(4) $E[|\varphi(\boldsymbol{X})|] < \infty,\ E[|\psi(\boldsymbol{X})|] < \infty$ のとき、$E[a\varphi(\boldsymbol{X}) + b\psi(\boldsymbol{X})] = aE[\varphi(\boldsymbol{X})] + bE[\psi(\boldsymbol{X})].$

$\boldsymbol{X}$ が離散型、絶対連続型のときの計算方法に関する定理を与える。$\varphi : \mathbf{R}^n \to \mathbf{R}$ を Borel 可測関数とする。

定理 4.4 $\boldsymbol{X}$ が離散型確率変数のとき、そのとりうる値を $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$ とすれば、

$$E[\varphi(\boldsymbol{X})] = \sum_{k=1}^{\infty} \varphi(\boldsymbol{x}_k) P(\boldsymbol{X} = \boldsymbol{x}_k). \tag{4.5}$$

証明: $\varphi \geq 0$ とする。§4.1 の 2nd step のように $\varphi_q(x) = \sum_{i=0}^{q2^q} \frac{i}{2^q} 1_{A_{q,i}}(\boldsymbol{X})$ と定義すると、

$$\begin{aligned}
E[\varphi_q(\boldsymbol{X})] &= \sum_{i=0}^{q2^q} \frac{i}{2^q} P(\boldsymbol{X} \in A_{q,i}) = \sum_{i=0}^{q2^q} \frac{i}{2^q} \sum_{k:\boldsymbol{x}_k \in A_{q,i}} P(\boldsymbol{X} = \boldsymbol{x}_k) \\
&= \sum_{i=0}^{q2^q} \sum_{k:\boldsymbol{x}_k \in A_{q,i}} \varphi_q(\boldsymbol{x}_k) P(\boldsymbol{X} = \boldsymbol{x}_k) = \sum_{k=1}^{\infty} \varphi_q(\boldsymbol{x}_k) P(\boldsymbol{X} = \boldsymbol{x}_k)
\end{aligned}$$

ここで、$\{\varphi_q\}$ は非負値単調増加列なので、$\lim_{q\to\infty}\varphi_q(\boldsymbol{x})=\varphi(\boldsymbol{x})$ に注意して単調収束定理 (個数測度として) を用いれば (4.5) を得る。一般の場合は、(4.4) として定義するのだから、$\varphi\geq 0$ の場合より成立する。 □

定理 4.5 $\boldsymbol{X}$ が絶対連続型確率変数のとき、その密度関数を $f(x_1,\ldots,x_p)$ とすれば、

$$E[\varphi(\boldsymbol{X})]=\int\cdots\int_{\mathbf{R}^p}\varphi(x_1,\ldots,x_p)f(x_1,\ldots,x_p)\,dx_1\cdots dx_p. \tag{4.6}$$

証明: 記号を簡単にするため 1 次元確率変数として記す。定理 4.4 の証明と同様に $\varphi\geq 0$ とし、§4.1 の 2nd step の $\varphi_q(x)$ を用いると、

$$\begin{aligned}E[\varphi_q(X)]&=\sum_{i=0}^{q2^q}\frac{i}{2^q}P(X\in A_{q,i})=\sum_{i=0}^{q2^q}\frac{i}{2^q}\int_{A_{q,i}}f(x)\,dx\\&=\sum_{i=0}^{q2^q}\int_{A_{q,i}}\varphi_q(x)f(x)\,dx=\int_{\mathbf{R}}\varphi_q(x)f(x)\,dx\end{aligned}$$

ここで、$\{\varphi_q(x)f(x)\}$ は非負値単調増加列なので、単調収束定理を用いて (4.6) を得る。一般の φ に対しては、(i) と同様に (4.4) として定義するのだから明らか。 □

注意 4.3 X,Y は独立で、X は離散型で取り得る値が $x_1,x_2,\ldots$, Y は絶対連続型でその密度関数が $f_Y(y)$ とする。このとき、

$$P(X=x_k,a<Y\leq b)=P(X=x_k)P(a<Y\leq b)=P(X=x_k)\int_a^b f_Y(y)\,dy$$

に注意すれば、定理 4.4, 4.5 と同様に次が証明できる。

$$E[\varphi(X,Y)]=\sum_{k=1}^{\infty}\int_{-\infty}^{\infty}\varphi(x_k,y)f_Y(y)\,dyP(X=x_k)=\int_{-\infty}^{\infty}\sum_{k=1}^{\infty}.\varphi(x_k,y)P(X=x_k)f_Y(y)\,dy.$$

確率変数の期待値と独立性の関係で、次は重要である。

定理 4.6 確率変数 $X_1,\ldots,X_n$ は独立であるとし、$\varphi_1,\cdots,\varphi_n$ は Borel 可測であるとする。このとき、もし、任意の j で $E[|\varphi_j(X_j)|]<\infty$ であれば、$E[\left|\varphi_1(X_1)\varphi_2(X_2)\cdots\varphi_n(X_n)\right|]<\infty$ であり、次が成立する。

$$E[\varphi_1(X_1)\varphi_2(X_2)\cdots\varphi_n(X_n)]=E[\varphi_1(X)]E[\varphi_2(X_2)]\cdots E[\varphi_n(X_n)]. \tag{4.7}$$

証明: 簡単のため $n=2$ の場合に証明する (一般の場合も同様に証明できる)。

$X_1=X,X_2=Y,\varphi_1=\varphi,\varphi_2=\psi$ と記す。

<u>1st step</u> $\varphi(x),\psi(y)$ が単関数 $\varphi(x)=\sum_{i=1}^M a_i1_{A_i}(x)$, $\psi(y)=\sum_{j=1}^N b_i1_{B_i}(y)$ のとき。$1_A(x)1_B(y)$ も $x\in A$ かつ $y\in B$ のときのみ 1 になる定義関数だから、$\varphi(x)\psi(y)=\sum_{i=1}^M\sum_{j=1}^N a_ib_j1_{A_i}(x)1_{B_j}(y)$ も単関数で、独立性の定義 2.8 とその下の注意により、

$$\begin{aligned}E[\varphi(X)\psi(Y)]&=\sum_{i=1}^M\sum_{j=1}^N a_ib_jP(X\in A_i,Y\in B_j)=\sum_{i=1}^M\sum_{j=1}^N a_ib_jP(X\in A_i)P(Y\in B_j)\\&=\sum_{i=1}^M a_iP(X\in A_i)\sum_{j=1}^N b_jP(Y\in B_j)=E[\varphi(X)]E[\psi(Y)].\end{aligned}$$

<u>2nd step</u> $\varphi\geq 0,\psi\geq 0$ のとき。§4.1 の 2nd step と同様に、各 $q\in\mathbf{N}$ に対して単関数 φ_q,ψ_q を定義する。このとき、1st step により

$$E[\varphi_q(X)\psi_q(Y)]=E[\varphi_q(X)]E[\psi_q(Y)]$$

となる。ここで、単調収束定理 (定理 4.2 (1)) により $q \to \infty$ とすることで (4.7) を得る。特に、$E[|\varphi(X)|] < \infty$ かつ $E[|\psi(Y)|] < \infty$ であれば、$E[|\varphi(X)\psi(Y)|] = E[|\varphi(X)|]E[|\psi(X)|] < \infty$ となることもわかった。

3rd step　一般の φ, ψ の場合。$(\varphi\psi)(x,y) = \varphi(x)\psi(y)$ と書くものとする。このとき、(4.3) のように、$\varphi^+, \varphi^-, \psi^+, \psi^-, (\varphi\psi)^+, (\varphi\psi)^-$ を定める。このとき、

$$(\varphi\psi)^+(x,y) = \varphi^+(x)\psi^+(y) + \varphi^-(x)\psi^-(y), \qquad (\varphi\psi)^-(x,y) = \varphi^+(x)\psi^-(y) + \varphi^-(x)\psi^+(y)$$

に注意する。これより、

$$\begin{aligned}
&E[\varphi(X)]E[\psi(Y)] = (E[\varphi^+(X)] - E[\varphi^-(X)])(E[\psi^+(Y)] - E[\psi^-(Y)]) \\
&\quad = (E[\varphi^+(X)]E[\psi^+(Y)] + E[\varphi^-(X)]E[\psi^-(Y)]) - (E[\varphi^+(X)]E[\psi^-(Y)] + E[\varphi^-(X)]E[\psi^+(Y)]) \\
&\quad = (E[\varphi^+(X)\psi^+(Y)] + E[\varphi^-(X)\psi^-(Y)]) - (E[\varphi^+(X)\psi^-(Y)] + E[\varphi^-(X)\psi^+(Y)]) \\
&\quad = E[(\varphi\psi)^+(X,Y)] - E[(\varphi\psi)^-(X,Y)] = E[\varphi(X)\psi(Y)]. \qquad \square
\end{aligned}$$

注意 4.4 事象 A に対し $E[X, A] = E[X 1_A]$ と定める。特に $E[|\varphi(X)|] < \infty$ で X の密度関数が $f_X(x)$ のとき

$$E[\varphi(X), a \le X \le b] = E[\varphi(X)1_{[a,b]}(X)] = \int_{-\infty}^{\infty} \varphi(x)1_{[a,b]}(x)f_X(x)\,dx = \int_a^b \varphi(x)f_X(x)\,dx, \quad a < b$$

となる。

4.3　積率 (モーメント)・分散

定義 4.1 X を確率変数、$r \in \mathbf{N}$ とする。それぞれの期待値が定義されるとする。
(1)　$E[X^r]$ を r 次の積率 (モーメント) という。特に、$r = 1$ のとき、$E[X]$ を X の平均という。
(2)　$\mu = E[X]$ とするとき、$E[(X-\mu)^r]$ を平均のまわりの r 次の積率という。特に、$r = 2$ のとき、$E[(X-\mu)^2]$ を X の分散といい $V(X)$ と表す。また、$\sigma(X) = \sqrt{V(X)}$ を X の標準偏差という。
(3)　$M_X(t) = E[e^{tX}]$, $t \in \mathbf{R}$, を X の積率母関数という。$\Lambda_X(t) = \log E[e^{tX}]$, $t \in \mathbf{R}$, を X のキュムラント母関数という。

定理 4.7 (Jensen の不等式) $\varphi(x)$ が下に凸で $E[|Y|] < \infty$ であれば、$E[\varphi(Y)] \ge \varphi(E[Y])$ となる。特に、$0 < s < r$ のとき $E[(|X|^s)^{r/s}] \ge (E[|X|^s])^{r/s}$, すなわち、$E[|X|^s] \le (E[|X|^r])^{s/r}$.

証明: $\mu = E[Y]$ とする。φ は下に凸なので a があって $\varphi(x) \ge a(x-\mu) + \varphi(\mu)$ とできる。従って、$E[\varphi(Y)] \ge a(E[Y]-\mu) + \varphi(\mu) = \varphi(\mu)$ と示される。後者は $Y = |X|^s$, $\varphi(x) = |x|^{r/s}$ とせよ。　$\square$

この定理により、$E[X^2] < \infty$ のとき、平均 $E[X]$ も存在する。また、このとき $(X-c)^2 = X^2 - 2cX + c^2$ となることから、分散 $V(X)$ が存在する。

定理 4.8 $E[X^2] < \infty$ とする。このとき次が成り立つ。
(1)　$V(X) = E[X^2] - E[X]^2$.　　(2)　$V(aX+b) = a^2V(X)$ (a, b は定数).
(3)　$V(X) = 0$ ならば $X = \mu$ a.e., ただし $\mu = E[X]$.

