第1章　確率・確率変数の基本事項

この章では, 統計的推測理論を理解する上で必要な確率論の基礎事項を簡潔に述べる．したがって, 証明はしない場合がある．測度論の内容には立ち入らない．数学的な厳密性を犠牲にして, 直観的な理解を目指す記述を心掛けた．節 1.1 では, 確率に関わる基本的な概念を導入する．節 1.2 では, 抽象的な確率空間と数量の世界を結ぶ確率変数を導入し, 確率変数の基本的な性質を説明する．節 1.3 では, 分布関数の逆関数のようなものである分位点を導入する．節 1.4 では, 1 変数確率変数の確率モデルを説明する．節 1.5 では, 2 次元確率変数の同時分布, 周辺分布, 条件付き分布を説明する．節 1.6 では, 多変数確率変数の確率モデルを導入する．節 1.7 では, 正規分布から誘導される重要な分布であるカンマ分布, χ^2 分布, t 分布, F 分布を導入する．

1.1　基礎的な確率規則

確率論はランダムな現象を扱う数学理論である．確率論で扱う行為を**試行**という．試行のありえる結果すべてを集めた集合を**標本空間**といい, Ω と記すことにする．Ω の部分集合[1]を**事象**という．事象には標本空間 Ω と空事象 $\varnothing$(何も起こらないという事象) も含める．事象をすべて集めた集合族を $\mathcal{A}$ と記す．$\mathcal{A}$ は以下で述べる σ **加法性** (完全加法性) をみたすことにする．

定義 1.1. Ω を空でない集合とし, $\mathcal{A}$ を Ω の部分集合族とする．$\mathcal{A}$ が次の 3 条件をみたすとき, σ **加法族**と呼ばれる．

(1) $\Omega \in \mathcal{A}$.

(2) $A \in \mathcal{A} \Longrightarrow A^c \in \mathcal{A}$.

(3) $A_n \in \mathcal{A}\,(n = 1, 2, \ldots) \Longrightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

[1] Ω が可算集合ならば, 事象はすべての部分集合としてよいが, Ω が連続濃度のときには, すべての部分集合を対象にすることはしない．

ただし $A^c = \{\omega \in \Omega;\ \omega \notin A\}$ である．Ω と $\mathcal{A}$ の組 $(\Omega, \mathcal{A})$ を**可測空間**と呼ぶ

補題 1.2. $(\Omega, \mathcal{A})$ を可測空間とする．このとき以下が成立する．

(1) $\varnothing \in \mathcal{A}$.

(2) $A_n \in \mathcal{A}\,(n = 1, 2, \ldots) \Longrightarrow \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$.

Proof. (1) $\varnothing = \Omega^c \in \mathcal{A}$.
(2) $A_n^c \in \mathcal{A}$ と De Morgan の法則から

$$\bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} (A_n^c)^c = \left(\bigcup_{n=1}^{\infty} A_n^c\right)^c \in \mathcal{A}.$$

最後は定義 1.1(3) を用いた．□

注意 1.3. $\mathcal{C}$ を Ω の部分集合族とする．$\mathcal{C}$ は σ 加法性をみたしてなくともよい．このとき, 集合族 $\sigma[\mathcal{C}]$ を

$$\sigma[\mathcal{C}] := \bigcap \big\{\mathcal{A};\ \mathcal{A} \supset \mathcal{C},\ \mathcal{A} \text{ は } \sigma \text{ 加法族}\big\}$$

で定める．すると $\sigma[\mathcal{C}]$ は σ 加法族となる．さらに $\mathcal{G}$ を $\mathcal{A}$ を含む σ 加法族としたとき

$$\sigma[\mathcal{C}] \subseteq \mathcal{G}$$

となる．すなわち $\sigma[\mathcal{C}]$ は $\mathcal{C}$ を含む最小 (包含関係の意味) の σ 加法族となる．□

定義 1.4. $\Omega = \mathbb{R}$ とし

$$\mathcal{O} = \{O \subset \mathbb{R};\ O \text{ は } \mathbb{R} \text{ の開集合}\}$$

とする．$\sigma[\mathcal{O}]$ を $\mathbb{R}$ の **Borel 集合族**と呼び, $\mathcal{B}(\mathbb{R})$ と記す．また

$$\mathcal{C} = \{(-\infty, x) \subset \mathbb{R};\ x \in \mathbb{R}\}$$

とする．このとき $\mathcal{C}$ は $\mathcal{O}$ の真部分集合であるが $\sigma[\mathcal{C}] = \sigma[\mathcal{O}]$ となる[2]．

注意 1.5. Ω が高々可算集合のときは, $\mathcal{F} = 2^{\Omega}$ と取る．ただし 2^{Ω} は Ω のすべての部分集合からなる集合族で冪(べき)集合という．□

[2] $\sigma[\mathcal{C}] \subset \sigma[\mathcal{O}]$ は明らかであるが, 逆の包含関係も示すことができる．

定義 1.6. $(\Omega, \mathcal{A})$ を可測空間とする．$\mathcal{A}$ 上の関数

$$\mathsf{Pr} : \mathcal{A} \ni A \mapsto \mathsf{Pr}(A) \in [0, 1]$$

が次の 2 条件をみたすとき，$(\Omega, \mathcal{A})$ 上の**確率測度**[3]と呼ばれる．

(1) $\mathsf{Pr}(\Omega) = 1$.

(2) 互いに排反[4]な事象列 $A_n \in \mathcal{A}\,(n = 1, 2, \ldots)$ に対して

$$\mathsf{Pr}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathsf{Pr}(A_n) := \lim_{N\to\infty} \sum_{n=1}^{N} \mathsf{Pr}(A_n).$$

これらの 3 つの組 $(\Omega, \mathcal{A}, \mathsf{Pr})$ を**確率空間**という．

補題 1.7. $(\Omega, \mathcal{A}, \mathsf{Pr})$ を確率空間とする．このとき以下が成立する．

(1) $\mathsf{Pr}(\varnothing) = 0$.

(2) $A \in \mathcal{A}$ に対して，$\mathsf{Pr}(A^c) = 1 - \mathsf{Pr}(A)$.

(3) $N \in \mathbb{N}$ とする．$\{A_n\}_{n=1}^{N} \subset \mathcal{A}$ が互いに排反ならば

$$\mathsf{Pr}\left(\bigcup_{n=1}^{N} A_n\right) = \sum_{n=1}^{N} \mathsf{Pr}(A_n).$$

(4) $A, B \in \mathcal{A}, A \subset B \Longrightarrow \mathsf{Pr}(B \setminus A) = \mathsf{Pr}(B) - \mathsf{Pr}(A)$.

(5) (Boole の定理/ユニオン・バウンド) $A_n \in \mathcal{A}\,(n = 1, 2, \ldots)$ に対して

$$\mathsf{Pr}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mathsf{Pr}(A_n).$$

(6) $A_n \in \mathcal{A}$ が $A_n \subset A_{n+1}\,(n = 1, 2, \ldots)$ をみたすならば

$$\lim_{n\to\infty} \mathsf{Pr}(A_n) = \mathsf{Pr}\left(\bigcup_{n=1}^{\infty} A_n\right).$$

(7) $A_n \in \mathcal{A}$ が $A_n \supset A_{n+1}\,(n = 1, 2, \ldots)$ をみたすならば

$$\lim_{n\to\infty} \mathsf{Pr}(A_n) = \mathsf{Pr}\left(\bigcap_{n=1}^{\infty} A_n\right).$$

---

[3]簡単に Ω 上の確率測度ともいう．

[4]$m \neq n$ ならば，$A_m \cap A_n = \varnothing$ が成立していること．

Proof. (1) $F_1 := \Omega$, $F_n := \varnothing\,(n \geq 2)$ とおくと $\{F_n\}_{n=1}^{\infty}$ は互いに排反な事象列となる．定義 1.6(2) を用いると

$$\mathsf{Pr}(\Omega) = \mathsf{Pr}\left(\bigcup_{n=1}^{\infty} F_n\right) = \sum_{n=1}^{\infty} \mathsf{Pr}(F_n) = \mathsf{Pr}(\Omega) + \sum_{n=2}^{\infty} \mathsf{Pr}(\varnothing)$$

を得る．よって

$$\mathsf{Pr}(\varnothing) = 0$$

がわかる．
(2) $F_1 := A$, $F_2 := A^c$, $F_n = \varnothing\,(n \geq 3)$ とおく．すると $\bigcup_{n=1}^{\infty} F_n = \Omega$ かつ $\{F_n\}_{n=1}^{\infty}$ は互いに排反な事象列となるので, 定義 1.6(1), (2) を用いると

$$\begin{aligned}
1 = \mathsf{Pr}(\Omega) &= \mathsf{Pr}\left(\bigcup_{n=1}^{\infty} F_n\right) = \sum_{n=1}^{\infty} \mathsf{Pr}(F_n) = \mathsf{Pr}(A) + \mathsf{Pr}(A^c) + \sum_{n=3}^{\infty} \mathsf{Pr}(\varnothing) \\
&= \mathsf{Pr}(A) + \mathsf{Pr}(A^c) \qquad (\because (1))
\end{aligned}$$

から, (2) は示せた．
(3) $F_i := A_i\,(i = 1, 2, \ldots, N)$ と $F_i = \varnothing\,(i \geq n+1)$ とおくと $\{F_n\}_{n=1}^{\infty}$ は互いに排反な事象列で $\bigcup_{n=1}^{\infty} F_n = \bigcup_{i=1}^{N} A_i$ となる．定義 1.6(2) を用いると

$$\begin{aligned}
\mathsf{Pr}\left(\bigcup_{i=1}^{N} A_i\right) &= \mathsf{Pr}\left(\bigcup_{n=1}^{\infty} F_n\right) = \sum_{i=1}^{\infty} \mathsf{Pr}(F_i) = \sum_{i=1}^{N} \mathsf{Pr}(A_i) + \sum_{i=N+1}^{\infty} \mathsf{Pr}(\varnothing) \\
&= \sum_{i=1}^{N} \mathsf{Pr}(A_i)
\end{aligned}$$

を得る．よって, (3) は示された．
(4) $B \setminus A = B \cap A^c$ かつ $B = (B \cap A^c) \cup A$ である．このことに注意して, $F_1 := B \cap A^c$, $F_2 = A$, $A_n = \varnothing\,(n \geq 3)$ とおくと $\{A_n\}_{n=1}^{\infty}$ は互いに排反な事象列で $\bigcup_{n=1}^{\infty} A_n = B$ となる．定義 1.6(2) を用いると

$$\begin{aligned}
\mathsf{Pr}(B) &= \mathsf{Pr}\left(\bigcup_{n=1}^{\infty} F_n\right) = \sum_{n=1}^{\infty} \mathsf{Pr}(F_n) = \mathsf{Pr}(B \cap A^c) + \mathsf{Pr}(A) + \sum_{n=3}^{\infty} \mathsf{Pr}(\varnothing) \\
&= \mathsf{Pr}(B \cap A^c) + \mathsf{Pr}(A)
\end{aligned}$$

より, (4) は示せた．
(5) $\mathsf{Pr}(A_1 \cup A_2) \leq \mathsf{Pr}(A_1) + \mathsf{Pr}(A_2)$ を示せばよい．まず, $A_1 \cup A_2 = A_1 \cup \Bigl(A_2 \setminus (A_1 \cap A_2\Bigr)$ かつ $A_1 \bigcap \Bigl(A_2 \setminus (A_1 \cap A_2)\Bigr) = \varnothing$ であることに注

意する. (3) から

$$\begin{aligned}
\mathsf{Pr}\big(A_1 \cup A_2\big) &= \mathsf{Pr}(A_1) + \mathsf{Pr}\Big(A_2 \setminus (A_1 \cap A_2)\Big) \\
&= \mathsf{Pr}(A_1) + \mathsf{Pr}(A_2) - \mathsf{Pr}(A_1 \cap A_2) \\
&\qquad \big(\because A_2 \supset A_1 \cap A_2 \text{ なので (4) を用いた}\big) \\
&\leq \mathsf{Pr}(A_1) + \mathsf{Pr}(A_2) \qquad \big(\because \mathsf{Pr}(A_1 \cap A_2) \geq 0\big)
\end{aligned}$$

(6) $F_1 := A_1$, $F_2 = A_2 \setminus A_1$, $\ldots$, $F_{n+1} := A_{n+1} \setminus A_n$ とおく. $\{F_n\}_{n=1}^{\infty}$ は互いに排反であり

$$A_n = \bigcup_{i=1}^{n} F_i \implies A := \bigcup_{i=1}^{\infty} F_i = \bigcup_{n=1}^{\infty} A_n$$

となる. よって

$$\begin{aligned}
\mathsf{Pr}(A) &= \mathsf{Pr}\left(\bigcup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} \mathsf{Pr}(F_i) \quad (\because \text{定義 } 1.6(2)) \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \mathsf{Pr}(F_i) = \lim_{n\to\infty} \mathsf{Pr}\left(\bigcup_{i=1}^{n} F_i\right) \quad (\because (3)) \\
&= \lim_{n\to\infty} \mathsf{Pr}(A_n)
\end{aligned}$$

を得る. よって, (6) は示せた.
(7) $A = \bigcap_{n=1}^{\infty} A_n$ とおく. すると $\{A_1 \setminus A_n\}_{n=1}^{\infty}$ は $A_1 \setminus A_2 \subset A_1 \setminus A_3 \subset A_1 \setminus A_n \subset \cdots$ となり,

$$\bigcup_{n=1}^{\infty} (A_1 \setminus A_n) = \bigcup_{n=1}^{\infty} (A_1 \cap A_n^c) = A_1 \cap \left(\bigcap_{n=1}^{\infty} A_n\right)^c = A_1 \cap A^c = A_1 \setminus A$$

となる. (4) と (6) を用いると

$$\begin{aligned}
\mathsf{Pr}(A_1) - \mathsf{Pr}(A) &= \mathsf{Pr}(A_1 \setminus A) \quad (A_1 \supset A \text{ なので, (4) を用いた}) \\
&= \lim_{n\to\infty} \mathsf{Pr}(A_1 \setminus A_n) \quad (\because (6) \text{ を用いた}) \\
&= \lim_{n\to\infty} \Big(\mathsf{Pr}(A_1) - \mathsf{Pr}(A_n)\Big) \\
&= \mathsf{Pr}(A_1) - \lim_{n\to\infty} \mathsf{Pr}(A_n)
\end{aligned}$$

からわかる. □

注意 1.8. 補題 1.7(6)(7) において, $\{\mathsf{Pr}(A_n)\}_{n=1}^{\infty}$ は有界な非減少列もしくは非増加列なので, 左辺は必ず収束することに注意せよ.

定義 1.9. $(\Omega, \mathcal{A}, \mathsf{Pr})$ を確率空間とし, $A, B \in \mathcal{A}$ とする.

(1) (独立性):

$$A \text{ と } B \text{ は}\textbf{独立} \iff \mathsf{Pr}(A \cap B) = \mathsf{Pr}(A)\mathsf{Pr}(B).$$

(2) (条件付き確率)] $\mathsf{Pr}(B) > 0$ のとき, B を与えたときの A の**条件付き確率**を

$$\mathsf{Pr}(A|\,B) := \frac{\mathsf{Pr}(A \cap B)}{\mathsf{Pr}(B)}$$

で定める.

注意 1.10. (1) $A, B \in \mathcal{A}$ とし, $\mathsf{Pr}(B) > 0$ とする. A と B が独立のとき

$$\mathsf{Pr}(A|\,B) = \mathsf{Pr}(A)$$

が成立する.
(2) $\mathsf{Pr}(B) > 0$ のとき

$$\mathsf{Pr}(A \cap B) = \mathsf{Pr}(B)\mathsf{Pr}(A|\,B)$$

となる. これを**乗法の公式**という.
(3) $\mathsf{Pr}(B) > 0$ のとき, 関数 $\mathsf{Pr}(\,\cdot\,|B) : \mathcal{A} \longrightarrow [0, 1]$ は定義 1.6(1) − (3) をみたす. すなわち $(\Omega, \mathcal{A})$ 上の確率測度である. □

問 1.1. 注意 1.10(1)(3) を証明せよ.

補題 1.11. (全確率の法則) $N \in \mathbb{N}$ とする. $A_1, A_2, \ldots, A_N \in \mathcal{A}$ は Ω の分割とする. すなわち, すべての $m \neq n$ に対し, $A_m \cap A_n = \varnothing$ で, $\bigcup_{n=1}^N A_n = \Omega$ が成立している. このとき任意の $B \in \mathcal{A}$ に対して

$$\mathsf{Pr}(B) = \mathsf{Pr}\left(\bigcup_{n=1}^N (A_n \cap B)\right) = \sum_{n=1}^N \mathsf{Pr}(A_n \cap B).$$

Proof. $B = B \cap \Omega = B \cap \left(\bigcup_{n=1}^N A_n\right) = \bigcup_{n=1}^N (A_n \cap B)$ と $\{A_i \cap B\}_{i=1}^n$ は互いに排反であることに注意して, 定義 1.6(2) を用いればよい. □

定理 1.12. (Bayes の定理) (1) $N \in \mathbb{N}$ とする. $A_1, A_2, \ldots, A_N, B \in \mathcal{A}$ とし, $A_1, A_2, \ldots, A_N$ は Ω の**分割**とする. $j = 1, 2, \ldots, N$ に対して, $\mathsf{Pr}(B) > 0$, $\mathsf{Pr}(A_j) > 0$ のとき

$$\mathsf{Pr}(A_j|\,B) = \frac{\mathsf{Pr}(A_j)\mathsf{Pr}(B|\,A_j)}{\sum_{n=1}^N \mathsf{Pr}(A_n)\mathsf{Pr}(B|\,A_n)}$$