証明: (1), (2) は演習問題とする。(3) は定理 4.1 (2), 注意 4.2 から自明。　$\square$

積率母関数に関して Lebesgue の収束定理を用いると、次が示せる。証明は (そんなにも難しくはないが) この授業の範囲を超えるので省略する。

定理 4.9 ある $a, b : a < 0 < b$ があって、$M_X(t) = E[e^{tX}] < \infty$ $(a < \forall t < b)$ とする。このとき $M_X(t)$ は (a, b) 上で C^∞ 級であり $M_X^{(k)} = E[X^k e^{tX}]$, $k \in \mathbb{N}$, となる。

問 4.1 定理 4.9 の仮定の下、X の cumulant 母関数 $\Lambda_X(t) = \log M_X(t)$ について以下を示せ。ただし $\mu = E[X]$ とした。注意: σ を X の標準偏差とするとき、$\Lambda'''_X(0)/\sigma^3$ を**歪度**、$\Lambda^{(4)}_X(0)/\sigma^4$ を**尖度**という。
(1) $\Lambda'_X(0) = \mu,$　　(2) $\Lambda''_X(0) = V(X),$　　(3) $\Lambda'''_X(0) = E[(X-\mu)^3],$
(4) $\Lambda^{(4)}_X(0) = E[(X-\mu)^4] - 3\{V(X)\}^2$
ヒント: $Y = X - \mu$ とおき、$\Lambda_X(t) = \mu t + \log M_Y(t)$ と変形し、$(\Lambda'_X(t) - \mu)M_Y(t) = M'_Y(t)$ を導き、左辺にライプニッツの公式を用いて両辺の微分、2 回微分、3 回微分を計算をすると容易に計算できます。

例 4.1 重要な分布の期待値と分散、積率母関数を例示する。

[A] 離散型確率変数の例
(1) 二項分布 $\mathrm{B}(n,p)$
$P(X=k) = \binom{n}{k} p^k q^{n-k}$, $k = 0, 1, \ldots, n$, $(q = 1-p)$ より

$$\begin{aligned} E[X] &= \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k q^{n-k} = np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-1-(k-1))!} p^{k-1} q^{n-1-(k-1)} \\ &= np(p+q)^{n-1} = np. \end{aligned}$$

また、

$$\begin{aligned} E[X(X-1)] &= \sum_{k=0}^{n} k(k-1) \frac{n!}{k!(n-k)!} p^k q^{n-k} = n(n-1)p^2 \sum_{k=2}^{n} \frac{(n-2)!}{(k-2)!(n-2-(k-2))!} p^{k-2} q^{n-2-(k-2)} \\ &= n(n-1)p^2(p+q)^{n-2} = n(n-1)p^2. \end{aligned}$$

となるから、$E[X^2] = E[X(X-1)+X] = E[X(X-1)] + E[X] = n(n-1)p^2 + np$. よって、$V(X) = E[X^2] - E[X]^2 = np(1-p)$. また、$M_X(t) = (e^t p + 1 - p)^n$, $E[(X - E[X])^3] = np(1-p)(1-2p)$.
(2) Poisson 分布 $\mathrm{Po}(\lambda)$
$P(X=k) = \dfrac{\lambda^k}{k!} e^{-\lambda}$, $k = 0, 1, \ldots$ より

$$E[X] = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda.$$

また、二項分布と同様にして、

$$E[X(X-1)] = \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda^2 e^{-\lambda} e^{\lambda} = \lambda^2.$$

となる。よって、$V(X) = E[X(X-1)] + E[X] - E[X]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$. また、$M_X(t) = e^{\lambda(e^t - 1)}$. ちなみに、$E[(X-\lambda)^3] = \lambda$, $E[(X-\lambda)^4] = \lambda + 3\lambda^2$ となるが、これには積率母関数もしくはキュムラント母関数を用いて計算しないときつい。

問 4.2 $X \sim \mathrm{Po}(\lambda)$ とする。問 4.1 を用いて $E[X]$, $V(X)$, $E[(X-\lambda)^3]$, $E[(X-\lambda)^4]$ の値を求めよ。

問 4.3 X, Y は独立で $X \sim \mathrm{Po}(\alpha)$, $Y \sim \mathrm{Po}(\beta)$ とするとき、$E[X^Y]$ を求めよ。

(3) 幾何分布 $\mathrm{Ge}(p)$
$P(X=k) = (1-p)^k p$, $k = 0, 1, \ldots$, まず、級数 $\sum\limits_{k=0}^{\infty} x^k$ の収束半径は 1 であったから、$|x| < 1$ のときこの級数は何度でも項別微分でき (cf. 系 A.2)、

$$\sum_{k=1}^{\infty} k x^{k-1} = \frac{d}{dx}\Big(\sum_{k=0}^{\infty} x^k\Big) = \frac{d}{dx}\Big(\frac{1}{1-x}\Big) = \frac{1}{(1-x)^2}$$

$$\sum_{k=2}^{\infty} k(k-1)x^{k-2} = \frac{d}{dx}\Big(\sum_{k=1}^{\infty} kx^{k-1}\Big) = \frac{d}{dx}\Big(\frac{1}{(1-x)^2}\Big) = \frac{2}{(1-x)^3}.$$

よって、

$$E[X] = \sum_{k=1}^{\infty} k(1-p)^k p = (1-p)\sum_{k=1}^{\infty} k(1-p)^{k-1} = \frac{p(1-p)}{(1-(1-p))^2} = \frac{1-p}{p},$$

$$V(X) = E[X(X-1)] + E[X] - E[X]^2 = \sum_{k=2}^{\infty} k(k-1)(1-p)^k p + E[X] - E[X]^2$$
$$= \frac{2(1-p)^2 p}{(1-(1-p))^3} + \frac{1-p}{p} - \Big(\frac{1-p}{p}\Big)^2 = \frac{1-p}{p^2}.$$

また、$M_X(t) = \dfrac{p}{1-e^t(1-p)}$, $t < \log\dfrac{1}{1-p}$, であり $E[(X-E[X])^3] = \dfrac{(1-p)^2+1-p}{p^3}$ となる。

問 4.4 (1) $X \sim \mathrm{Ge}(p)$ のとき、$\Lambda(t)$ を求め $E[(X-E[X])^3]$ を計算せよ。また $E[3^{-X}]$, $E[X3^{-X}]$ を求めよ。(2) X, Y は独立でともに $\mathrm{Ge}(p)$ に従うとき $P(\min\{X,Y\} \geq k)$, $k = 0, 1, \ldots$, と $E[\min\{X,Y\}]$ および $E[\max\{X,Y\}]$ を求めよ。ヒント: 最後の期待値は $\max\{x,y\} = x + y - \min\{x,y\}$ を用いよ。

(4) 負の二項分布 $\mathrm{NB}(\alpha, p)$ $(\alpha > 0, 0 < p < 1)$
$P(X = n) = \dbinom{\alpha+n-1}{n} p^\alpha (1-p)^n$, $n = 0, 1, 2, \ldots$.
まず、$\dbinom{\alpha+n-1}{n} = \dfrac{(\alpha+n-1)(\alpha+n-2)\cdots\alpha}{n!} = (-1)^n \dfrac{(-\alpha)(-\alpha-1)\cdots(-\alpha-n+1)}{n!} = (-1)^n \dbinom{-\alpha}{n}$
と $(1+x)^{-\alpha} = \sum_{n=0}^{\infty}\dbinom{-\alpha}{n} x^n$ $(|x| < 1)$ より、$\sum_{n=0}^{\infty} P(X = n) = 1$ となることに注意する。このとき、

$$E[X] = \sum_{n=1}^{\infty} n\binom{\alpha+n-1}{n} p^\alpha (1-p)^n = \sum_{n=1}^{\infty} (-1)^n n\binom{-\alpha}{n} p^\alpha (1-p)^n = p^\alpha \sum_{n=1}^{\infty} n\binom{-\alpha}{n}(p-1)^n$$
$$= p^\alpha(-\alpha)(p-1)\sum_{n=1}^{\infty}\binom{-\alpha-1}{n-1}(p-1)^{n-1} = p^\alpha(-\alpha)(p-1)(1+p-1)^{-\alpha-1} = \frac{\alpha(1-p)}{p}.$$

同様に、$n(n-1)\dbinom{-\alpha}{n} = (-\alpha)(-\alpha-1)\dbinom{-\alpha-2}{n-2}$ より $E[X(X-1)]$ が求まり、$V(X) = \dfrac{\alpha(1-p)}{p^2}$ となる。また、$M_X(t) = \dfrac{p^\alpha}{(1-e^t q)^\alpha}$, $t < \log\dfrac{1}{q}$, であり $E[(X-E[X])^3] = \dfrac{\alpha(q^2+q)}{p^3}$ となる $(q = 1-p$ とした)。

級数 $(1+x)^{-\alpha} = \sum_{n=0}^{\infty}\dbinom{-\alpha}{k} x^n$ の収束半径が 1 であるから、項別微分を利用することで、幾何分布と同様に $E[X]$ や $E[X(X-1)]$ を計算できることにも注意する。

(5) 平均が存在しない例
とり得る値を $x_k = (-1)^{k-1}k$, $k = 1, 2, \ldots$, その確率を $P(X = (-1)^{k-1}k) = \dfrac{1}{k(k+1)}$, $k = 1, 2, \ldots$ とすると、

$$\sum_{k=1}^{\infty} P(X = (-1)^{k-1}k) = \sum_{k=1}^{\infty}\Big(\frac{1}{k} - \frac{1}{k+1}\Big) = 1$$

となり、X は確かに確率変数となるが、X は正の奇数と負の偶数を値にとることに注意して、

$$E[\max\{X, 0\}] = \sum_{l=1}^{\infty}(2l-1)\frac{1}{(2l-1)(2l-1+1)} = \sum_{l=1}^{\infty}\frac{1}{2l} = \infty,$$
$$E[\max\{-X, 0\}] = \sum_{l=1}^{\infty}\{-(-2l)\}\frac{1}{2l(2l+1)} = \sum_{l=1}^{\infty}\frac{1}{2l+1} = \infty$$

となり平均は存在しない。この場合、

$$\sum_{k=1}^{\infty}(-1)^{k-1}k\frac{1}{k(k+1)} = \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{k+1} = 1 - \sum_{l=1}^{\infty}\frac{(-1)^{l-1}}{l} = 1 - \log 2$$

と条件収束し、一見期待値が存在しそうだが、この場合は平均が存在しないとする。

問 4.5 A で X が偶数であるという事象とするとき、$1+(-1)^X = 21_A$ となることを示し、$X \sim \mathrm{B}(n,p)$ のとき X が偶数となる確率を求めよ。また、$X \sim \mathrm{Po}(\lambda)$ と $X \sim \mathrm{NB}(\alpha,p)$ のときも求めよ。

問 4.6 $0<p<1$ とする。確率変数 X が $P(X=k) = c\cdot\dfrac{p^k}{k}$, $k=1,2,\cdots,$ を満たすとき次の問いに答えよ。
(1) $\dfrac{1}{1-x} = \sum\limits_{k=1}^{\infty} x^{k-1}$ に定理 A.3 (2) を用いて定数 c を定めよ。　(この分布を対数級数分布 $\mathrm{LS}(p)$ という。)
(2) 平均 $E[X]$, 分散 $V(X)$ と積率母関数 $M_X(t) = E[e^{tX}]$ を求めよ。
(3) $E[3^{-X}]$ および $E[X3^{-X}]$ を求めよ。

[B] 絶対連続型確率変数の例

(1) 一様分布 $\mathrm{U}(a,b)$

$E[X] = \displaystyle\int_a^b x\frac{1}{b-a}\,dx = \frac{a+b}{2}.$

$V(X) = E[\Big(X-\dfrac{a+b}{2}\Big)^2] = \displaystyle\int_a^b \Big(x-\frac{a+b}{2}\Big)^2\frac{1}{b-a}\,dx = \frac{1}{b-a}\Big[\frac{1}{3}\Big(x-\frac{a+b}{2}\Big)^3\Big]_a^b = \frac{(b-a)^2}{12}.$

問 4.7 X,Y が独立でともに $U(0,1)$ に従うとき $E[X^Y]$ を求めよ。

(2) ガンマ分布 $\Gamma(\alpha,\beta)$

$E[X^c] = \displaystyle\int_0^\infty x^c\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}\,dx = \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty x^{c+\alpha-1}e^{-\beta x}\,dx = \frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\Gamma(c+\alpha)}{\beta^{c+\alpha}} = \frac{\Gamma(c+\alpha)}{\beta^c\Gamma(\alpha)},\ c>-\alpha.$

$E[X] = \dfrac{\beta^\alpha}{\Gamma(\alpha)}\dfrac{\Gamma(1+\alpha)}{\beta^{1+\alpha}} = \dfrac{\alpha}{\beta},\ E[X^2] = \dfrac{\beta^\alpha}{\Gamma(\alpha)}\dfrac{\Gamma(2+\alpha)}{\beta^{2+\alpha}} = \dfrac{(\alpha+1)\alpha}{\beta^2}$ より $V(X) = E[X^2]-(E[X])^2 = \dfrac{\alpha}{\beta^2}.$

$M_X(t) = \displaystyle\int_0^\infty e^{tx}\frac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}\,dx = \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty x^{c+\alpha-1}e^{-(\beta-t)x}\,dx = \frac{\beta^\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha)}{(\beta-t)^\alpha} = \Big(\frac{\beta}{\beta-t}\Big)^\alpha,\ t<\beta.$

指数分布 $Y\sim\mathrm{Ex}(\lambda)$ はガンマ分布で $\alpha=1,\ \beta=\lambda$ の場合だから、

$E[Y] = \dfrac{1}{\lambda},\ V(Y) = \dfrac{1}{\lambda^2},\ M_Y(t) = \dfrac{\lambda}{\lambda-t},\ t<\lambda.$

(3) 正規分布 $\mathrm{N}(\mu,\sigma^2)$

標準正規分布 $Z\sim\mathrm{N}(0,1)$ の場合を考える。

$$\begin{aligned} E[|Z|^\alpha] &= \int_{-\infty}^\infty |z|^\alpha\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\,dz = \frac{2}{\sqrt{2\pi}}\int_0^\infty |z|^\alpha e^{-\frac{z^2}{2}}\,dz \\ &= \frac{2}{\sqrt{2\pi}}\int_0^\infty (2t)^{\frac{\alpha-1}{2}}e^{-t}\,dt = \frac{2^{\alpha/2}}{\sqrt{\pi}}\Gamma\Big(\frac{\alpha+1}{2}\Big). \end{aligned}$$

ただし $\alpha>-1$ とした。$\alpha\le-1$ のときは $E[|Z|^\alpha]=\infty$ である。これより、すべての $n\in\mathbf{N}$ に対して $E[Z^n]$ は存在し、n が奇数のとき密度関数が週関数なので $E[Z^n]=0$ となる。$n=2m$ が偶数のとき

$$E[Z^{2m}] = \frac{2^m}{\sqrt{\pi}}\Gamma\Big(\frac{2m+1}{2}\Big) = \frac{2^m}{\sqrt{\pi}}\frac{2m-1}{2}\frac{2m-3}{2}\cdots\frac{1}{2}\Gamma\Big(\frac{1}{2}\Big) = (2m-1)!!$$