が成立する．
(2) $A, B \in \mathcal{A}$ とし，$\mathsf{Pr}(A) > 0$, $\mathsf{Pr}(B) > 0$ とする．このとき

$$\mathsf{Pr}(A|\, B)\mathsf{Pr}(B) = \mathsf{Pr}(B|\, A)\mathsf{Pr}(A)$$

が成立する．

Proof. 条件付き確率の定義，補題 1.11，定義 1.4(2) を用いればよい． □

例 1.13. 病気 D に対する検査の結果を $+$ と $-$ とし，確率が以下であるとする．

$$\mathsf{Pr}(+|\, D) = 0.9, \quad \mathsf{Pr}(-|\, D^c) = 0.9, \quad \mathsf{Pr}(D) = 0.01.$$

Bayes の定理を用いて，検査で $+$ に人が本当に D である確率を求めると

$$\begin{aligned}\mathsf{Pr}(D|\, +) = \frac{\mathsf{Pr}(+\cap D)}{\mathsf{Pr}(+)} &= \frac{\mathsf{Pr}(D)\mathsf{Pr}(+|\, D)}{\mathsf{Pr}(D)\mathsf{Pr}(+|, D) + \mathsf{Pr}(D^c)\mathsf{Pr}(+|\, D^c)}\\ &= \frac{0.01\times 0.9}{0.01\times 0.9 + (1-0.01)\times\{1-0.9\}} \approx 0.83.\end{aligned}$$

1.2　確率変数

前節では，確率と事象を記述する数学的なモデルを導入した．しかし，現実の現象を扱い統計学の対象は，事象には直接結びつかないかもしれない数量的な情報である．以下で定義する確率変数は，事象と数量の間の橋渡しをする．

定義 1.14. $(\Omega, \mathcal{A})$ と $(\mathbb{X}, \mathcal{B})$ を可測空間とする．関数 $X : \Omega \longrightarrow \mathbb{X}$ は $(\Omega, \mathcal{A})$ から $(\mathbb{X}, \mathcal{B})$ への**可測写像**であるとは

$$X^{-1}(B) := \big\{\omega \in \Omega :\ X(\omega) \in B\big\} \in \mathcal{A} \qquad \big(\forall B \in \mathcal{B}\big)$$

をみたすときをいう．

注意 1.15. (i)．$d \geq 2\,(d \in \mathbb{N})$ とする．$\big(\mathbb{X}, \mathcal{B}\big) = \big(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\big)$ のとき，X は**確率ベクトル**と呼べれる．
(ii)．$d = 1$ のとき，X は**確率変数**と呼べれる． □

定理 1.16. $\big(\mathbb{X}, \mathcal{B}\big)$ を可測空間とし，$X : \Omega \longrightarrow \mathbb{X}$ を写像とし，$\mathcal{C}$ を $\mathbb{X}$ の集合族とする．$\forall C \in \mathcal{C}$ に対して，$\big\{\omega \in \Omega :\ X(\omega) \in C\big\} \in \mathcal{A}$ であり，$\mathcal{C}$ が $\mathcal{B}$ を生成する[5]とき，X は可測となる．

[5] $\mathcal{B}$ は $\mathcal{C}$ を含む最小の σ 集合族．

Proof. $B \in \mathcal{B}$ に対して, $\bigl\{\omega \in B : X(\omega) \in B\bigr\}$ を $\{X \in B\}$ と書くことにする. $\{B\}$, $\{B_n\}_{n=1}^{\infty} \subset \mathcal{B}$ に対して

$$\Bigl\{X \in \bigcup_{n=1}^{\infty} B_n\Bigr\} = \bigcup_{n=1}^{\infty}\{X \in B_N\}$$
$$\bigl\{X \in B^c\bigr\} = \bigl\{X \in B\bigr\}^c$$

となる. したがって, 集合族 $\mathcal{D} := \{B \subset \mathbb{X} : \{X \in B\} \in \mathcal{A}\}$ は σ 集合族となる. よって, $\mathcal{D} \subset \mathcal{C}$ であり, $\mathcal{C}$ は $\mathcal{B}$ を生成するので, $\mathcal{D} \subset \mathcal{B}$ となる.
□

注意 1.17. (i). $\bigl(\mathbb{X}, \mathcal{B}\bigr) = \bigl(\mathbb{R}, \mathcal{B}(\mathbb{R})\bigr)$ に対して, 定理 1.16 における $\mathcal{C}$ の選択として, $\{(-\infty, r] : r \in \mathbb{R}\}$ と $\{(-\infty, q] : q \in \mathbb{Q}\}$ などがある.
(ii). $d \geq 2\,(d \in \mathbb{N})$ とする. $\bigl(\mathbb{X}, \mathcal{B}\bigr) = \bigl(\mathbb{R}^d, \mathcal{B}(\mathbb{R})\bigr)$ に対して, 定理 1.16 における $\mathcal{C}$ の選択として

$$\{(a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_d, b_d) : -\infty < a_i < b_i < \infty\,(i = 1, 2, \ldots, d)\}$$

がある. □

定理 1.18. $\bigl(\Omega, \mathcal{A}\bigr)$, $\bigl(\mathbb{X}, \mathcal{B}\bigr)$, $\bigl(\mathbb{Y}, \mathcal{C}\bigr)$ を可測空間とする. $X : \Omega \longrightarrow \mathbb{X}$ と $f : \mathbb{X} \longrightarrow \mathbb{Y}$ は可測写像のとき, $f(X) : \Omega \longrightarrow \mathbb{Y}$ は可測写像となる.

Proof. $C \in \mathcal{C}$ に対して

$$\bigl\{\omega \in \Omega : f(X(\omega)) \in B\bigr\} = \bigl\{\omega \in \Omega : X(\omega) \in f^{-1}(C)\bigr\} \in \mathcal{A}$$

となる. なぜならば, f の可測性から $f^{-1}(C) \in \mathcal{B}$ となることからわかる.
□

定理 1.19. $n \in \mathbb{N}$ とし, $X_1, X_2, \ldots, X_n$ は確率変数とし, $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ は可測とする. このとき, $f(X_1, X_2, \ldots, X_n)$ は確率変数となる.

Proof. 定理 1.18 から $(X_1, X_2, \ldots, X_n)$ は確率ベクトルであることを示せばよい. $A_1, A_2, \ldots, A_n \in \mathcal{B}(\mathbb{R})$ に対して

$$\bigl\{(X_1, X_2, \ldots, X_n) \in A_1 \times A_2 \times \cdots \times A_n\bigr\} = \bigcap_{i=1}^{n}\{X_i \in A_i\} \in \mathcal{A}$$

である. 集合族 $\{A_1 \times A_2 \times \cdots \times A_n : A_i \in \mathcal{B}(\mathbb{R})\,(i = 1, 2, \ldots, n)\}$ は $\mathcal{B}(\mathbb{R}^n$ を生成するので, 定理 1.16) から $\bigl(\mathbb{X}, \mathcal{B}\bigr) = \bigl(\mathbb{R}, \mathcal{B}(\mathbb{R})\bigr)$ に対して, $(X_1, X_2, \ldots, X_n)$ は確率ベクトルであることがわかる. □

定理 1.20. $X_1, X_2, \ldots, X_n$ は確率変数のとき, $X_1 + X_2 + \cdots + X_n$ も確率変数となる.

Proof. 定理 1.19) から $\mathbb{R}^n$ 上の実数値関数 $f(x_1, x_2, \ldots, x_n) = x_1 + x_2 + \cdots + x_n$ が可測であることを示せばよい. $\forall r \in \mathbb{R}$ に対して

$$\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_1 + x_2 + \cdots + x_n < r\}$$

は $\mathbb{R}^n$ の開集合となるので

$$\{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : x_1 + x_2 + \cdots + x_n < r\} \in \mathcal{B}(\mathbb{R}^n)$$

となる. $\{(-\infty, r) : r \in \mathbb{R}\}$ は$\mathcal{B}(\mathbb{R})$ を生成するので, f は可測である. □

注意 1.21. X_1, X_2 は確率変数のとき, X_1X_2 も確率変数であることがわかる. さらに, $X_2 \neq 0$ のとき, X_1/X_2 も確率変数であることもわかる. □

定理 1.22. $X_1, X_2, \ldots$ は確率変数列のとき

$$\inf_n X_n \quad \sup_n X_n \quad \limsup_n X_n \quad \liminf_n X_n$$

も確率変数となる.

Proof. $r \in \mathbb{R}$ に対して

$$\left\{\omega \in \Omega : \inf_n X_n(\omega) < r\right\} = \bigcup_{n=1}^{\infty} \left\{\omega \in \Omega : X_n(\omega) < r\right\} \in \mathcal{A}$$

が $\inf_n X_n$ は確率変数であることがわかる. 同様に

$$\left\{\omega \in \Omega : \sup_n X_n(\omega) < r\right\} = \bigcap_{n=1}^{\infty} \left\{\omega \in \Omega : X_n(\omega) < r\right\}$$

から $\sup_n X_n$ も確率変数であることがわかる. 次に

$$\liminf_n X_n = \sup_n \left(\inf_{m \geq n} X_m\right)$$

に注意する. $Y_n := \inf_{m \geq n} X_m$ は各 $n \in \mathbb{N}$ に対して確率変数なので, $\sup_n Y_n$ も確率変数となる. 同様に

$$\limsup_n X_n = \inf_n \left(\sup_{m \geq n} X_m\right)$$

から $\limsup_n X_n$ も確率変数であることはわかる. □

注意 1.23. (i). $X_1, X_2, \ldots\ldots$ を確率変数列とする．定理 1.22 と定理 1.19 から

$$\begin{aligned}\Omega_0 :=& \big\{\omega \in \Omega :\ \lim_{n\to\infty} X_n \text{ は存在}\big\}\\ =& \big\{\omega \in \Omega :\ \limsup_{n\to\infty} X_n - \liminf_{n\to\infty} X_n = 0\big\}\end{aligned}$$

は可測集合となる．$\mathsf{Pr}(\Omega_0) = 1$ のとき，確率変数列 $\{X_n\}_{n=1}^{\infty}$ は**ほとんど確実に収束**するといい，a.s. と記す．$\mathbb{R} \cup \{\pm\infty\}$ に値を取る確率変数を X_∞ とおくと

$$X_\infty = \lim_{n\to\infty} X_n \qquad \text{(a.s.)}$$

と書くける．
(ii). 任意の $a, b \in \mathbb{R}$ とし，$\{-\infty\} \cup (-\infty, a), (a, b), (b, \infty) \cup \{+\infty\}$ によって生成される σ 集合族を $\mathcal{B}(\mathbb{R} \cup \{\pm\infty\})$ と書く．集合 $D \subset \Omega$ が定義域で，$\mathbb{R} \cup \{\pm\infty\}$ を値域ともつ関数 X は可測であるとは，$\forall B \in \mathcal{B}(\mathbb{R} \cup \{\pm\infty\})$ に対して

$$X^{-1}(B) = \{\omega \in \Omega :\ X(\omega) \in B\} \in \mathcal{A}$$

をみたすときをいう． □

注意 1.24. 上の定義の条件は

$$X^{-1}(B) \in \mathcal{A} \qquad (\forall B \in \mathcal{B}(\mathbb{R}))$$

と同値であることを示すことができる．補遺の定理 C.26 を参照せよ． □

定義 1.25. (1) 確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の確率変数 X に対して

$$\mathsf{F}_X(x) := \mathsf{Pr}(X \leq x) := \mathsf{Pr}(\{\omega \in \Omega;\ X(\omega) \leq x\}) \qquad (x \in \mathbb{R})$$

を X **の累積分布関数** (cumulative distribution function(c.d.f.)) という．また

$$\mathsf{P}_X(B) := \mathsf{Pr}(X \in B) := \mathsf{Pr}(\{\omega \in \Omega;\ X(\omega) \in B\}) \qquad (B \in \mathcal{B}(\mathbb{R}))$$

を X の**分布**という．$\mathsf{P}_X((-\infty, x]) = \mathsf{F}_X(x)$ である．
(2) X, Y を $(\Omega, \mathcal{A}, \mathsf{Pr})$ の確率変数とし，それぞれの c.d.f. を $\mathsf{F}_X, \mathsf{F}_Y$ とする．このとき

$$\mathsf{F}_X(x) = \mathsf{F}_Y(x) \qquad (\forall x \in \mathbb{R})$$

が成立するとき，X と Y の分布は同じであるという．これを $X \stackrel{d}{=} Y$ と書く．
(3) X が c.d.f. F を持つとき，$X \sim \mathsf{F}$ と書く．

注意 1.26. (1) P_X は可測空間 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度である．すなわち P_X は定義 1.6 をみたすことがわかる．
(2) c.d.f. F が与えられると

$$\mathsf{F}(x) = \mathsf{P}((-\infty, x]) \qquad (x \in \mathbb{R})$$

をみたす $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度 P が一意的に定まることが知られている．証明は補遺の命題 C.29 を参照せよ．このことにより X の c.d.f. と分布を同一視する．さらに $X \sim \mathsf{F}$ とし, c.d.f. F から定まる確率測度を P としたとき, $X \sim \mathsf{P}$ とも書く．
(3) X と Y の分布が同じとき

$$\mathsf{P}_X(B) = \mathsf{P}_Y(B) \qquad (\forall B \in \mathcal{B}(\mathbb{R}))$$

も成立する．

問 1.2. 注意 1.26(1) を示せ．逆像の性質を利用すること．

定理 1.27. X を確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の確率変数とし, F を X の c.d.f. とする．このとき F は次をみたす．

(1) F は非減少関数: $x < y \Longrightarrow \mathsf{F}(x) \leq \mathsf{F}(y)$.

(2) $\lim_{x\to\infty} \mathsf{F}(x) = 1$: $\lim_{x\to-\infty} \mathsf{F}(x) = 0$.

(3) F は右連続関数: $\lim_{y\to x+0} \mathsf{F}(y) = \mathsf{F}(x)$.

Proof. X の分布を P とする．すなわち $\mathsf{P}(B) := \mathsf{Pr}(X \in B)\ \big(\forall B \in \mathcal{B}(\mathbb{R})\big)$ である．
(1) $(-\infty, x]) \subset (-\infty, y]$ に注意して, 補題 1.7(4) を P に適用すればよい．
(2) $A_n = (-\infty, n]$ と $A_n = (-\infty, -n]$ として, P に補題 1.7(6)(7) を適用すればよい．
(3) $A_n = \big(-\infty, x + \frac{1}{n}\big]$ として, P に補題 1.7(7) を適用すればよい． □

問 1.3. 固定した $x \in \mathbb{R}$ に対して

$$(-\infty, x] = \bigcap_{n=1}^{\infty} \left(-\infty, x + \frac{1}{n}\right)$$

を示せ．

問 1.4. 定理 1.27(1) – (3) の証明を詳しく書き下せ．

定理 1.28. X を確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の確率変数とし, X の c.d.f. を F とする．このとき次の (1) – (3) は同値である．

(1) F は $\mathbb{R}$ 上の連続関数．

(2) $\mathsf{F}(x) = \mathsf{F}(x-)\,(\forall x \in \mathbb{R})$. ただし, $\mathsf{F}(x-) := \sup\{\mathsf{F}(y);\, y < x\}$ とした.

(3) $\mathsf{Pr}(X = x) = 0\,(\forall x \in \mathbb{R})$.

Proof. (1) ⇒ (2) の証明: $x_n = x - \dfrac{1}{n}\,(n = 1, 2, \ldots)$ とすれば

$$\lim_{n\to\infty} \mathsf{F}(x_n) = F(x-)$$

である. この式から, F が連続ならば, $\mathsf{F}(x) = \mathsf{F}(x-)$ となることがわかる.

(2) ⇒ (3) の証明:

$$\{X \le x_n\} \subset \{X \le x_{n+1}\}\,(n = 1, 2, \ldots), \quad \bigcup_{n=1}^{\infty}\{X \le x_n\} = \{X < x\}$$

であるので, 補題 1.7(6) より

$$\begin{aligned} \mathsf{F}(x-) &= \lim_{n\to\infty} \mathsf{F}(x_n) = \lim_{n\to\infty} \mathsf{Pr}(X \le x_n) = \mathsf{Pr}\left(\bigcup_{n=1}^{\infty}\{X \le x_n\}\right) \\ &= \mathsf{Pr}(X < x) \end{aligned}$$

を得る. さらに補題 1.7(4) より

$$\begin{aligned} \mathsf{Pr}(X = x) &= \mathsf{Pr}\Big(\{X \le x\} \setminus \{X < x\}\Big) = \mathsf{Pr}(X \le x) - \mathsf{Pr}(X < x) \\ &= F(x) - F(x-) \end{aligned} \tag{1.1}$$

であることからわかる. よって, $\mathsf{Pr}(X = 0) = 0$ となる.
(3) ⇒ (1) の証明: (1.1) から

$$\mathsf{Pr}(X = x) \implies \mathsf{F}(x) = \mathsf{F}(x-)$$

である. このことと F は右連続であることから F は連続となる. 以上から 3 つの条件は同値であることが示せた. □

問 1.5. 定理 1.28 の記号を踏襲する. 固定した $x \in \mathbb{R}$ に対して

$$\begin{aligned} \bigcup_{n=1}^{\infty}\left\{X \le x - \frac{1}{n}\right\} &= \bigcup_{n=1}^{\infty}\left\{\omega \in \Omega;\, X(\omega) \le x - \frac{1}{n}\right\} = \{\omega \in \Omega;\, X(\omega) < x\} \\ &= \{X < x\} \end{aligned}$$

を示せ.