となる。一般の場合 $X\sim\mathrm{N}(\mu,\sigma^2)$ とすると $Z=\dfrac{X-\mu}{\sigma}\sim\mathrm{N}(0,1)$ なので、

$$E[X] = E[\sigma Z+\mu] = \mu, \qquad V(X) = E[(X-\mu)^2] = E[\sigma^2 Z^2] = \sigma^2$$

を得る。また、

$$M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t(\sigma z+\mu)} e^{-\frac{z^2}{2}}\, dz$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{t\mu} e^{-\frac{1}{2}\{(z-t\sigma)^2-(t\sigma)^2\}}\, dz = e^{t\mu+\frac{1}{2}t^2\sigma^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(z-t\sigma)^2}\, dz = e^{t\mu+\frac{1}{2}t^2\sigma^2}.$$

問 4.8 $X \sim \mathrm{N}(\mu, \sigma^2)$ のとき、$Y = e^X$ を考える ($cf.$ 問 2.7 (2))。
(1) $E[Y], V(Y)$ を求めよ。 (2) $P(Y \le y_m) = 1/2$ となる y_m (Y の中央値) を求めよ。
(3) Y の密度関数 $f_Y(y)$ を最大とする点 $y = y_{\max}$ (Y の最頻値) を求めよ。
注意: これより (Y の最頻値)<(Y の中央値)<(Y の平均) となる。所得額や貯蓄額がこの対数正規分布に従うと考えられている。

(4) 自由度 n の t 分布 t_n
T が自由度 n の t 分布 t_n に従うとき、定理 3.6 (1) より $T = Z/\sqrt{\frac{Y}{n}}$, $Z \sim \mathrm{N}(0,1)$, $Y \sim \chi^2_n$, Z と Y は独立と表せる。ここで、独立性と自由度 n のカイ二乗分布 χ^2_n はガンマ分布 $\Gamma(\frac{n}{2}, \frac{1}{2})$ と同じ分布に注意すると

$$E[|T|^\alpha] = E[|Z|^\alpha] n^{\frac{\alpha}{2}} E[Y^{-\frac{\alpha}{2}}] = \frac{2^{\alpha/2}}{\sqrt{\pi}} \Gamma\Big(\frac{\alpha+1}{2}\Big) n^{\frac{\alpha}{2}} \frac{\Gamma(\frac{n-\alpha}{2})}{\Gamma(\frac{n}{2})(1/2)^{-\frac{\alpha}{2}}} = \frac{n^{\frac{\alpha}{2}}\Gamma(\frac{\alpha+1}{2})\Gamma(\frac{n-\alpha}{2})}{\sqrt{\pi}\Gamma(\frac{n}{2})}.$$

よって $k \in \mathbf{N}$ に対して $n \ge k+1$ であれば $E[T^k]$ は定義され、k が奇数なら $E[T^k] = 0$ となる。$n \ge 3$ のとき

$$V(T) = E[T^2] = \frac{n\Gamma(\frac{3}{2})\Gamma(\frac{n-2}{2})}{\sqrt{\pi}\Gamma(\frac{n}{2})} = \frac{n \cdot \frac{1}{2}\Gamma(\frac{1}{2})\Gamma(\frac{n-2}{2})}{\sqrt{\pi}\frac{n-2}{2}\Gamma(\frac{n-2}{2})} = \frac{n}{n-2}$$

となる。$n \ge 5$ のとき $E[T^4] = \dfrac{3n^2}{(n-2)(n-4)}$ となる。また、積率母関数は存在しない。

問 4.9 $n > 4$ とする。$X \sim F^m_n$ ($cf.$ 定理 3.6 (2)) のとき $E[X], V(X)$ を求めよ

問 4.10 事象 A が $P(A) > 0$ を満たすとき、条件つき期待値を $E[X|A] = \dfrac{E[X, A]}{P(A)}$ と定義する。次を求めよ。
(1) $X \sim \mathrm{Ge}(p)$, $0 < p < 1$, のとき $E[X-k|X \ge k]$, $k \in \mathbb{N}$.
(2) $X \sim \mathrm{Ex}(\lambda)$ のとき $E[X-a|X \ge a]$ と $E[(X-a)^2|X \ge a]$. ただし $\lambda > 0$, $a > 0$..
(3) $Y \sim \Gamma(2, \lambda)$ のとき $E[Y-a|Y \ge a]$ と $E[(Y-a)^2|Y \ge a]$. ただし $\lambda > 0$, $a > 0$.

問 4.11 1 から N ($N \ge 3$) までのカードがそれぞれ 1 枚ずつある。それを 3 回復元試行で引き、1 回目、2 回目、3 回目のカードに書かれた数字を X, Y, Z とする。以下を求めよ。
(1) $E[X]$ (2) $V(X)$ (3) $E\Big[\dfrac{1}{X(X+1)}\Big]$ (4) $E[\max\{X-2, 0\}]$ (5) $E[\max\{X, Y\}]$
(6) $E[\min\{X, Y\}]$ (7) $E[\max\{X, Y, Z\}]$ (8) $E[\min\{X, Y, Z\}]$ (9) $E[X|X \ge k]$, $1 \le k \le N$,
(10) $E[X|X \ge Y]$ (11) $E[X|X < Y, X < Z]$ (12) $E[Y|X < Y < Z]$

問 4.12 1 から N ($N \ge 3$) までのカードがそれぞれ 1 枚ずつある。それを 3 回非復元試行で引き、1 回目、2 回目、3 回目のカードに書かれた数字を X, Y, Z とする。以下を求めよ。
(1) $E[XY]$ (2) $E[\max\{X, Y\}]$ (3) $E[\min\{X, Y\}]$ (4) $E[\max\{X, Y, Z\}]$
(5) $E[X|X \ge Y]$ (6) $E[X|X < Y, X < Z]$ (7) $E[Y|X < Y < Z]$ (8) $E[XYZ]$

命題 4.1 確率変数 X が非負整数値であれば、$E[X] = \sum\limits_{n=1}^{\infty} P(X \ge n)$ となる。

証明: $E[X] = \sum\limits_{k=1}^{\infty} kP(X = k) = \sum\limits_{k=1}^{\infty} \sum\limits_{n=1}^{k} P(X = k) = \sum\limits_{n=1}^{\infty} \sum\limits_{k=n}^{\infty} P(X = k) = \sum\limits_{n=1}^{\infty} P(X \ge n).$ □

問 4.13 確率変数 X が非負整数値であれば、$E\Big[\dfrac{X(X+1)}{2}\Big] = \sum_{n=1}^{\infty} nP(X \geq n)$ となることを示せ。

4.4 共分散と相関係数

以下、$L^2(\Omega)$ で、確率変数 X で $E[X^2] < \infty$ を満たすもの全体を表すとする。

定理 4.10 (Cauchy-Schwarz の不等式) $X, Y \in L^2(\Omega)$ のとき、$E[XY]$ は定義され、

$$|E[XY]|^2 \leq E[X^2]E[Y^2] \tag{4.8}$$

が成立する。等号成立のための必要十分条件は、$(a,b) \neq (0,0)$ なる定数 a, b が存在して $aX + bY = 0$ a.e. となることである。

証明: $|XY| \leq \frac{1}{2}(X^2 + Y^2)$ より、$E[|XY|] \leq \frac{1}{2}(E[X^2] + E[Y^2]) < \infty$ となり $E[XY]$ は定義される。$E[X^2] = 0$ のとき、定理 4.1, 注意 4.2 より $X = 0$ a.e. なので (4.8) は $0 = 0$ として明らかに成立し、また、$a = 1$, $b = 0$ として $aX + bY = 0$ a.e. も成立する。$E[X^2] > 0$ のとき、

$$0 \leq E[(tX+Y)^2] = t^2E[X^2] + 2tE[XY] + E[Y^2] = E[X^2]\Big\{t + \frac{E[XY]}{E[X^2]}\Big\}^2 - \frac{(E[XY])^2}{E[X^2]} + E[Y^2] \tag{4.9}$$

が任意の t に対して成立するので、(4.8) は成立する。一方、(4.8) で等号が成立するとき、(4.9) で $t = -E[XY]/E[X^2]$ ととれば $E[(tX+Y)^2] = 0$ となる。よって、$tX + Y = 0$ a.e. を得る。逆に、$aX + bY = 0$ a.e. とすると、$P(X \neq 0) > 0$ より $b \neq 0$ となることに注意する。このとき、$Y = aX/b$ a.e. より

$$|E[XY]|^2 = \Big(\frac{a}{b}\Big)^2 E[X^2]^2, \qquad E[X^2]E[Y^2] = E[X^2]\Big(\frac{a}{b}\Big)^2 E[X^2]$$

となるので、(4.8) の等号が成立する。 □

系 4.1 $X, Y \in L^2(\Omega)$ のとき $E[(X-\mu_1)(Y-\mu_2)]$ は定義され、

$$|E[(X-\mu_1)(Y-\mu_2)]|^2 \leq V(X)V(Y) \tag{4.10}$$

が成立する。ただし、$E[X] = \mu_1$, $E[Y] = \mu_2$ とした。(4.10) で等号成立のための必要十分条件は、$(a,b) \neq (0,0)$ なる定数 a, b が存在して $a(X-\mu_1) + b(Y-\mu_2) = 0$ a.e. となることである。

証明: 定理 4.8 より $X, Y \in L^2(\Omega)$ のとき $X - \mu_1, Y - \mu_2 \in L^2(\Omega)$ に注意する。よって、定理 4.10 の X, Y を $|X-\mu_1|, |Y-\mu_2|$ に置き換えると、

$$E[|(X-\mu_1)(Y-\mu_2)|]^2 \leq E[(X-\mu_1)^2]E[(Y-\mu_2)^2] < \infty$$

より、$E[(X-\mu_1)(Y-\mu_2)]$ は定義される。他は、定理 4.10 の X, Y を $X - \mu_1, Y - \mu_2$ に置き換えれよい。□

定義 4.2 $X, Y \in L^2(\Omega)$ なる確率変数 X, Y の共分散 $\mathrm{Cov}(X,Y)$、相関係数 $\rho(X,Y)$ を次で定義する。

$$\begin{aligned}\mathrm{Cov}(X,Y) &= E[(X-\mu_1)(Y-\mu_2)] = E[XY] - \mu_1E[X] - \mu_2E[Y] + \mu_1\mu_2 = E[XY] - \mu_1\mu_2, \\ \rho(X,Y) &= \frac{\mathrm{Cov}(X,Y)}{\sqrt{V(X)V(Y)}}\end{aligned}$$

ただし、$\mu_1 = E[X], \mu_2 = E[Y]$ とし、相関係数は $V(X)V(Y) > 0$ のときのみに定義されるものとする。

定理 4.11 $X, Y \in L^2(\Omega)$ のとき、相関係数 $\rho(X,Y)$ について、$-1 \leq \rho(X,Y) \leq 1$ であり、

$$\rho(X,Y) = 1 \text{ のとき } \frac{X-\mu_1}{\sqrt{V(X)}} = \frac{Y-\mu_2}{\sqrt{V(Y)}} \text{ a.e.,} \quad \rho(X,Y) = -1 \text{ のとき } \frac{X-\mu_1}{\sqrt{V(X)}} = -\frac{Y-\mu_2}{\sqrt{V(Y)}} \text{ a.e.}$$

証明: $-1 \le \rho(X,Y) \le 1$ は定義と系 4.1 より明らか。$\rho(X,Y) = \pm 1$ のとき系 4.1 により $(a,b) \ne (0,0)$ があって $a(X-\mu_1) + b(Y-\mu_2) = 0$ a.e. となる。従って、$a^2(X-\mu_1)^2 = b^2(Y-\mu_2)^2$ の期待値をとって、$a^2V(X) = b^2V(Y)$. ここで、$V(X)V(Y) > 0$ より a,b はともに 0 ではことがわかる。(一方が 0 なら双方とも 0 となるため。) よって、$Y-\mu_2 = -a(X-\mu_1)/b$ a.e. より、

$$V(Y) = \left(\frac{a}{b}\right)^2 V(X), \quad \rho(X,Y) = \frac{E[\frac{-a}{b}(X-\mu_1)^2]}{V(X)^{1/2}|\frac{a}{b}|V(X)^{1/2}} = -\frac{\frac{a}{b}}{|\frac{a}{b}|}.$$

よって、$\rho(X,Y) = 1$ なら、$\dfrac{a}{b} < 0$ で $\dfrac{a}{b} = -\dfrac{\sqrt{V(Y)}}{\sqrt{V(X)}}$, $\rho(X,Y) = -1$ なら、$\dfrac{a}{b} > 0$ で $\dfrac{a}{b} = -\dfrac{\sqrt{V(Y)}}{\sqrt{V(X)}}$ となり主張を得る。 □

例題 4.1 同時密度関数が $f(x,y) = \begin{cases} 24xy & (0 \le x, 0 \le y, x+y \le 1) \\ 0 & (\text{その他}) \end{cases}$ である確率変数 X,Y について、その共分散 $\mathrm{Cov}(X,Y)$ と相関係数 $\rho(X,Y)$ を求める。

解 $f_X(x)$, $f_Y(y)$ を X,Y の周辺密度関数とする。$0 \le x \le 1$ のとき、

$$f_X(x) = \int_0^{1-x} 24xy\,dy = 12x(1-x)^2$$

より、$f_X(x) = \begin{cases} 12x(1-x)^2 & (0 \le x \le 1) \\ 0 & (\text{その他}) \end{cases}$. 全く同様に、$f_Y(y) = \begin{cases} 12y(1-y)^2 & (0 \le y \le 1) \\ 0 & (\text{その他}) \end{cases}$ である。よって、補題 2.3 を用いて計算すると、

$$E[X] = \int_0^1 x\,12x(1-x)^2\,dx = 12\int_0^1 x^2(1-x)^2\,dx = 12\,\mathrm{B}(3,3) = 12\frac{\Gamma(3)\Gamma(3)}{\Gamma(6)} = 12\frac{2\cdot 2}{5!} = \frac{2}{5}\ (= E[Y])$$
$$E[X^2] = \int_0^1 x^2\,12x(1-x)^2\,dx = 12\int_0^1 x^3(1-x)^2\,dx = 12\,\mathrm{B}(4,3) = 12\frac{\Gamma(4)\Gamma(3)}{\Gamma(7)} = 12\frac{3!\cdot 2!}{6!} = \frac{1}{5} \text{ より}$$
$$V(X) = E[X^2] - E[X]^2 = \frac{1}{5} - \left(\frac{2}{5}\right)^2 = \frac{1}{25}\ (= V(Y)).$$

(注意: $f_X(x)$ を求めず、$E[\varphi(X)] = \displaystyle\iint_D \varphi(x)\,24xy\,dxdy$ から求めてもよい。) 一方 $D:0 \le y \le 1-x,\ 0 \le x \le 1$ とすると、

$$\begin{aligned} E[XY] &= \iint_D xy\,24xy\,dxdy = \int_0^1 dx \int_0^{1-x} 24x^2y^2\,dy = \int_0^1 \Big[8x^2y^3\Big]_{y=0}^{y=1-x} dx \\ &= \int_0^1 8x^2(1-x)^3\,dx = 8\,\mathrm{B}(3,4) = \cdots = \frac{2}{15} \text{ より} \\ \mathrm{Cov}(X,Y) &= E[XY] - E[X]E[Y] = \frac{2}{15} - \left(\frac{2}{5}\right)^2 = -\frac{2}{75} \\ \rho(X,Y) &= \frac{\mathrm{Cov}(X,Y)}{\sqrt{V(X)V(Y)}} = \frac{-2/75}{1/25} = -\frac{2}{3}. \qquad \square \end{aligned}$$

問 4.14 (1) (X,Y) の同時密度関数が $f(x,y) = \begin{cases} 3y & (0 \le x, 0 \le y, x^2+y^2 \le 1) \\ 0 & (\text{その他}) \end{cases}$ であるとき、$\mathrm{Cov}(X,Y)$ を求めよ。

(2) (X,Y) の同時密度関数が $f(x,y) = \begin{cases} 2xy & (0 \le y \le 2x \le 2) \\ 0 & (\text{その他}) \end{cases}$ であるとき、$E[X^2]$ と $E[XY]$ を求めよ。

問 4.15 $X \sim U(0,10)$ とし $Y = \lfloor X \rfloor$ とするとき、$\mathrm{Cov}(X,Y)$ を求めよ ($\lfloor a \rfloor$ は a 以下の最大整数)。

命題 4.2 $X, Y, X_k \in L^2(\Omega)$, $k = 1, 2, \dots, n$, と定数 $a, b, a_k \in \mathbf{R}$, $k = 1, 2, \dots, n$, に対して次が成立する。

(0) $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$,

(i) $\mathrm{Cov}\Big(\sum_{k=1}^{n} a_k X_k, Y\Big) = \sum_{k=1}^{n} a_k \mathrm{Cov}(X_k, Y)$,

(ii) $V\Big(\sum_{k=1}^{n} a_k X_k\Big) = \sum_{k=1}^{n} {a_k}^2 V(X_k) + 2 \sum_{1 \le k < l < n} a_k a_l \mathrm{Cov}(X_k, X_l)$,

特に $V(aX + bY) = a^2 V(X) + 2ab \mathrm{Cov}(X, Y) + b^2 V(Y)$.

証明: (0) は明らか。(i) は $\widetilde{X_k} = X_k - E[X_k]$, $\widetilde{Y} = Y - E[Y]$ とすると、

$$\begin{aligned}\mathrm{Cov}\Big(\sum_{k=1}^{n} a_k X_k, Y\Big) &= E\Big[\Big\{\sum_{k=1}^{n} a_k X_k - E\Big[\sum_{k=1}^{n} a_k X_k\Big]\Big\}(Y - E[Y])\Big] \\ &= E\Big[\sum_{k=1}^{n} a_k \tilde{X_k} \widetilde{Y}\Big] = \sum_{k=1}^{n} a_k E[\widetilde{X_k} \widetilde{Y}] = \sum_{k=1}^{n} a_k \mathrm{Cov}(X_k, Y).\end{aligned}$$

(ii) は $\Big(\sum_{k=1}^{n} a_k \tilde{X}_k\Big)^2 = \sum_{k=1}^{n} (a_k \tilde{X}_k)^2 + 2 \sum_{1 \le k < l < n} a_k \tilde{X}_k a_l \tilde{X}_l$ に注意すれば (i) と同様に示せる。 □

定理 4.12 確率変数 X, Y について、次の (i)–(iv) は同値である。

(i) $\mathrm{Cov}(X, Y) = 0$. (ii) X と Y は無相関、即ち、$\rho(X, Y) = 0$ となる。

(iii) $E[XY] = E[X]E[Y]$. (iv) $V(X + Y) = V(X) + V(Y)$.

証明: (i) $\Longleftrightarrow$ (ii): 定義より明らか。

(i) $\Longleftrightarrow$ (iii): $\mathrm{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$ よりわかる。

(i) $\Longleftrightarrow$ (iv): $V(X + Y) = V(X) + V(Y) + 2\mathrm{Cov}(X, Y)$ より従う。 □

注意 4.5 X, Y が独立であれば、無相関である。実際、X, Y が独立であれば、定理 4.6 より (iii) が成立する。よって、定理 4.12 より無相関となる。

逆は成立しない。例えば、$0 < q < 1$ とし、$P((X, Y) = (0, 1)) = q$, $P((X, Y) = (1, 0)) = P((X, Y) = (-1, 0)) = (1 - q)/2$ となる確率変数 X, Y を考えると、

$$P(X = 0) = P(Y = 1) = P((X, Y) = (0, 1)) = q$$

より $P(X = 0, Y = 1) = P((X, Y) = (0, 1)) \neq P(X = 0)P(Y = 1)$ となり X と Y は独立ではない。

一方、$P(XY = 0) = P((X, Y) = (0, 1)) + P((X, Y) = (1, 0)) + P((X, Y) = (-1, 0)) = 1$ より $E[XY] = 0$, $P(X = 1) = P((X, Y) = (1, 0)) = (1 - q)/2, P(X = -1) = P((X, Y) = (-1, 0)) = (1 - q)/2$ より $E[X] = 1 \cdot P(X = 1) + (-1) \cdot P(X = -1) + 0 \cdot P(X = 0) = 0$ となり、$E[XY] = E[X]E[Y]$ が成立するので $\mathrm{Cov}(X, Y) = \rho(X, Y) = 0$.

ただし、(X, Y) が 2 次元正規分布であれば X, Y が独立であることと、無相関であることが同値となる (*cf.* 例 2.5, 定理 4.13, 系 4.2)。

例 4.2 2 次元確率変数 (X, Y) が次のような多項分布に従うとき、共分散 $\mathrm{Cov}(X, Y)$ を求めよ:

$$P(X = k, Y = l) = \frac{n!}{k!l!(n - k - l)!} p^k q^l r^{n-k-l}, \quad k, l \in \mathbb{Z}; 0 \le k, l,\ k + l \le n. \tag{4.11}$$

ただし、$0 < p, q, r < 1$, $p + q + r = 1$ とする。

解: X, Y の周辺分布がそれぞれ二項分布 $\mathrm{B}(n, p)$, $\mathrm{B}(n, q)$ に従うから、$E[X] = np$, $E[Y] = nq$. また、

$$E[XY] = \sum_{k=0}^{n} \sum_{l=0}^{n-k} kl \frac{n!}{k!l!(n - k - l)!} p^k q^l r^{n-k-l} = \sum_{k=0}^{n} k \frac{n!}{k!(n - k)!} p^k \sum_{l=1}^{n-k} l \frac{(n - k)!}{l!(n - k - l)!} q^l r^{n-k-l}$$

$$
\begin{aligned}
&= \sum_{k=0}^{n} k \frac{n!}{k!(n-k)!} p^k \sum_{l=1}^{n-k} (n-k) q \binom{n-k-1}{l-1} q^{l-1} r^{n-k-l} \\
&= \sum_{k=0}^{n} k(n-k) \frac{n!}{k!(n-k)!} p^k q (q+r)^{n-k-1} = n(n-1)pq \sum_{k=1}^{n-1} \binom{n-2}{k-1} p^{k-1} (q+r)^{n-k-1} \\
&= n(n-1)pq(p+q+r)^{n-2} = n(n-1)pq.
\end{aligned}
$$

よって、$\mathrm{Cov}(X,Y) = -npq.$ □

相関係数について $\rho(X,Y) > 0$ のとき X と Y は正の相関がある、$\rho(X,Y) < 0$ のとき負の相関があるという。X と Y は正の相関があるのとき X が大きくなれば Y も大きくなる、負の相関があるとき X が大きくなれば Y は小さくなるという傾向を持つ。これを以下の例で見よう。

例 4.3 n 回サイコロを投げるとき、i の目の出た回数を X_i とし、奇数の目の出た回数を Y, 偶数の目の出た回数を Z とある。相関係数 $\rho(X_1, Y)$, $\rho(X_1, Z)$ を求めよ。また、$V(X_1 - X_2)$ を求めよ。

解: $i \neq j$ とすると、X_i, X_j は (4.11) を $p = q = 1/6$ で満たす。よって、$\mathrm{Cov}(X_i, X_j) = -n/36$ となる。また、X_i は $\mathrm{B}(n, \frac{1}{6})$ に従うから $E[X_i] = n/6$, $V(X_i) = 5n/36$ である。よって、$Y = X_1 + X_3 + X_5$, $Z = X_2 + X_4 + X_6$ であるから、

$$
\begin{aligned}
\mathrm{Cov}(X_1, Y) &= \mathrm{Cov}(X_1, X_1) + \mathrm{Cov}(X_1, X_3) + \mathrm{Cov}(X_1, X_5) = \frac{5n}{36} - \frac{n}{36} - \frac{n}{36} = \frac{n}{12}, \\
\mathrm{Cov}(X_1, Z) &= \mathrm{Cov}(X_1, X_2) + \mathrm{Cov}(X_1, X_4) + \mathrm{Cov}(X_1, X_6) = -\frac{n}{36} - \frac{n}{36} - \frac{n}{36} = -\frac{n}{12}
\end{aligned}
$$

また、Y, Z はともに $\mathrm{B}(n, \frac{1}{2})$ に従うから、$V(Y) = V(Z) = \dfrac{n}{4}$. よって、$\rho(X_1, Y) = \dfrac{1}{\sqrt{5}}$, $\rho(X_1, Z) = -\dfrac{1}{\sqrt{5}}$ となる。一方、

$$
\begin{aligned}
V(X_1 - X_2) &= \mathrm{Cov}(X_1 - X_2, X_1 - X_2) = \mathrm{Cov}(X_1 - X_2, X_1) - \mathrm{Cov}(X_1 - X_2, X_2) \\
&= \mathrm{Cov}(X_1, X_1) - \mathrm{Cov}(X_2, X_1) - \mathrm{Cov}(X_1, X_2) + \mathrm{Cov}(X_2, X_2) \\
&= V(X_1) - 2\,\mathrm{Cov}(X_1, X_2) + V(X_2) = \frac{5n}{36} - 2\Big(-\frac{n}{36}\Big) + \frac{5n}{36} = \frac{n}{3} \qquad \square
\end{aligned}
$$

問 4.16 公正なサイコロを n 回投げたとき、1 の目の出た回数を X とし偶数の目が出た回数を Y とするとき、$M(s,t) = E[e^{sX+tY}]$, $s, t \in \mathbf{R}$, を求めよ。ヒント: $(a+b+c)^n$ の展開式を考えよ。

問 4.17 1 から n までの数字が 1 つだけ書かれたカードがそれぞれ 1 枚ずつ計 n 枚ある。このカードをよく切って 2 枚取出し、書かれた数字のうち、小さいほうを X, 大きいほうを Y とする。
(1) k, ℓ を $1 \le k < \ell \le n$ となる自然数とするとき、$P(X = k, Y = \ell)$ を求めよ。さらに、$\ell = 2, \ldots, n$ とするとき、$P(Y = \ell)$ を求めよ。
(2) $\mathrm{Cov}(X, Y)$ を求めよ。

問 4.18 赤玉 M 個と白玉 N 個をよく混ぜたあと、一つずつ空になるまで取り出す試行を考える。ここで、$M \ge 2, N \ge 2$ とする。
(1) $1 \le j \le M + N$ に対して確率変数 X_j を j 番目に取り出した玉が赤だったら 1, 白だったら 0 と定める。このとき、$1 \le i \le N + M$ に対して $P(X_i = 1)$, $E[X_i]$ と分散 $V(X_i)$ を、$1 \le i < j \le N + M$ に対して $P(X_i = 1, X_j = 1)$, $E[X_i X_j]$ と共分散 $\mathrm{Cov}(X_i, X_j)$ を求めよ。
(2) $3 \le L \le M + N$ とする。最初の L 個に含まれる赤玉の数を Y とするとき、$E[Y]$ と $V(Y)$ を求めよ。