注意 1.29. F が $\mathbb{R}$ 上で連続のとき, $\mathsf{Pr}(X = x) = 0\,(\forall x \in \mathbb{R})$ なので, $a < b$ に対して

$$\mathsf{Pr}(a \le X \le b) = \mathsf{Pr}(a < X < b) = \mathsf{Pr}(a < X \le b) = \mathsf{Pr}(a \le X < b)$$

である. □

定義 1.30. X を確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の確率変数とし, その c.d.f. を F と書く.
(1) X が高々可算個の集合 $\{x_1, x_2, \ldots\}$ 上にしか値を取らないとき, X を離散型であるという. この場合には

$$\mathsf{p}(x) := \mathsf{Pr}(X = x) = \mathsf{F}(x) - \mathsf{F}(x-) \qquad (x \in \{x_1, x_2, \ldots\})$$

で X の分布が特徴付けられる[6]. p を X の**確率関数** (probability mass function(p.m.f.)) と呼ぶ.
(2) $\mathsf{Pr}(X = x) = 0\,(\forall x \in \mathbb{R})$ のとき, X を連続型であるという. さらにある非負値関数 p で

$$\mathsf{F}(x) = \int_{-\infty}^{x} \mathsf{p}(t)\,dt \qquad (\forall x \in \mathbb{R})$$

をみたすものが存在するとき, p を X の**確率密度関数** (probability density function(p.d.f.)) という. 特に F がほとんどいたるところ[7]で微分可能ならば, ほとんどいたるところで

$$\dot{\mathsf{F}}(x) = \frac{d\mathsf{F}}{dx}(x) = \mathsf{p}(x)$$

となる.

注意 1.31. 確率変数 X の分布を P とする.
(1) X を離散型とし, $S = \{x \in \mathbb{R}; \mathsf{p}(x) > 0\}$ とおく. すると任意の Borel 集合 $B \subset \mathbb{R}$ に対して

$$\mathsf{P}(X \in B) = \sum_{x \in B \cap S} \mathsf{p}(x)$$

が成り立つことを示すことができる.
(2) X を連続型とし, その p.d.f. p が定義されるとする. このとき, 任意の Borel 集合 $B \subset \mathbb{R}$ に対して

$$\mathsf{P}(X \in B) = \int_B \mathsf{p}(x)\,dx$$

[6] $\mathsf{p}(x) = 0\,(x \notin \{x_1, x_2, \ldots\})$ となるので, p は $\mathbb{R}$ 上の関数であり, $0 \le \mathsf{p}(x) \le 1$ となる.

[7] この授業では, $\mathbb{R}$ から可算個の点を除いた集合上で微分可能と理解して差し支えない.

が成り立つことを示すことができる．厳密な証明は測度論の知識が必要となるので, 証明は省略する．

例 1.32. (1) $\Omega = \{0, 1\}$, $\mathcal{A} = 2^{\Omega}$, $\mathsf{Pr}(\{0\}) = \mathsf{Pr}(\{1\}) = 1/2$ とし

$$X : \Omega \ni \omega \mapsto X(\omega) = \omega \in \mathbb{R}$$

と定義すれば

$$\mathsf{Pr}(X = 0) = \mathsf{Pr}(X = 1) = \frac{1}{2}$$

となる．このとき, X の c.d.f. F と p.m.f. p はそれぞれ

$$\mathsf{F}(x) = \begin{cases} 0 & (x < 0) \\ \frac{1}{2} & (0 \leq x < 1) \\ 1 & (x \geq 1) \end{cases} \qquad \mathsf{p}(x) = \begin{cases} \frac{1}{2} & (x = 0,\, 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

となる．

(2) $\Omega = [0, 1]$, $\mathcal{A} = \mathcal{B}(\mathbb{R}) \cap [0, 1]$ とし

$$\mathsf{Pr}((a, b]) = b - a \qquad (0 \leq a < b \leq 1)$$

とする． さらに

$$X : \Omega \ni \omega \mapsto X(\omega) = \omega \in \mathbb{R}$$

と定義すれば, X の c.d.f. F と p.d.f. p はそれぞれ

$$\mathsf{F}(x) = \begin{cases} 0 & (x < 0) \\ x & (0 \leq x \leq 1) \\ 1 & (x > 1) \end{cases} \ ; \quad \mathsf{p}(x) = \begin{cases} 1 & (0 \leq x \leq 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

となる．

1.3 分位点関数

定義 1.33. c.d.f. F に対して, **分位点関数** (quantile function) $\mathsf{F}^{-} : (0, 1) \longrightarrow \mathbb{R}$ を

$$\mathsf{F}^{-}(y) := \inf\{x \in \mathbb{R} : \mathsf{F}(x) \geq y\}$$

で定義する．また, $y \in (0, 1)$ に対して, $\mathsf{F}^{-1}(y)$ を F **の** y **分位点**とよぶ．$1/2$ 分位点を**メディアン** (median) と呼ぶ．

注意 1.34. c.d.f. F が連続かつ $\operatorname{supp} \mathsf{F}$ 上で狭義単調増加のとき, F^{-} は F の逆関数になる． □

注意 1.35. F^- の定義から，任意の $0 < y < 1$ に対して，単調減少列 $\{x_n\}_{n=1}^\infty$ が存在して

$$\mathsf{F}(x_n) \geq y \quad \text{かつ} \quad \lim_{n\to\infty} x_n = \mathsf{F}^{-1}(y)$$

とできる．このとき，$\{x_n\}_{n=1}^\infty$ の単調減少性と F の右連続性から

$$\mathsf{F}\bigl(\mathsf{F}^-(y)\bigr) = \mathsf{F}\bigl(\lim_{n\to\infty} x_n\bigr) = \lim_{n\to\infty} \mathsf{F}(x_n) \geq y \tag{1.2}$$

となることがわかる．よって，F^- の定義における inf は達成される． □

注意 1.36. c.d.f. F は連続とする．c.d.f. の性質

$$\lim_{n\to\infty} \mathsf{F}(x) = 1 \quad \text{かつ} \quad \lim_{x\to-\infty} \mathsf{F}(x) = 0$$

と F の連続性から，中間値の定理を用いると任意の $0 < y < 1$ に対して，ある $x \in \mathbb{R}$ が存在して

$$\mathsf{F}(x) = y$$

をみたす．ここで，F^- の定義から

$$\mathsf{F}^-(y) \leq x \tag{1.3}$$

となる．なぜならば，$\mathsf{F}^-(y)$ は条件 $\mathsf{F}(z) \geq y$ をみたす $z \in \mathbb{R}$ の inf である．一方，$\mathsf{F}(x) = y$ だから $x \in \{z \in \mathbb{R} : \mathsf{F}(z) \geq y\}$ なので，$x \geq \mathsf{F}^-(y)$ がわかる．(1.3) と F の非減少性から

$$\mathsf{F}\bigl(\mathsf{F}^-(y)\bigr) \leq \mathsf{F}(x) = y \tag{1.4}$$

となる．一方，F^- の定義から

$$\mathsf{F}\bigl(\mathsf{F}^-(y)\bigr) \geq y \tag{1.5}$$

となる．(1.4) と (1.5) を合わせると

$$\mathsf{F}\bigl(\mathsf{F}^-(y)\bigr) = y$$

を得る．したがって

$$\mathsf{F}\text{ が連続} \Longrightarrow \mathsf{F}\bigl(\mathsf{F}^-(y)\bigr) = y\,(\forall 0 < y < 1)$$

となる．しかし，F が不連続ならば

$$\mathsf{F}\bigl(\mathsf{F}^-(y)\bigr) = y$$

とは限らない． □

例 1.37. $a < b\,(a, b \in \mathbb{R})$ とし，$X \sim \mathsf{U}(a, b)$ とする．すると X の c.d.f. は

$$\mathsf{F}(x) = \begin{cases} 0 & (x \leq a) \\ \dfrac{x-a}{b-a} & (a < x < b) \\ 1 & (x \geq b) \end{cases}$$

である．よって，F の分位点関数は

$$\mathsf{F}^{-}(y) = a + (b-a)y \qquad (y \in (0,\,1))$$

となる． □

例 1.38. $\mathsf{Pr}(X = 0) = \mathsf{Pr}(X = 1) = \dfrac{1}{2}$ とする．このとき，X の c.d.f. は

$$\mathsf{F}(x) = \begin{cases} 0 & (x < 0) \\ \dfrac{1}{2} & (0 \leq x < 1) \\ 1 & (x \geq 1) \end{cases}$$

となる．よって，F の分位点関数は

$$\mathsf{F}^{-}(y) = \begin{cases} 0 & \left(0 < y \leq \dfrac{1}{2}\right) \\ 1 & \left(\dfrac{1}{2} < y < 1\right) \end{cases}$$