定義 4.3 n 次元確率変数 $\boldsymbol{X} = (X_1, \ldots, X_n)$ について、ベクトル $E[\boldsymbol{X}] = (E[X_1], \ldots, E[X_n])$ を $\boldsymbol{X}$ の平均ベクトルという。また、$\mathrm{Cov}(X_i, X_j)$ をその (i, j) 成分とする n 次対称行列を、$\boldsymbol{X}$ の共分散行列といい、$V(\boldsymbol{X})$ で表す。

定理 4.13 $\boldsymbol{X} = (X_1, \ldots, X_n)'$ を n 次元正規分布 $\mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ に従うとする。$\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)' \in \mathbf{R}^n$, $\boldsymbol{\Sigma} = (\sigma_{ij})$ は正定値対称行列であった。このとき、$E[\boldsymbol{X}] = \boldsymbol{\mu}$, $V(\boldsymbol{X}) = \boldsymbol{\Sigma}$ となる。

証明: $E[X_k] = \mu_k$, $1 \le k \le n$, $\mathrm{Cov}(X_k, X_l) = \sigma_{kl}$, $1 \le k, l \le n$ を示せばよい。例 2.3 (ii) の記号を用いと、$\boldsymbol{X} = (X_1, \ldots, X_n)'$ の密度関数は

$$f(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}(\det \boldsymbol{\Sigma})^{1/2}} e^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}, \quad \boldsymbol{x} = (x_1, \ldots, x_n)'.$$

$\boldsymbol{\Sigma}$ は正定値対称行列なので、線形代数の対角化に関する定理から、直交行列 $P = (p_{ij})$ と対角成分がすべて正の対角行列 $D = (\lambda_{ij})$ がとれて、$P'\boldsymbol{\Sigma}P = D$ とできる。このとき、$\boldsymbol{Y} = (Y_1, \ldots, Y_n)' = P'(\boldsymbol{X} - \boldsymbol{\mu})$ とし、この密度関数 $g(y_1, \ldots, y_n)$ を求めよう。まず、$\boldsymbol{y} = (y_1, \ldots, y_n)' = P'(\boldsymbol{x} - \boldsymbol{\mu})$ とすると、

$$(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = \boldsymbol{y}'P'\boldsymbol{\Sigma}^{-1}P\boldsymbol{y} = \boldsymbol{y}'D^{-1}\boldsymbol{y} = \sum_{j=1}^{n} \frac{1}{\lambda_{jj}} y_j^2.$$

また、$\boldsymbol{x} = P\boldsymbol{y} + \boldsymbol{\mu}$ より $\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{y}} = \det P = \pm 1$, $\det \boldsymbol{\Sigma} = \det D = \prod_{j=1}^{n} \lambda_{jj}$. よって、

$$g(y_1, \ldots, y_n) = \frac{1}{(2\pi)^{n/2}(\lambda_{11}\cdots\lambda_{nn})^{1/2}} e^{-\frac{1}{2}\sum_{j=1}^{n}\frac{1}{\lambda_{jj}}y_j^2} \left|\frac{\partial \boldsymbol{x}}{\partial \boldsymbol{y}}\right| = \prod_{j=1}^{n} \frac{1}{(2\pi\lambda_{jj})^{1/2}} e^{-\frac{1}{2}\frac{y_j^2}{\lambda_{jj}}}.$$

これは、$Y_1, \ldots, Y_n$ が独立で、各 Y_j は正規分布 $\mathrm{N}(0, \lambda_{jj})$ に従うことを意味している。特に、

$$E[Y_i] = 0,\ E[Y_i^2] = \lambda_{ii},\ 1 \le i \le n, \qquad E[Y_iY_j] = E[Y_i]E[Y_j] = 0,\ 1 \le i < j \le n.$$

今、$\boldsymbol{X} = P\boldsymbol{Y} + \boldsymbol{\mu}$ より、$X_k = \sum_{i=1}^{n} p_{ki}Y_i + \mu_k$ なので、

$$E[X_k] = \sum_{i=1}^{n} p_{ki}E[Y_i] + \mu_k = \mu_k,$$

$$\mathrm{Cov}(X_k, X_l) = E[(X_k - \mu_k)(X_l - \mu_l)] = \sum_{i=1}^{n}\sum_{j=1}^{n} p_{ki}p_{lj}E[Y_iY_j] = \sum_{i=1}^{n} p_{ki}p_{li}\lambda_{ii} = \sigma_{kl}$$

を得る。最後の等号は $\boldsymbol{\Sigma} = PDP'$ であるが、DP' の (i,l)-成分が $\lambda_{ii}p_{li}$ であるから、PDP' の (k,l)-成分が $\sum_{i=1}^{n} p_{ki}\lambda_{ii}p_{li}$ となることを用いた。 □

系 4.2 $\boldsymbol{X} = (X_1, \ldots, X_n)'$ を n 次元正規分布 $\mathrm{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ に従うとする。このとき、$X_1, \ldots, X_n$ が独立であるための必要十分条件は $\boldsymbol{\Sigma}$ が対角行列であることである。

証明: 必要性は例 2.5 で示した。十分性について。$\boldsymbol{\Sigma} = (\sigma_{ij})$ とかく。$i \ne j$ ならば、X_i, X_j は独立なので定理 4.12 により $\mathrm{Cov}(X_i, X_j) = 0$. よって、定理 4.13 により $\sigma_{ij} = \mathrm{Cov}(X_i, X_j) = 0$ となる。 □

例 4.4 2 次元確率変数 (X, Y) の同時密度関数が $f(x,y) = \dfrac{1}{\pi}\exp\Big\{-\dfrac{1}{2}Q(x,y)\Big\}$, $(x,y) \in \mathbf{R}^2$, であるとする。ただし、

$$Q(x,y) = x^2 + 2xy + 5y^2 - 2x + 2y + 2$$

とする。このとき、$V(X)$, $V(Y)$ と $\mathrm{Cov}(X,Y)$ を求めよ。

解: $Q(x,y) = (x+y-1)^2 - (y-1)^2 + 5y^2 + 2y + 2 = (x+y-1)^2 + 4(y+\frac{1}{2})^2$ に注意して、$S = X + Y - 1$, $T = 2(Y + \frac{1}{2})$ とおく。このとき、(S,T) の同時密度関数 $g(s,t)$ を求めると、$s = x + y - 1, t = 2(y + \frac{1}{2})$ より $y = \frac{1}{2}(t-1), x = s - \frac{1}{2}t + \frac{3}{2}$ に注意して、

$$g(s,t) = f\Big(s - \frac{1}{2}t + \frac{3}{2}, \frac{1}{2}(t-1)\Big)\left|\frac{\partial(x,y)}{\partial(s,t)}\right| = \frac{1}{\pi}\exp\Big\{-\frac{1}{2}(s^2+t^2)\Big\}\frac{1}{2} = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}s^2}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}t^2}.$$

これは、S,T は独立でともに標準正規分布 $\mathrm{N}(0,1)$ に従うことを示している。よって、

$$E[S]=E[T]=0, \qquad E[S^2]=V(S)=E[T^2]=V(T)=1, \qquad E[ST]=\mathrm{Cov}(S,T)=0$$

となる。これより、$Y=\frac{1}{2}(T-1), X=S-\frac{1}{2}T+\frac{3}{2}$ より、

$$\begin{aligned}
E[Y]&=\frac{1}{2}\left(E[T]-1\right)=-\frac{1}{2}, \qquad E[X]=E[S]-\frac{1}{2}E[T]+\frac{3}{2}=\frac{3}{2},\\
V(Y)&=E[(Y-E[Y])^2]=\frac{1}{4}E[T^2]=\frac{1}{4},\\
V(X)&=E[(X-E[X])^2]=E\Big[\Big(S-\frac{1}{2}T+\frac{3}{2}-\frac{3}{2}\Big)^2\Big]=E[S^2]-2\cdot\frac{1}{2}E[ST]+\frac{1}{4}E[T^2]=\frac{5}{4},\\
\mathrm{Cov}(X,Y)&=E[(X-E[X])\,(Y-E[Y])]=E\Big[\Big(S-\frac{1}{2}T\Big)\frac{1}{2}T\Big]=\frac{1}{2}E[ST]-\frac{1}{4}E[T^2]=-\frac{1}{4}. \qquad \Box
\end{aligned}$$

問 4.19 (X,Y) の同時密度関数が $f(x,y)=K\exp\Big\{-\dfrac{1}{2}Q(x,y)\Big\}$, $(x,y)\in\mathbf{R}^2$, であるとする。ただし、$Q(x,y)=3x^2+3xy+2y^2-6x+2y$ とする。

(1) 定数 $a,b,c,\alpha,\beta,\gamma$ を $Q(x,y)=a(x+by+c)^2+\alpha(y+\beta)^2+\gamma$ を満たすように定め、確率変数 S,T を $S=\sqrt{a}(X+bY+c)$, $T=\sqrt{\alpha}(Y+\beta)$ とする。このとき、定数 K を定め、(S,T) の同時密度関数 $g(s,t)$ を求めることで、S,T が独立でそれぞれ標準正規分布 $\mathrm{N}(0,1)$ に従うことを示せ。

(2) 平均 $E[Y]$, $E[X]$ と共分散 $\mathrm{Cov}(X,Y)$ を求めよ。

4.5 条件つき期待値

$P(Y\le y|X=x)$ がすべての y で定義されていれば、条件つき期待値 $E[h(Y)|X=x]$ が定義できる。例えば、条件 $X=x$ の下での Y の条件つき密度関数 $f_{Y|X}(y|x)$ がわかっていれば

$$E[h(Y)|X=x]=\int_{-\infty}^{\infty}h(y)f_{Y|X}(y|x)\,dx$$

と、$X=x$ の下で Y のとり得る値が離散型で $y_1,y_2,\dots$ のとき $E[h(Y)|X=x]=\sum_{l=1}^{\infty}h(y_l)P(Y=y_l|X=x)$ と定義される。さらに、条件つき分散を $V(Y|X=x)=E[(Y-E[Y|X=x])^2|X=x]=E[Y^2|X=x]-(E[Y|X=x])^2$ で定める。また、$E[h(Y)|X=x]$ は x の関数であるがそれを $\psi(x)$ で表すとき $\psi(X)$ は確率変数であるがこれを単に $E[h(Y)|X]$ と表す。同様に $V(Y|X)$ も定義され、これは確率変数となる。

例題 4.2 (1) (X,Y) の密度関数が $f(x,y)=4e^{-2x-y}$ $(0\le 2x\le y)$, $f(x,y)=0$ (その他) とする。$x>0$ のときの $f_{Y|X}(y|x)$ 及び $E[Y|X]$, $V(Y|X)$ を求めよ。

(2) Λ は $\Gamma(\alpha,\beta)$ に従い、$\Lambda=\lambda$ の条件のもと Y が $\mathrm{N}(0,1/\lambda)$ に従うとき (*cf*. 例題 2.5)、$E[\Lambda|Y]$ および $V(\Lambda|Y)$ を求めよ。

解: (1) $f_X(x)=\displaystyle\int_{2x}^{\infty}4e^{-2x-y}\,dy=4e^{-4x}$ より、$f_{Y|X}(y|x)=\dfrac{4e^{-2x-y}}{4e^{-4x}}=e^{2x-y}$ $(y\ge 2x)$, $f_{Y|X}(y|x)=0$ (その他). よって、$x>0$ のとき $t=y-2x$ とおくと

$$\begin{aligned}
E[Y|X=x]&=\int_{-\infty}^{\infty}yf_{Y|X}(y|x)\,dy=\int_{2x}^{\infty}ye^{2x-y}\,dy=\int_0^{\infty}(t+2x)e^{-t}\,dt=1+2x,\\
E[Y^2|X=x]&=\int_0^{\infty}(t+2x)^2e^{-t}\,dt=\Big[-\{(t+2x)^2+2(t+2x)+2\}e^{-t}\Big]_0^{\infty}=4x^2+4x+2,\\
V(Y|X=x)&=E[Y^2|X=x]-(E[Y|X=x]^2)=4x^2+4x+2-(1+2x)^2=1
\end{aligned}$$

より、$E[Y|X]=2X+1$, $V(Y|X)=1$.

(2) 例題 2.5 より $\lambda > 0$ のとき

$$f_{\Lambda|Y}(\lambda|y) = \frac{\left(\beta + \frac{1}{2}y^2\right)^{\alpha+\frac{1}{2}}}{\Gamma(\alpha+\frac{1}{2})}\lambda^{\alpha+\frac{1}{2}-1}e^{-(\beta+\frac{1}{2}y^2)\lambda},$$

で $\lambda \leq 0$ のとき $f_{\Lambda|Y}(\lambda|y) = 0$, 即ち、$Y = y$ の条件のもと Λ はガンマ分布 $\Gamma\left(\alpha + \frac{1}{2}, \beta + \frac{1}{2}y^2\right)$ に従う。よって、

$$E[\Lambda|Y = y] = \frac{\alpha + \frac{1}{2}}{\beta + \frac{1}{2}y^2}, \qquad V(\Lambda|Y = y) = \frac{\alpha + \frac{1}{2}}{(\beta + \frac{1}{2}y^2)^2}.$$

従って、$E[\Lambda|Y] = \dfrac{\alpha + \frac{1}{2}}{\beta + \frac{1}{2}Y^2}$, $V(\Lambda|Y) = \dfrac{\alpha + \frac{1}{2}}{(\beta + \frac{1}{2}Y^2)^2}$. $\square$

問 4.20 (X, Y) の同時密度関数が $f(x,y) = \begin{cases} 3(x+y) & (0 \leq x, 0 \leq y, x+y \leq 1) \\ 0 & (\text{その他}) \end{cases}$ であるとする。

(1) X の周辺密度関数 $f_X(x)$ と $E[X]$ を求めよ。

(2) $0 < x < 1$ のとき $f_{Y|X}(y|x)$ と $E[Y|X = x]$ を求めよ。

(3) $E[E[Y|X]]\left(= \displaystyle\int_{-\infty}^{\infty} E[Y|X = x] f_X(x)\,dx\right)$ を求めよ。

問 4.21 (X, Y) の同時密度関数が $f(x,y) = \begin{cases} \frac{4}{\pi}xe^{-(1+y^2)x} & (x > 0, y > 0) \\ 0 & (\text{その他}) \end{cases}$ であるとする。

(1) Y の周辺密度関数 $f_Y(y)$ と $E[Y]$ を求めよ。 (2) $E[XY]$ を求めよ。

(3) $y > 0$ に対して、$f_{X|Y}(x|y)$ と $E[X|Y = y]$ を求めよ。

問 4.22 X はガンマ分布 $\Gamma(\alpha, 1)$ に従い、$X = x$ の下で $Y \sim \mathrm{Ex}(x)$ とするとき、以下を求めよ。

(1) (X, Y) の同時密度関数 $f(x,y)$ と Y の周辺密度関数 $f_Y(y)$ を求めよ。

(2) $f_{X|Y}(x|y)$ $(y > 0)$ を求め、$E[X|Y]$ と $V(X|Y)$ を求めよ。

定理 4.14 条件つき期待値について以下が成立する。

(1) $E[E[Y|X]] = E[Y]$.

(2) $E[aY + bZ|X] = aE[Y|X] + bE[Z|X]$. ($a, b$ は定数。)

(3) $E[g(X)h(Y)|X] = g(X)E[h(Y)|X]$, 特に $E[g(X)|X] = g(X)$.

(4) X, Y が独立なら $E[Y|X] = E[Y]$.

(5) $E[E[Z|X, Y]|X] = E[Z|X]$.

(6) $E[(Y - g(X))^2]$ を最小にするのは $g(X) = E[Y|X]$ である。これは確率変数からなるベクトル空間に $\langle X, Y\rangle = E[XY]$ で内積を定義するとき、Y から X への正射影が $E[Y|X]$ であることを表している。

(7) $V(Y) = E[V(Y|X)] + V(E[Y|X])$.

ただし、 (3) で g, h は “よい” 関数とし、厳密には (2)–(6) については “a.s.” として成立する。

証明: 一部のみ示す。(1) について (X, Y) が絶対連続型であれば (2.10) を用いて

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

と証明できる。(2)–(5) は省略する。(6) について “よい” 関数 h に対して

$$E[(X - E[X|Y])h(Y)] \overset{(1)}{=} E\big[E[(X - E[X|Y])h(Y)|Y]\big] \overset{(3)}{=} E\big[E[X - E[X|Y]|Y]h(Y)\big]$$

$$\overset{(2)}{=} E[(E[X|Y]-E[E[X|Y]|Y])h(Y)] \overset{(3)}{=} E[(E[X|Y]-E[X|Y])h(Y)]=0.$$

ここで 2 行目の二つ目の等号は $E[X|Y]$ は Y の関数だから (3) より $E[E[X|Y]|Y]=E[X|Y]$ となることを用いた。上式で $h(Y)=E[X|Y]-g(Y)$ とみなして

$$\begin{aligned}&E\,[(X-g(Y))^2]=E[(X-E[X|Y]+E[X|Y]-g(Y))^2]\\&=E[(X-E[X|Y])^2]+2E[(X-E[X|Y])(E[X|Y]-g(Y))]+E[(E[X|Y]-g(Y))^2]\\&=E[(X-E[X|Y])^2]+E[(E[X|Y]-g(Y))^2].\end{aligned}$$

$E[(Y-g(X))^2]$ を最小にするのは $g(X)=E[Y|X]$ である。(7) について:

$$\begin{aligned}E[V(Y|X)]+V(E[Y|X])&=E[E[Y^2|X]-(E[Y|X])^2]+E[(E[Y|X])^2]-(E[E[Y|X]])^2\\&=E[E[Y^2|X]]-(E[E[Y|X]])^2\overset{(1)}{=}E[Y^2]-(E[Y])^2=V(Y).\end{aligned}$$

(7) は次のように示される。

$$\begin{aligned}V(Y)&=E[Y^2]-(E[Y])^2\overset{(1)}{=}E[E[Y^2|X]]-(E[E[Y|X]])^2\\&=E[E[Y^2|X]-(E[Y|X])^2]+E[(E[Y|X])^2]-(E[E[Y|X]])^2=E[V(Y|X)]+V(E[Y|X]).\end{aligned}$$ □

例題 4.3 $X_i,\,i\in\mathbf{N},\,N$ は独立な確率変数で、$X_i,\,i\in\mathbf{N}$, は同じ分布に従い、N のとりうる値は $0,1,2,\ldots$ とする。このとき、

$$S=\begin{cases}0, & N=0\\ \sum_{i=1}^N X_i, & N\ge 1\end{cases} \tag{4.12}$$

とすると、$E[X_1^2]<\infty$ かつ $E[N^2]<\infty$ であれば、次が成立する。

(1) $E[S]=E[X_1]E[N]$, (2) $V(S)=V(X_1)E[N]+(E[X_1])^2V(N)$.

(3) また、$E[e^{tX_1}]<\infty$ かつ $E[e^{tN}]<\infty,\,\forall t\in\mathbf{R}$ であれば、$M_S(t)=M_N(\log M_{X_1}(t))$ となる。

証明: (1) $n\ge 1$ とする。$X_1,\cdots,X_n$ は N の独立で $X_1,\ldots,X_n$ は同分布なので

$$E[S|N=n]=E[X_1+\cdots+X_n|N=n]=E[X_1+\cdots+X_n]=nE[X_1].$$

これは $n=0$ でも $0=0$ として成り立つ。よって、$E[S|N]=NE[X_1]$ なので、

$$E[S]=E[E[S|N]=E[NE[X_1]]=E[X_1]E[N].$$

(2) (1) と同様に $X_1,\cdots,X_n$ は N の独立で $X_1,\ldots,X_n$ は i.i.d. なので

$$V(S|N=n)=V(X_1+\cdots+X_n|N=n)=V(X_1+\cdots+X_n)=nV(X_1),\qquad ,n\ge 1.$$

これは $n=0$ でも成り立つので $V(S|N)=NV(X_1)$. よって、定理 4.14 (7) より

$$V(S)=E[V(S|N)]+V(E[S|N])=E[NV(X_1)]+V(NE[X_1])=V(X_1)E[N]+(E[X_1])^2V(N).$$

(3) (2) と同様に $X_1,\cdots,X_n$ は N の独立で $X_1,\ldots,X_n$ は i.i.d. なので、$n\ge 1$ のとき

$$E[e^{tS}|N=n]=E[e^{t(X_1+\cdots+X_n)}|N=n]=E[e^{t(X_1+\cdots+X_n)}]=E[e^{tX_1}]\times\cdots\times E[e^{tX_n}]=M_{X_1}(t)^n.$$

これは $n=0$ でも成り立つので $E[e^{tS}|N]=M_{X_1}(t)^N$. よって、

$$M_S(t)=E[E[e^{tS}|N]]=E[M_{X_1}(t)^N]=E[e^{N\log M_{X_1}(t)}]=M_N(\log M_{X_1}(t)).$$ □

問 4.23 $X_i,\,i\in\mathbf{N},\,N$ は独立な確率変数で、各 X_i は $\mathrm{Ex}(\lambda)$ に N は $\mathrm{Ge}(p)$ に従うとし、S を (4.12) で定義する。このとき S の積率母関数 $M_S(t)=E[e^{tS}]$ と分布関数 $F_S(y)=P(S\le y)$ を求めよ。(*cf*. 問 2.12(3).)

A　Appendix

参考図書: [OY] 岡安, 吉野, 高橋, 武元 微分積分学入門 裳華房
[SS] 吹田, 新保 著 理工系の微分積分学 学術図書

A.1　関数の級数展開について

離散確率分布に用いる、関数のべき級数展開について述べる ($cf.$ [OY] pp.64—, pp.143—)。次の Rolle の定理の証明は省略する ($cf.$ [OY] p.50)。

定理 A.1 (Rolle の定理)　関数 $f(x)$ が閉区間 $[a,b]$ で連続、(a,b) で微分可能とする。このとき、$f(a)=f(b)$ ならば、$f'(c)=0$ を満たす c が存在する。

定理 A.2 (Taylor の定理)　関数 $f(x)$ が閉区間 $[a,b]$ で n 回微分可能とするとき、R_n を

$$f(b)=f(a)+\frac{f'(a)}{1!}(b-a)+\frac{f''(a)}{2!}(b-a)^2+\cdots+\frac{f^{(n-1)}(a)}{(n-1)!}(b-a)^{n-1}+R_n \tag{A.1}$$

で定める。このとき、R_n は次のように表せる。

(a) $R_n=\dfrac{f^{(n)}(c)}{n!}(b-a)^n$ $(a<c<b)$ を満たす c が存在する。この R_n を **Lagrange の剰余項**という。

(b) $R_n=\dfrac{f^{(n)}(c)}{(n-1)!}(b-c)^{n-1}(b-a)$ $(a<c<b)$ を満たす c が存在する。この R_n を **Cauchy の剰余項**という。

証明: p を自然数とし

$$\varphi(x)=f(x)+\sum_{k=1}^{n-1}f^{(k)}(x)\frac{(b-x)^k}{k!}+R_n\cdot\left(\frac{b-x}{b-a}\right)^p$$

とおくと、$\varphi(a)=\varphi(b)=f(b)$ より Rolle の定理から、$\varphi'(c)=0$ となる $a<c<b$ が存在する。ここで、

$$\begin{aligned}\varphi'(x)&=f'(x)+\sum_{k=1}^{n-1}\left\{f^{(k+1)}(x)\frac{(b-x)^k}{k!}-f^{(k)}(x)\frac{k(b-x)^{k-1}}{k!}\right\}-R_n\cdot\frac{p(b-x)^{p-1}}{(b-a)^p}\\&=f^{(n)}(x)\frac{(b-x)^{n-1}}{(n-1)!}-R_n\cdot\frac{p(b-x)^{p-1}}{(b-a)^p}.\end{aligned}$$

ここで $\varphi'(c)=0$ を $p=n$ として書き直せば (a) の主張を、$p=1$ として書き直せば (b) の主張を得る。　□

ここで、$a>b$ の場合も上記が成立することに注意し、さらに、b を x と置き換え、c は a と x の間にあるので、$x=a+\theta(x-a)$ $(0<\theta<1)$ と表せることに注意して次の系を得る。

系 A.1 関数 $f(x)$ が $x=a$ を含む区間で n 回微分可能とするとき、その区間に属する x に対して

$$f(x)=f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots+\frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1}+R_n$$

とおくと、

Lagrange の剰余項は　$$R_n=\frac{f^{(n)}(a+\theta(x-a))}{n!}(x-a)^n\qquad(0<\theta<1) \tag{A.2}$$

Cauchy の剰余項は　$$R_n=\frac{f^{(n)}(a+\theta(x-a))}{(n-1)!}(1-\theta)^{n-1}(x-a)^n\qquad(0<\theta<1) \tag{A.3}$$

と表せる。

この系でもし $R_n \to 0$ $(n \to \infty)$ であれば

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n-1)}(a)}{(n-1)!}(x-a)^{n-1} + \cdots = \sum_{n=0}^{\infty} f^{(n)}(x)\frac{(x-a)^n}{n!}$$

と表せる。これを $f(x)$ の $x=a$ における Taylor 展開といい、特に $a=0$ のときを Maclaurin 展開という。

例 A.1 主な関数の Maclaurin 展開をあげる。

(1) $e^x = 1 + x + \dfrac{x^2}{2!} + \cdots + \dfrac{x^n}{n!} + \cdots = \displaystyle\sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad (-\infty < x < \infty)$

(2) $\cos x = 1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} - \cdots + (-1)^n \dfrac{x^{2n}}{(2n)!} + \cdots \qquad (-\infty < x < \infty)$

(3) $\sin x = x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \cdots + (-1)^n \dfrac{x^{2n+1}}{(2n+1)!} + \cdots \qquad (-\infty < x < \infty)$

(4) $\log(1+x) = x - \dfrac{x^2}{2} + \dfrac{x^3}{3} - \cdots + (-1)^{n-1}\dfrac{x^n}{n} + \cdots \qquad (-1 < x \leq 1)$

(5) $(1+x)^\beta = 1 + \beta x + \dfrac{\beta(\beta-1)}{2!}x^2 + \cdots + \dfrac{\beta(\beta-1)\cdots(\beta-n+1)}{n!}x^n + \cdots = \displaystyle\sum_{n=0}^{\infty}\binom{\beta}{n}x^n \qquad (|x|<1)$

ただし、β は任意の実数で、$\dbinom{\beta}{n} = \dfrac{\beta(\beta-1)\cdots(\beta-n+1)}{n!}$, $n = 0, 1, 2, \ldots,$ と定める。