となる．この場合，$y \neq \dfrac{1}{2}$ のとき

$$\mathsf{F}\big(\mathsf{F}^{-}(y)\big) \neq y$$

となる． □

定理 1.39. 分位点関数は以下の性質 (i) ～ (iii) をもつ．
(i) F^{-} は非減少である．
(ii) F^{-} は左連続である．すなわち，$\{y_n\}_{n=1}^{\infty}$ を任意の非減少列で $\lim_{n\to\infty} y_n = y$ としたとき

$$\lim_{n\to\infty} \mathsf{F}^{-}(y_n) = \mathsf{F}^{-}(y)$$

が成立する．
(iii) $\mathsf{F}^{-}(y) \leq x \iff y \leq \mathsf{F}(x)$ が成立する．

Proof. (i) $y_1 < y_2\,(y_1,\, y_2 \in (0,\, 1))$ に対して

$$\{x \in \mathbb{R} :\, \mathsf{F}(x) \geq y_1\} \supset \{x \in \mathbb{R} :\, \mathsf{F}(x) \geq y_2\}$$

となる．両辺の inf をとると

$$\mathsf{F}^{-}(y_1) = \inf\{x \in \mathbb{R} :\, \mathsf{F}(x) \geq y_1\} \leq \inf\{x \in \mathbb{R} :\, \mathsf{F}(x) \geq y_2\} = \mathsf{F}^{-}(y_2)$$

となる．よって, F^{-} の非減少性が証明できた.
(ii) $x_n := \mathsf{F}^{-}(y_n)\,(n \in \mathbb{N})$ とおく．数列 $\{y_n\}_{n=1}^{\infty}$ は非減少なので, (i) から $\{x_n\}_{n=1}^{\infty}$ も非減少列となる．さらに, $y_n \leq y$ から

$$x_n = \mathsf{F}^{-}(y_n) \leq \mathsf{F}^{-}(y) =: x_0 \tag{1.6}$$

となる．よって, (1.6) の両辺の極限を取ると

$$x := \lim_{n\to\infty} x_n \leq x_0 \tag{1.7}$$

となる．ここで, $x < x_0$ を仮定して矛盾を導く．そのために

$$\epsilon := \frac{1}{2}(x_0 - x) > 0 \iff x = x_0 - 2\epsilon$$

とおく．F^{-} の定義から導出された (1.5) に注意すると

$$y_n \leq \mathsf{F}\big(\mathsf{F}^{-}(y_n)\big) \leq \mathsf{F}(x_n) \leq \mathsf{F}(x + \epsilon) = \mathsf{F}(x_0 - \epsilon)$$

となる．すると

$$y = \lim_{n\to\infty} y_n \leq \mathsf{F}(x_0 - \epsilon)$$

となる．しかし, F^{-} の定義から

$$\mathsf{F}(x_0 - \epsilon) < y$$

となる．なぜならば, $\mathsf{F}^{-}(y) = x_0$ なので, x_0 は条件 $\mathsf{F}(z) \geq y$ をみたす z の inf である．$x_0 > x_0 - \epsilon$ だから, $x_0 - \epsilon \notin \{z \in \mathbb{R} :\, \mathsf{F}(z) \geq y\}$ となる．よって, $\mathsf{F}(x_0 - \epsilon) < y$ となることがわかる．以上から $y \leq \mathsf{F}(x_0 - \epsilon)$ かつ $y > \mathsf{F}(x_0 - \epsilon)$ となるので, 矛盾．よって

$$\mathsf{F}^{-}(y) = x_0 = x = \lim_{n\to\infty} x_n = \lim_{n\to\infty} \mathsf{F}^{-}(y_n)$$

となるので, F^{-} の左連続性がわかる.
(iii) F^{-} の非減少性から

$$y \leq \mathsf{F}(x) \implies \mathsf{F}^{-}(y) \leq \mathsf{F}^{-}\big(\mathsf{F}(x)\big) \leq x$$

となる. 最後の不等号は

$$\mathsf{F}^{-}\big(\mathsf{F}(x)\big) = \inf\big\{z \in \mathbb{R} :\ \mathsf{F}(z) \geq \mathsf{F}(x)\big\}$$

である. しかし, $x \in \big\{z \in \mathbb{R} :\ \mathsf{F}(z) \geq \mathsf{F}(x)\big\}$ であるので, $x \geq \mathsf{F}^{-}\big(\mathsf{F}(x)\big)$ がわかる.

逆に, $\mathsf{F}^{-}(y) \leq x$ ならば, (1.5) と F の非減少性から

$$y \leq \mathsf{F}\big(\mathsf{F}^{-}(y)\big) = \mathsf{F}(x)$$

がわかる. よって, (iii) は証明された. □

系 1.40. F を c.d.f. とする. このとき, $U \sim \mathsf{U}(0, 1)$ に対して

$$X := \mathsf{F}^{-}(U) \sim \mathsf{F}$$

となる.

Proof. 定理 1.39(iii) から, $\forall x \in \mathbb{R}$ に対して

$$\{X \leq x\} \iff \{U \leq \mathsf{F}(x)\}$$

となる. よって

$$\mathsf{Pr}(X \leq x) = \mathsf{Pr}(Y \leq \mathsf{F}(x)) = \mathsf{F}(x)$$

を得る. □

1.4 主な 1 次元分布

1.4.1 離散型確率変数

Bernoulli 分布

$0 \leq \theta \leq 1$ とする. 確率変数 X は**母数** θ **の** Bernoulli **分布**に従うとは, X の p.m.f. $\mathsf{p}(\,\cdot\,|\,\theta)$ が

$$\mathsf{p}(x|\,\theta) = p(x) = \begin{cases} \theta^{x}(1-\theta)^{1-x} & (x = 0,\, 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

のときをいう. このことを $X \sim \mathsf{Ber}(\theta)$ と記す.

2 項分布

$n \in \mathbb{N},\, 0 \le \theta \le 1$ とする．確率変数 X は**母数** $(n,\, \theta)$ **の 2 項分布**に従うとは, X の p.m.f. $\mathsf{p}(\,\cdot\,|\, n,\, \theta)$ が

$$\mathsf{p}(x|\, n,\, \theta) = \mathsf{p}(x) = \begin{cases} \begin{pmatrix} n \\ x \end{pmatrix} \theta^x (1-\theta)^{n-x} & (x = 0,\, 1,\, \ldots,\, n) \\ 0 & (\text{その他の場合}) \end{cases}$$

のときをいう．ただし

$$\begin{pmatrix} n \\ x \end{pmatrix} = \frac{n!}{x!\,(n-x)!}, \qquad 0! = 1.$$

このことを $X \sim \mathsf{Bino}(n,\, \theta)$ と記す.

幾何分布

$0 < \theta < 1$ とする．確率変数 X は (X は**母数** θ **の幾何分布**に従うとは, X の p.m.f. $\mathsf{p}(\,\cdot\,|\, \theta)$ が

$$\mathsf{p}(x|\, \theta) = \mathsf{p}(x) = \begin{cases} \theta(1-\theta)^{x-1} & (x = 1,\, 2,\, \ldots) \\ 0 & (\text{その他の場合}) \end{cases}$$

のときをいう．このことを $X \sim \mathsf{Geom}(\theta)$ と記す.

Poisson 分布

$\theta > 0$ とする．確率変数 X は**母数** θ **の Poisson 分布**に従うとは, X の p.m.f. $\mathsf{p}(\,\cdot\,|\, \theta)$ が

$$\mathsf{p}(x|\, \theta) = \mathsf{p}(x) = \begin{cases} e^{-\theta} \dfrac{\theta^x}{x!} & (x = 0,\, 1,\, 2,\, \ldots) \\ 0 & (\text{その他の場合}) \end{cases}$$

のときをいう．このことを $X \sim \mathsf{Poi}(\theta)$ と記す.

1.4.2　連続型確率変数

正規分布

$\mu \in \mathbb{R},\, 0 < \sigma < \infty$ とする．確率変数 X は**平均** μ**, 分散** σ^2 **の正規分布**に従うとは, X の p.d.f. $\mathsf{p}(\,\cdot\,|\, \mu,\, \sigma^2)$ が

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

のときをいう．ただし $\exp(x) = e^x$ である．このことを $X \sim \mathsf{N}(\mu,\, \sigma^2)$ と記す. $\mu = 0,\, \sigma = 1$ のときの分布を**標準正規分布**という.

ガンマ分布

$\alpha > 0, \beta > 0$ とする. 確率変数 X は**母数** (α, β) **のガンマ分布**に従うとは, X の p.d.f. $\mathsf{p}(\cdot\,|\,\alpha, \beta)$ が

$$\mathsf{p}(x|\,\alpha, \beta) = \mathsf{p}(x) = \begin{cases} \dfrac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & (x > 0) \\ 0 & (\text{その他の場合}) \end{cases}$$

のときをいう. ただし

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx \qquad (\alpha > 0)$$

である. このことを $X \sim \mathsf{Ga}(\alpha, \beta)$ と記す. $\lambda > 0$ とし, $\alpha = 1, \beta = \dfrac{1}{\lambda}$ とおく. このとき, $\mathsf{Ga}(1, 1/\lambda)$ を**母数** λ **の指数分布**といい, $\mathsf{Exp}(\lambda)$ と書く. また, $p \in \mathbb{N}$ とし, $\alpha = \dfrac{p}{2}, \beta = 2$ とおく. このとき, $\mathsf{Ga}(p/2, 2)$ を**自由度** p **の** χ^2 **分布**といい, χ_p^2 と記す.