証明: (1), (5) のみ示す。(2), (3) は (1) と同様に、(4) は $\displaystyle\sum_{k=0}^{n}(-t)^k = \frac{1-(-t)^{n+1}}{1+t}$ で $\displaystyle\int_0^x \frac{(-t)^n}{1+t}\,dt$ が $-1 < x \leq 1$ のとき 0 に収束することに注意して、両辺を 0 から x までの積分すれば証明できる ($cf.$ 問 A.2)。
(1) Lagrange の剰余項 (A.2) より $R_n = \dfrac{e^{\theta x}}{n!}x^n$ で $0 < \theta < 1$ より、$|R_n| \leq e^{|x|}\dfrac{|x|^n}{n!} \to 0$ $(n \to \infty)$ と示せる。
(5) $f(x) = (1+x)^\beta$ とすると $f^{(n)}(x) = \beta(\beta-1)\cdots(\beta-n+1)(1+x)^{\beta-n}$ より $\dfrac{f^{(n)}(0)}{n!} = \dbinom{\beta}{n}$ となるので、$R_n \to 0$ $(n \to \infty)$ を証明すればよい。Cauchy の剰余項 (A.3) より

$$\begin{aligned} |R_n| &= \left| \frac{\beta(\beta-1)\cdots(\beta-n+1)(1+\theta x)^{\beta-n}}{(n-1)!}(1-\theta)^{n-1}x^n \right| \\ &= \left| \frac{\beta(\beta-1)\cdots(\beta-n+1)}{(n-1)!}x^n \right| \left(\frac{1-\theta}{1+\theta x} \right)^{n-1} (1+\theta x)^{\beta-1}. \end{aligned}$$

ここで、$|x| < 1$ であるとき、$0 < \theta < 1$ より $1 - |x| \leq 1 + \theta x \leq 1 + |x|$ であり、よって、

$\beta - 1 \geq 0$ なら $(1+\theta x)^{\beta-1} \leq (1+|x|)^{\beta-1}$, $\quad \beta - 1 < 0$ なら $(1+\theta x)^{\beta-1} \leq (1-|x|)^{\beta-1}$ であり、

$0 < \dfrac{1-\theta}{1+\theta x} < 1$ より、$C_n = \dfrac{\beta(\beta-1)\cdots(\beta-n+1)}{(n-1)!}x^n$ とおいて、$C_n \to 0$ $(n \to \infty)$ を示せばよい。ここで、

$$\lim_{n\to\infty}\left|\frac{C_{n+1}}{C_n}\right| = \lim_{n\to\infty}\left|\frac{\beta-n}{n}x\right| = |x| < 1$$

となるので、次の問 A.1 より $C_n \to 0$ を得る。 □

注意 A.1 例 A.1 (5) を **Newton の一般化された二項定理**という。特に β が自然数のとき n が β より大きい自然数であれば $\dbinom{\beta}{n} = 0$ となり、通常の二項定理と一致する。

問 A.1 数列 $\{a_n\}$ について、極限 $\displaystyle\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = c$ が存在し $c < 1$ であるとする。このとき $\displaystyle\lim_{n\to\infty} a_n = 0$ となることを示せ。

問 A.2 例 A.1 (2), (3), (4) を示せ。

A.2 Stirling の公式

Stirling の公式 ([OY] pp.103–) とその準備として Wallis の公式 ([Y] p.94) について述べる。ここでは、Stirling の公式については参考文献 [SS] pp.121– の証明法を紹介する。[OY] では台形公式とその誤差を用いて証明している。各自その証明も勉強しておくこと。

命題 A.1 $I_n = \displaystyle\int_0^{\pi/2} \sin^n x\,dx = \begin{cases} \dfrac{(n-1)!!}{n!!}\dfrac{\pi}{2}, & n \text{ は偶数}, \\ \dfrac{(n-1)!!}{n!!}, & n \text{ は奇数}. \end{cases}$, ただし、$n!! = \begin{cases} n(n-2)\cdots 4\cdot 2, & n \text{ は偶数}, \\ n(n-2)\cdots 3\cdot 1, & n \text{ は奇数}. \end{cases}$

証明: $I_0 = \dfrac{\pi}{2}$, $I_1 = 1$ と、部分積分により $I_n = \dfrac{n-1}{n} I_{n-2}$, $n \geq 2$, となることより従う。(証明は演習問題とする。) □

定理 A.3 (Wallis の公式) $\sqrt{\pi} = \displaystyle\lim_{n\to\infty} \frac{1}{\sqrt{n}} \frac{(2n)!!}{(2n-1)!!} = \lim_{n\to\infty} \frac{2^{2n}}{\sqrt{n}} \frac{(n!)^2}{(2n)!}.$

証明: 命題 A.1 の I_n について、$\sin^{2n+1} x < \sin^{2n} x < \sin^{2n-1} x$, $(0 < x < \frac{\pi}{2})$ より $I_{2n+1} < I_{2n} < I_{2n-1}$. 即ち、

$$\frac{(2n)!!}{(2n+1)!!} < \frac{(2n-1)!!}{(2n)!!}\frac{\pi}{2} < \frac{(2n-2)!!}{(2n-1)!!}.$$

ここで、(左辺) $= \dfrac{1}{2n+1}\dfrac{(2n)!!}{(2n-1)!!}$, (右辺) $= \dfrac{1}{2n}\dfrac{(2n)!!}{(2n-1)!!}$ なので、$\dfrac{(2n)!!}{(2n-1)!!}$ を移項して、

$$\frac{1}{2n+1}\frac{2}{\pi} < \left\{\frac{(2n-1)!!}{(2n)!!}\right\}^2 < \frac{1}{2n}\frac{2}{\pi}, \quad \text{よって、} \quad \sqrt{\pi} < \frac{1}{\sqrt{n}}\frac{(2n)!!}{(2n-1)!!} < \sqrt{\frac{2n+1}{n}\frac{\pi}{2}}.$$

これより、最初の等式を得る。第 2 の等式は $(2n-1)!! = \dfrac{(2n)!}{(2n)!!}$ と $(2n)!! = 2^n \cdot n!$ より従う。 □

定理 A.4 (Stirling の公式) $\displaystyle\lim_{n\to\infty} \frac{n!}{n^{n+\frac{1}{2}}e^{-n}} = \sqrt{2\pi}.$

証明: $a_n = \dfrac{n!}{n^{n+\frac{1}{2}}e^{-n}}$ とおくと、

$$\log\frac{a_n}{a_{n+1}} = \log\frac{n!}{n^{n+\frac{1}{2}}e^{-n}}\frac{(n+1)^{n+1+\frac{1}{2}}e^{-n-1}}{(n+1)!} = \left(n+\frac{1}{2}\right)\log\left(1+\frac{1}{n}\right) - 1.$$

ここで、$y = 1/x$ の $[n, n+1]$ の部分の面積を上下の台形 (下の台形は $x = n + 1/2$ における接線を用いる) の面積比較により、

$$\frac{1}{2}\cdot 2\cdot\frac{1}{n+1/2} < \int_n^{n+1}\frac{1}{x}\,dx\left(= \log\left(1+\frac{1}{n}\right)\right) < \frac{1}{2}\left(\frac{1}{n}+\frac{1}{n+1}\right).$$

よって、$0 < \log\dfrac{a_n}{a_{n+1}} < \left(n+\dfrac{1}{2}\right)\dfrac{1}{2}\left(\dfrac{1}{n}+\dfrac{1}{n+1}\right) - 1 = \dfrac{1}{4n(n+1)} = \dfrac{1}{4}\left(\dfrac{1}{n}-\dfrac{1}{n+1}\right)$ となるので、

$$0 < \log\frac{a_n}{a_{2n}} = \sum_{k=n}^{2n-1}\log\frac{a_k}{a_{k+1}} < \sum_{k=n}^{2n-1}\frac{1}{4}\left(\frac{1}{k}-\frac{1}{k+1}\right) < \frac{1}{4n}.$$

これより、$\displaystyle\lim_{n\to\infty}\frac{a_n}{a_{2n}} = 1$ を得る。ここで、Wallis の公式を用いるため、

$$\frac{2^{2n}}{\sqrt{n}}\frac{(n!)^2}{(2n)!} = \frac{2^{2n}(a_n n^{n+\frac{1}{2}}e^{-n})^2}{\sqrt{n}\cdot a_{2n}(2n)^{2n+\frac{1}{2}}e^{-n}} = \frac{a_n^2}{\sqrt{2}a_{2n}}$$

と変形すれば、$\displaystyle\lim_{n\to\infty} a_n = \lim_{n\to\infty}\sqrt{2}\frac{a_{2n}}{a_n}\frac{2^{2n}}{\sqrt{n}}\frac{(n!)^2}{(2n)!} = \sqrt{2}\cdot 1\cdot\sqrt{\pi} = \sqrt{2\pi}$ となり主張を得る。 □

A.3　べき級数

べき級数 $\sum_{n=0}^{\infty} a_n(x-c)^n$ についての収束性とその項別微分等について復習する ($cf.$ [OY] pp.136—, pp.224—)。簡単のため、$c=0$ の場合

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots$$

について述べる。(べき級数のことを整級数ともいう。)

定理 A.5 (Abel の補題)　整級数 $\sum_{n=0}^{\infty} a_n x^n$ が $x = x_0 (\neq 0)$ で収束すれば、$|x| < |x_0|$ なるすべての x について絶対収束する。

証明: 仮定より $a_n x_0^n \to 0$ $(n \to \infty)$ より $\{a_n x_0^n\}$ は有界な数列。$|a_n x_0^n| \leq M$, $n = 0, 1, \ldots,$ とすると、$|x| < |x_0|$ ならば

$$\sum_{n=0}^{\infty} |a_n x^n| = \sum_{n=0}^{\infty} |a_n x_0^n| \frac{|x|^n}{|x_0|^n} \leq \sum_{n=0}^{\infty} M\Big(\frac{|x|}{|x_0|}\Big)^n = \frac{M}{1 - |x|/|x_0|} < \infty$$

となり絶対収束することが示された。　□

整級数 $\sum_{n=0}^{\infty} a_n x^n$ の収束、発散について Abel の補題より次がわかる。

(i)　$x = 0$ 以外のすべての x について発散する。

(ii)　すべての実数 x について収束する。

(iii)　ある $0 < R < \infty$ があって、$|x| < R$ なるすべての x に対して絶対収束、$|x| > R$ ならすべての x に対して発散する。

(iii) の R を整級数 $\sum_{n=0}^{\infty} a_n x^n$ の**収束半径**という。(i) の場合には $R = 0$, (ii) の場合には $R = \infty$ と規約する。

例 A.2　$\sum_{n=0}^{\infty} a^n x^n$ のとき $R = 1/a$, ($a \neq 0$ は定数), $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ のとき $R = \infty$, $\sum_{n=0}^{\infty} n^n x^n$ のとき $R = 0$.

定理 A.6　(1) $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = r$ または、(2) $\lim_{n\to\infty} \sqrt[n]{|a_n|} = r$ $(0 \leq r \leq \infty)$ が存在するとき、$R = 1/r$ である。ただし、$1/0 = \infty$, $1/\infty = 0$ とする。

証明: (1) $\lim_{n\to\infty} \frac{|a_{n+1} x^{n+1}|}{|a_n x^n|} = r|x|$, (2) $\lim_{n\to\infty} \sqrt[n]{|a_n x^n|} = r|x|$ であるから、それぞれ次の補題 A.1(1), (2) より、整級数 $\sum_{n=0}^{\infty} a_n x^n$ は $r|x| < 1$ なら収束、$r|x| > 1$ なら発散する。　□

補題 A.1　級数 $\sum_{n=0}^{\infty} a_n$ について、(1) $\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} = c$ または、(2) $\lim_{n\to\infty} \sqrt[n]{|a_n|} = c$ $(0 \leq c \leq \infty)$ が存在するとき、$c < 1$ なら絶対収束し、$c > 1$ なら発散する。

証明: (1) のみ示す。(2) は演習問題とする。$c < 1$ のとき、$0 < \varepsilon < 1 - c$ に対して、ある $N \in \mathbf{N}$ があって、$n \geq N$ なら $\left| \frac{|a_{n+1}|}{|a_n|} - c \right| < \varepsilon$, 即ち、$|a_{n+1}| \leq (c+\varepsilon)|a_n|$. 特に、$|a_n| \leq (c+\varepsilon)^{n-N}|a_N|$, $n \geq N$ を得る。よって、$0 < c + \varepsilon < 1$ に注意して、

$$\sum_{n=0}^{\infty} |a_n| \leq \sum_{n=0}^{N-1} |a_n| + \sum_{n=N}^{\infty} (c+\varepsilon)^{n-N}|a_N| = \sum_{n=0}^{N-1} |a_n| + \frac{|a_N|}{1-(c+\varepsilon)} < \infty.$$

$c>1$ のとき、同様に $0<\varepsilon<c-1$ に対して、ある $N\in\mathbf{N}$ があって、$n\geq N$ なら $\left|\frac{|a_{n+1}|}{|a_n|}-c\right|<\varepsilon$, 即ち、$|a_{n+1}|\geq(c-\varepsilon)|a_n|$. 特に、$c-\varepsilon>1$ に注意して、$|a_n|\geq(c-\varepsilon)^{n-N}|a_N|\to\infty$, $n\to\infty$ となるので、級数 $\sum_{n=0}^{\infty}a_n$ は発散する。 □

問 A.3 上記の補題 A.1(2) を示せ。また、例 A.2 を確かめよ。さらに、次の整級数の収束半径 R を求めよ。ただし、$p\in\mathbf{R}$, $\alpha>0$ は定数とする。解: (1) $R=1$, (2) $R=1$, (3) $R=1/e$, (4) $R=\alpha^{-1/2}$.

(1) $\sum\limits_{n=0}^{\infty}(n+1)^p x^n$, (2) $\sum\limits_{n=0}^{\infty}\binom{-\alpha}{n}x^n$, (3) $\sum\limits_{n=0}^{\infty}\frac{n^n}{n!}x^n$, (4) $\sum\limits_{n=0}^{\infty}\alpha^n x^{2n}$,

定理 A.7 整級数 $\sum\limits_{n=0}^{\infty}a_nx^n$ の収束半径を R $(0<R\leq\infty)$ とし、$f(x)=\sum\limits_{n=0}^{\infty}a_nx^n$ とする。

(1) この級数は $|x|<R$ で絶対収束し、さらに、$(-R,R)$ に含まれる任意の閉区間で一様収束する。特に、$f(x)=\sum\limits_{n=0}^{\infty}a_nx^n$ は $(-R,R)$ で連続である。

(2) $(-R,R)$ に含まれる任意の閉区間 $[a,b]$ で項別積分できる。すなわち、

$$\int_a^b f(x)\,dx=\int_a^b\sum_{n=0}^{\infty}a_nx^n\,dx=\sum_{n=0}^{\infty}\int_a^b a_nx^n\,dx=\sum_{n=0}^{\infty}\left[\frac{a_n}{n+1}x^{n+1}\right]_a^b.$$

(3) $f(x)$ は $(-R,R)$ で微分可能で、項別微分できる。

$$f'(x)=\left(\sum_{n=0}^{\infty}a_nx^n\right)'=\sum_{n=1}^{\infty}na_nx^{n-1}.$$

項別微分してできる整級数の収束半径も R である。

証明には有限和 $f_n(x)=\sum\limits_{k=0}^{n}a_kx^k$ に対して、次の関数列の収束に関する定理を用いる (cf. [OY] pp.224—)。証明は省略する。

定理 A.8 (1) 関数列の一様収束極限は連続である。

(2) 閉区間 $[a,b]$ 上連続な関数列 $\{f_n(x)\}$ が $[a,b]$ 上 $f(x)$ の一様収束するとき、$\lim\limits_{n\to\infty}\int_a^b f_n(x)\,dx=\int_a^b f(x)\,dx$.

(3) 区間 I 上の C^1-関数の列 $\{f_n(x)\}$ がある 1 点 $x_0\in I$ で収束し、さらに $\{f_n'(x)\}$ が I 上一様収束するとき、$\{f_n(x)\}$ は I 上のすべての点で収束し、極限関数 $f(x)$ も I 上で C^1 級であって、$\lim\limits_{n\to\infty}f_n'(x)=f'(x)$ となる。

また、関数項級数に対する次の一様収束に関する定理を用いる (cf. [OY] p.227)。証明は演習問題とする。

定理 A.9 (Weierstrass の判定法) 区間 I 上の関数項級数 $\sum\limits_{n=1}^{\infty}f_n(x)$ に関して、数列 $\{M_n\}$ で

$$|f_n(x)|\leq M_n\quad(n\in\mathbf{N},\ x\in I),\qquad \text{かつ}\quad \sum_{n=1}^{\infty}M_n<\infty$$

を満たすものがあるとき、$\sum\limits_{n=1}^{\infty}f_n(x)$ は I 上一様収束する。

定理 A.7 の証明: (1) 任意の $0<r<R$ をとり、$|x|\leq r$ で一様収束することを示せばよい。$r<R$ より、$\sum a_nx^n$ は $x=r$ で絶対収束する。すなわち、$\sum|a_n|r^n<\infty$. $|x|\leq r$ のとき $|a_nx^n|\leq|a_n|r^n$ であるから、Weierstrass の判定法より、$\sum a_nx^n$ は一様収束する。よって、定理 A.8 (1) より、$f(x)$ の連続性が従う。

(2) (1) と定理 A.8 (2) より成立する。

(3) まず、$\sum\limits_{n=1}^{\infty}na_nx^{n-1}$ の収束半径 R' が R に等しいことを示す。$\sum\limits_{n=1}^{\infty}na_nx^n$ の収束半径も明らかに R' である。$|a_nx^n|\leq|na_nx^n|$ であるから、$R'\leq R$. 次に、$|x|<R$ なる任意の x を固定して、$|x|<\xi<R$ なる ξ

をとると、$|na_nx^n| = n|x/\xi|^n|a_n\xi^n|$. ここで、$|x/\xi| < 1$ より $n|x/\xi|^n \to 0$ $(n \to \infty)$, 従って、$\{n|x/\xi|^n\}$ は有界なので、$\sum |a_n\xi^n|$ の収束性から、$\sum |a_nx^n|$ の収束性が従う。よって、$|x| \leq R'$. これが $|x| < R$ なる任意の x について成立するので $R \leq R'$. 以上より $R = R'$.

よって、(1) より $\sum\limits_{n=1}^{\infty} na_nx^n$ は $(-R, R)$ に含まれる任意の閉区間で一様収束する。これと定理 A.8 (3) より $\sum\limits_{n=1}^{\infty} a_nx^n$ は項別微分してよいことがわかる。 □

定理 A.7 (3) を繰り返し用いることにより、次の系を得る。

系 A.2 整級数 $\sum\limits_{n=0}^{\infty} a_nx^n$ の収束半径を R $(0 < R \leq \infty)$ とすると、$(-R, R)$ において $\sum\limits_{n=0}^{\infty} a_nx^n$ は何回でも微分可能で、しかも項別微分できる。

A.4 Cantor 集合と Cantor 関数

注意 2.3 で述べた Cantor 関数と Cantor 集合を紹介する。

参考図書: 伊藤清三 著 ルベーグ積分入門 裳華房 pp.41–45.

$[0, 1]$ の閉部分集合の減少列 K_n, $n = 1, 2, \ldots,$ を次のように定め、$K = \bigcap_{n=1}^{\infty} K_n$ とおく。

$$K_1 = \Big[0, \frac{1}{3}\Big] \cup \Big[\frac{2}{3}, 1\Big] = [0, 1] \backslash \Big(\frac{1}{3}, \frac{2}{3}\Big), \qquad K_n = K_{n-1} \backslash \bigcup_{k=0}^{3^{n-1}-1} \Big(\frac{k}{3^{n-1}} + \frac{1}{3^n}, \frac{k}{3^{n-1}} + \frac{2}{3^n}\Big),\ n = 2, 3, \ldots.$$

ここで、各 K_n はコンパクトなので $K \neq \emptyset$ であることに注意する。この K を **Cantor 集合**という。

定理 A.10 Cantor 集合 K は連続濃度をもつ長さ 0 の集合となる。

証明: 各 K_n は 2^n 個の長さ $1/3^n$ の区間の和集合なので、K_n の長さは $2^n/3^n$ となる。特に、K の長さは 0 となる。

K が連続濃度をもつことを示すためには K から $[0, 1]$ への全射 F を定義すればよい。

各 $n \in \mathbf{N}$ に対して、$F_n : \mathbf{R} \to [0, 1]$ を $F_n(0) = 0$ で K_n 上でのみ傾き $(3/2)^n$ で増大し、それ以外では傾き 0 となる折れ線と定義する。例えば、$F_1(x) = 0$, $x < 0$, $F_1(x) = 1$, $x > 1$ で

$$F_1(x) = \frac{3}{2}x,\ 0 \leq x \leq \frac{1}{3}, \quad F_1(x) = \frac{1}{2},\ \frac{1}{3} < x < \frac{2}{3}, \quad F_1(x) = \frac{3}{2}\Big(x - \frac{2}{3}\Big) + \frac{1}{2},\ \frac{1}{3} \leq x \leq 1,$$

となる。このとき、構成法より $|F_{n+1}(x) - F_n(x)| \leq 1/2^n$ となることに注意する。よって、$m > n$ とすると、

$$|F_m(x) - F_n(x)| \leq \sum_{k=n}^{m-1} |F_{k+1}(x) - F_k(x)| \leq \sum_{k=n}^{m-1} \frac{1}{2^k} = 2\Big(\frac{1}{2^n} - \frac{1}{2^m}\Big) \to 0, \qquad m, n \to \infty$$

となり、$\{F_n(x)\}$ は一様収束の意味で Cauchy 列をなすので、連続関数 $F(x)$ が存在して一様収束する。

構成法より $x \notin K_n$ のとき $m \geq n$ なら $F_m(x) = F_n(x) = k/2^n$ となる $k = 0, 1, \ldots, 2^n$ があるので、$F(x)$ は K_n の補集合の各開区間では定数関数となる。すなわち、$F(x)$ は K 上でのみ増加する連続関数なので、F が K から $[0, 1]$ への全射となっていることが示された。 □

注意 A.2 定理 A.10 の証明で定義した連続関数 $F(x)$ を **Cantor 関数**という。$F(x)$ は $F(x) \equiv 0$, $x < 0$, $F(x) \equiv 1$, $x > 1$ と拡張すると、$F(x)$ はある確率変数の分布関数となるが、密度関数は持たない。K の外で定数なので、$F'(x) = 0$ となる。すなわち K の Lebesgue 測度が 0 であるから、$F' = 0$ a.e. となるので、$\displaystyle\int_0^1 F'(t)\,dt = 0$, 一方 $F(1) - F(0) = 1$ となるためである。

A.5　カイ 2 乗分布、t 分布表、標準正規分布の上側 α 点について

自由度 n の χ^2 分布の上側 α 点: $\chi_n^2(\alpha)$

n \ α	0.975	0.950	0.050	0.025
1	0.0010	0.0039	3.8415	5.0239
2	0.0506	0.1026	5.9915	7.3778
3	0.2158	0.3518	7.8147	9.3484
4	0.4844	0.7107	9.4877	11.1433
5	0.8312	1.1455	11.0705	12.8325
6	1.2373	1.6354	12.5916	14.4494
7	1.6899	2.1673	14.0671	16.0128
8	2.1797	2.7326	15.5073	17.5345
9	2.7004	3.3251	16.9190	19.0228
10	3.2470	3.9403	18.3070	20.4832
11	3.8157	4.5748	19.6751	21.9200
12	4.4038	5.2260	21.0261	23.3367
13	5.0088	5.8919	22.3620	24.7356
14	5.6287	6.5706	23.6848	26.1189
15	6.2621	7.2609	24.9958	27.4884
16	6.9077	7.9616	26.2962	28.8454
17	7.5642	8.6718	27.5871	30.1910
18	8.2307	9.3905	28.8693	31.5264
19	8.9065	10.1170	30.1435	32.8523
20	9.5908	10.8508	31.4104	34.1696
21	10.2829	11.5913	32.6706	35.4789
22	10.9823	12.3380	33.9244	36.7807
23	11.6886	13.0905	35.1725	38.0756
24	12.4012	13.8484	36.4150	39.3641
25	13.1197	14.6114	37.6525	40.6465
26	13.8439	15.3792	38.8851	41.9232
27	14.5734	16.1514	40.1133	43.1945
28	15.3079	16.9279	41.3371	44.4608
29	16.0471	17.7084	42.5570	45.7223
30	16.7908	18.4927	43.7730	46.9792
31	17.5387	19.2806	44.9853	48.2319
32	18.2908	20.0719	46.1943	49.4804
33	19.0467	20.8665	47.3999	50.7251
34	19.8063	21.6643	48.6024	51.9660
35	20.5694	22.4650	49.8018	53.2033
36	21.3359	23.2686	50.9985	54.4373
37	22.1056	24.0749	52.1923	55.6680
38	22.8785	24.8839	53.3835	56.8955
39	23.6543	25.6954	54.5722	58.1201
40	24.4330	26.5093	55.7585	59.3417

自由度 n の t 分布の上側 α 点: $t_n(\alpha)$

n \ α	0.100	0.050	0.025
1	3.0777	6.3138	12.7062
2	1.8856	2.9200	4.3027
3	1.6377	2.3534	3.1824
4	1.5332	2.1318	2.7764
5	1.4759	2.0150	2.5706
6	1.4398	1.9432	2.4469
7	1.4149	1.8946	2.3646
8	1.3968	1.8595	2.3060
9	1.3830	1.8331	2.2622
10	1.3722	1.8125	2.2281
11	1.3634	1.7959	2.2010
12	1.3562	1.7823	2.1788
13	1.3502	1.7709	2.1604
14	1.3450	1.7613	2.1448
15	1.3406	1.7531	2.1314
16	1.3368	1.7459	2.1199
17	1.3334	1.7396	2.1098
18	1.3304	1.7341	2.1009
19	1.3277	1.7291	2.0930
20	1.3253	1.7247	2.0860
21	1.3232	1.7207	2.0796
22	1.3212	1.7171	2.0739
23	1.3195	1.7139	2.0687
24	1.3178	1.7109	2.0639
25	1.3163	1.7081	2.0595
26	1.3150	1.7056	2.0555
27	1.3137	1.7033	2.0518
28	1.3125	1.7011	2.0484
29	1.3114	1.6991	2.0452
30	1.3104	1.6973	2.0423

標準正規分布の上側 α 点 $u(\alpha)$ について

$$u(0.05) = 1.645$$
$$u(0.025) = 1.960$$
$$u(0.01) = 2.326$$
$$u(0.005) = 2.576$$

注意. Excel で cell に「=CHIINV(α,n)」,「=T.INV($1-\alpha$, n)」として作成した。