1.5　2 次元の分布

1.5.1　同時確率関数と確率密度関数

確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の離散型確率変数 (X, Y) の対に対して, 同時 p.m.f. p を

	Y = 0	Y = 1	
X = 0	1/9	2/9	1/3
X = 1	2/9	4/9	2/3
	1/3	2/3	

で定める. たとえば

$$\mathsf{p}(1, 1) = \mathsf{Pr}(X = 1, Y = 1) = \frac{4}{9}$$

である.

定義 1.41. 確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の確率変数 X, Y は連続型とする. $\mathbb{R}^2$ 上の非負値実数値関数 p が確率ベクトル (X, Y) の **同時確率密度関数** (同時 p.d.f.) であるとは, 次の条件をみたすときをいう.

(1) $\mathsf{p}(x, y) \geq 0\, (\forall (x, y) \in \mathbb{R}^2)$.

(2) $\int_{-\infty}^\infty \int_{-\infty}^\infty \mathsf{p}(x, y)\, dxdy = 1$.

(3) $\forall A \in \mathcal{B}(\mathbb{R}^2)$ に対して

$$\mathsf{Pr}((X,\,Y) \in A) = \int\int_A \mathsf{p}(x,\,y)\,dxdy.$$

注意 1.42. $(X,\,Y)$ の同時累積分布関数 (同時 c.d.f.) F を

$$\begin{aligned}\mathsf{F}(x,\,y) :&= \mathsf{Pr}(X \le x,\, Y \le y) &\quad (1.8)\\ &:= \mathsf{Pr}(\{\omega \in \Omega;\, X(\omega) \le x,\, Y(\omega) \le y\}) \qquad (\forall (x,\,y) \in \mathbb{R}^2)\end{aligned}$$

で定義する．$(X,\,Y)$ の同時分布 P を

$$\mathsf{P}(A) = \mathsf{Pr}\big((X,\,Y) \in A\big) \quad (\forall A \in \mathcal{B}(\mathbb{R}^2)) \tag{1.9}$$

で定義する．どの確率変数の c.d.f. と分布であるかを明示したいときには, $\mathsf{F}_{(X,Y)}$ または $\mathsf{P}_{(X,Y)}$ と記す．細かなことであるが, (1.8) の F によって $(\mathbb{R}^2,\,\mathcal{B}(\mathbb{R}^2))$ 上の確率測度 P を一意的に定めていることができる．この点に関しては補遺 C 章の測度論の知識が必要となる． □

例 1.43. 連続型確率変数のベクトル $(X,\,Y)$ は同時 p.d.f.

$$\mathsf{p}(x,\,y) = \begin{cases} x + y & (0 \le x \le 1,\, 0 \le y \le 1)\\ 0 & (\text{その他の場合})\end{cases}$$

を持つとする．このとき集合 $\{(x,\,y);\, \mathsf{p}(x,\,y) > 0\}$ に注意すれば

$$\begin{aligned}\int_0^1\int_0^1 \mathsf{p}(x,\,y)\,dxdy &= \int_0^1\left[\int_0^1 x\,dx\right]dy + \int_0^1\left[\int_0^1 y\,dy\right]dx\\ &= \int_0^1 \frac{1}{2}\,dy + \int_0^1 \frac{1}{2}\,dx = 1.\end{aligned}$$

□

1.5.2　周辺分布

定義 1.44. $(X,\,Y)$ を確率空間 $(\Omega,\,\mathcal{A},\,\mathsf{Pr})$ 上で定義された確率ベクトルとする．(1) $(X,\,Y)$ が離散型で同時 p.m.f. p を持つとする．X の**周辺確率関数** (周辺 p.m.f.) を

$$\mathsf{p}_X(x) = \mathsf{Pr}(X = x) = \sum_{y \in S_Y} \mathsf{Pr}(X = y,\, Y = y) = \sum_{y \in S_Y} \mathsf{p}(x,\,y) \qquad (\forall x \in S_X)$$

で定義する．ただし

$$\begin{aligned}S_X &:= \{x \in \mathbb{R};\, \text{ある } y \in \mathbb{R} \text{ に対して},\, \mathsf{p}(x,\,y) > 0\}\\ S_Y &:= \{y \in \mathbb{R};\, \text{ある } x \in \mathbb{R} \text{ に対して},\, \mathsf{p}(x,\,y) > 0\}\end{aligned}$$

である.
(2) (X, Y) は連続型とし, 同時 p.d.f. p を持つとする. このとき X の**周辺確率密度関数** (周辺 p.d.f.) を

$$\mathsf{p}_X(x) = \int_{-\infty}^{\infty} \mathsf{p}(x, y)\, dy \qquad (\forall x \in \mathbb{R})$$

で定義し, Y の**周辺確率密度関数** (周辺 p.d.f.) を

$$\mathsf{p}_Y(x) = \int_{-\infty}^{\infty} \mathsf{p}(x, y)\, dx \qquad (\forall y \in \mathbb{R})$$

で定義する.

注意 1.45. 連続型確率ベクトル (X, Y) に対して

$$\begin{aligned}\mathsf{F}_X(x) := \mathsf{Pr}(X \leq x) &= \int\int_{(s,\, t) \in \mathbb{R}^2;\, s \leq x} \mathsf{p}(s, t)\, dsdt \\ &= \int_{-\infty}^{x} \left\{ \int_{-\infty}^{\infty} \mathsf{p}(s, t)\, ds \right\} dt = \int_{-\infty}^{x} \mathsf{p}_X(s)\, ds\end{aligned}$$

となるので, X の周辺 p.d.f. と X の p.d.f. は同じである. □

注意 1.46. F を確率ベクトル (X, Y) の同時 c.d.f. とし, F_X を X の周辺 c.d.f. とする. このとき $x \in \mathbb{R}$ に対して

$$\mathsf{F}_X(x) = \lim_{y \to \infty} \mathsf{F}(x, y); \ \mathsf{F}_Y(y) = \lim_{x \to \infty} \mathsf{F}(x, y)$$

が成立する. □

問 1.6. 注意 1.46 の式を証明せよ.

1.5.3 独立性な分布と条件付き分布

定義 1.47. 確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の 2 つの確率変数 X と Y は**独立**であるとは, $\forall A, B \in \mathcal{B}(\mathbb{R})$ に対して

$$\mathsf{Pr}(X \in A, Y \in B) = \mathsf{Pr}(X \in A)\mathsf{Pr}(Y \in B)$$

が成り立つときをいう. 独立でないとき X と Y は**従属**であるという.

定理 1.48. p を確率ベクトル (X, Y) の同時 p.d.f. とし, p_X と p_Y を X と Y それぞれの周辺 p.d.f. とする. このとき

$$X \text{ と } Y \text{ は独立} \iff \mathsf{p}(x, y) = \mathsf{p}_X(x)\mathsf{p}_Y(y) \quad (\forall x, y \in \mathbb{R}).$$

Proof. 独立ならば, 同時 p.d.f. が周辺 p.d.f. の積で表現されることの証明は易しい. 逆はこの講義の範囲 (測度の拡張の議論が必要となる!) を超えるので, 信じることにする. □

注意 1.49. (1) 連続型の場合には, 測度 0 の集合を除いて[8]p.d.f. は定義されるので, 定義 1.47 の書き方はやや数学的な厳密性に欠ける.
(2) 離散型確率変数の場合には, p.d.f. を p.m.f. に置き換えればよい. □

例 1.50. 連続型確率変数 X と Y は独立で同じ p.d.f.

$$\mathsf{p}(x) = \begin{cases} 2x & (0 \leq x \leq 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

を持つとする. このとき確率 $\mathsf{Pr}(X+Y \leq 1)$ を求めてみよう. 独立性より (X, Y) の同時 p.d.f. は

$$\mathsf{p}_{(X,Y)}(x, y) = \begin{cases} 4xy & (0 \leq x \leq 1,\, 0 \leq y \leq 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

となる. したがって

$$\begin{aligned} \mathsf{Pr}(X+Y \leq 1) &= \int\int_{x+y\leq 1} \mathsf{p}_{(X,Y)}(x, y)\, dxdy = 4\int_0^1 x \left\{ \int_0^{1-x} y\, dy \right\} dx \\ &= 4\int_0^1 x\frac{(1-x)^2}{2}\, dx = \frac{1}{6}. \end{aligned}$$

□

定理 1.51. 連続型確率ベクトル (X, Y) は同時 p.d.f. $\mathsf{p}(x, y)$ を持つとする. ただし $\{(x, y) \in \mathbb{R}^2;\, \mathsf{p}(x, y) > 0\}$ は矩形[9]とする. このときある非負値関数 q と r が存在して

$$\mathsf{p}(x, y) = q(x)r(y)$$

と書けるとき, X と Y は独立である.

Proof. 積分をして確かめればよい. □

例 1.52. 連続型確率ベクトル (X, Y) は同時 p.d.f.

$$\mathsf{p}(x, y) = \begin{cases} 2e^{-(x+y)} & (x \geq 0,\, y \geq 0) \\ 0 & (\text{その他の場合}) \end{cases}$$

[8]この用語は定義されていない. 「有限個の点を除いて」と理解してもこの講義の内容の範囲では問題ない.
[9]有界でなくともよい.

を持つとする. $\{(x,\,y)\in\mathbb{R}^2;\,\mathsf{p}(x,\,y)>0\}=[0,\,\infty)^2$ なので

$$q(x)=\begin{cases}2e^{-x} & (x\geq 0)\\ 0 & (\text{その他の場合})\end{cases}\qquad r(y)=\begin{cases}e^{-y} & (x\geq 0)\\ 0 & (\text{その他の場合})\end{cases}$$

と取れば

$$\mathsf{p}(x,\,y)=q(x)r(y)$$

となるので, X と Y は独立. □

定義 1.53. (1) 離散型確率ベクトル $(X,\,Y)$ は同時 p.m.f. $\mathsf{p}(x,\,y)$ を持つとする. $\mathsf{p}_Y(y)>0$ なる y に対して, $Y=y$ を与えたときの X の**条件付き確率関数** (条件付き p.m.f.) を

$$\mathsf{p}_{X|Y}(x|\,y)=\mathsf{Pr}(X=x|\,Y=y)=\frac{\mathsf{Pr}(X=x,\,Y=y)}{\mathsf{Pr}(Y=y)}=\frac{\mathsf{p}(x,\,y)}{\mathsf{p}_Y(y)}\quad(x\in\mathbb{R})$$

で定義する.
(2) 連続型確率ベクトル $(X,\,Y)$ は同時 p.d.f. $\mathsf{p}(x,\,y)$ を持つとする. $\mathsf{p}_Y(y)>0$ なる y に対して, $Y=y$ を与えたときの X の**条件付き確率密度関数** (条件付き p.d.f.) を

$$\mathsf{p}_{X|Y}(x|\,y)=\frac{\mathsf{p}(x,\,y)}{\mathsf{p}_Y(y)}\quad(x\in\mathbb{R})$$

で定義する.

注意 1.54. $(X,\,Y)$ が連続型確率ベクトルのとき, $\mathsf{p}_Y(y)>0$ なる $y\in\mathbb{R}$ に対して

$$\mathsf{Pr}(X\in A|\,Y=y):=\int_A\mathsf{p}_{X|Y}(x|\,y)\,dx\qquad\big(\forall A\in\mathcal{B}(\mathbb{R})\big)$$

と形式的に定義する. □

例 1.55. 連続型確率ベクトル $(X,\,Y)$ は同時 p.d.f.

$$\mathsf{p}(x,\,y)=\begin{cases}x+y & (0\leq x\leq 1,\,0\leq y\leq 1)\\ 0 & (\text{その他の場合})\end{cases}$$

を持つとする. このとき $\mathsf{Pr}(X\leq 1/4|\,Y=1/3)$ を求めてみよう. まず $0<y<1$ に対して

$$\mathsf{p}_Y(y)=\int_{-\infty}^{\infty}\mathsf{p}(x,\,y)\,dx=y+\int_0^1 y\,dx=y+\frac{1}{2}.$$

したがって

$$\mathsf{p}_Y(y) = \begin{cases} y + \frac{1}{2} & (0 \leq y \leq 1), \\ 0 & (\text{その他の場合}) \end{cases}$$

となる．よって $0 < y < 1$ と $0 \leq x \leq 1$ に対して

$$\mathsf{p}_{X|Y}(x|\,y) = \frac{\mathsf{p}(x,\,y)}{\mathsf{p}_Y(y)} = \frac{x+y}{y+\frac{1}{2}}$$

これより $0 < y < 1$ のとき，

$$\mathsf{p}_{(X|Y)}(x|\,y) = \begin{cases} \dfrac{x+y}{y+\frac{1}{2}} & (0 \leq x \leq 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

となる．注意 1.54 より

$$\mathsf{Pr}\left(X \leq \frac{1}{4} \middle| Y = \frac{1}{3}\right)6 = \int_0^{1/4} p_{X|Y}\left(x \middle| \frac{1}{3}\right) dx = \int_0^{1/4} \frac{x+\frac{1}{3}}{\frac{1}{3}+\frac{1}{2}}\, dx = \frac{11}{80}.$$

□

1.6　多次元分布と i.i.d. 標本

$X_1,\, X_2,\, \ldots,\, X_n$ を確率空間 $(\Omega,\, \mathcal{A},\, \mathsf{Pr})$ 上の確率変数とし

$$\boldsymbol{X} = (X_1,\, X_2,\, \ldots,\, X_n)$$

と書く．$\boldsymbol{X}$ を確率ベクトルという．

定義 1.56. $X_1,\, X_2,\, \ldots,\, X_n$ が独立であるとは，すべての Borel 集合 $A_1,\, A_2,\, \ldots,\, A_n \in \mathcal{B}(\mathbb{R})$ に対して

$$\mathsf{Pr}(X_1 \in A_1,\, X_2 \in A_2,\, \ldots,\, X_n \in A_n) = \prod_{j=1}^{n} \mathsf{Pr}(X_j \in A_j) \tag{1.10}$$

が成立するときである．

$\boldsymbol{X}$ の同時 p.d.f. を $\mathsf{p}(x_1,\, x_2,\, \ldots,\, x_n)$ と書き，各 $X_j\ (j = 1,\, 2,\, \ldots,\, n)$ の周辺 p.d.f. を p_{X_j} と書くことにする．

注意 1.57. (1.10) を示すためには

$$\mathsf{p}(x_1,\, x_2,\, \ldots,\, x_n) = \prod_{j=1}^{n} \mathsf{p}_{X_j}(x_j) \quad (\forall x_1,\, x_2,\, \ldots,\, x_n \in \mathbb{R})$$

を示せばよい．□

定義 1.58. $X_1, X_2, \ldots, X_n$ は独立で, 各 $X_j\,(j = 1, 2, \ldots, n)$ は同じ c.d.f. F を持つとき, $X_1, X_2, \ldots, X_n$ は**独立同一分布に従う** (i.i.d. =identically and independently distributed) といい

$$X_1, X_2, \ldots, X_n \sim \text{i.i.d.}\, \mathsf{F}$$

と書く[10]. $X_1, X_2, \ldots, X_n$ は累積分布関数 F から標本の大きさが n のランダム標本ともいう.

1.6.1　重要な多次元分布モデル

この本で取り上げる代表的な多次元分布モデルをあげておく. 離散型分布は多項分布, 連続型分布は多変量正規分布である.

多項分布

2 項分布を多変量にしたものが多項分布である. $d, n \in \mathbb{N}$ とし, $\boldsymbol{p} = (p_1, p_2, \ldots, p_d)$, $p_j\,(0 \le p_j \le 1; , j = 1, 2, \ldots, d)$, $\sum_{j=1}^{d} p_j = 1$ とする. $\boldsymbol{X} = (X_1, X_2, \ldots, X_d)$ は**母数** $(d, \boldsymbol{p})$ **の多項分布**に従うとは, $\boldsymbol{X}$ の p.m.f. が

$$\mathsf{p}(\boldsymbol{x}|\, n, d, \boldsymbol{p}) = \begin{cases} \begin{pmatrix} d \\ x_1, x_2, \ldots, x_d \end{pmatrix} p_1^{x_1} p_2^{x_2} \times \cdots \times p_d^{x_d} & (x_j = 0, 1, \ldots, n;\ j = 1, 2, \ldots, d) \\ 0 & (\text{その他の場合}) \end{cases}$$

で与えられたときをいう. ただし

$$\begin{pmatrix} d \\ x_1 x_2, \ldots, x_d \end{pmatrix} = \frac{n!}{x_1! x_2! \cdots x_d!}, \quad 0! = 1$$

である. これを $\boldsymbol{X} \sim \mathsf{Multi}_d(n, \boldsymbol{p})$ と記す.

補題 1.59. $\boldsymbol{X} \sim \mathsf{Multi}_d(n, \boldsymbol{p})$ とする. ただし, $\boldsymbol{p} = (p_1, p_2, \ldots, p_d)$ と $\boldsymbol{X} = (X_1, X_2, \ldots, X_d)$ とする. このとき $X_j\,(j = 1, 2, \ldots, d)$ の周辺分布は $\mathsf{Bino}(n, p_j)$ である.

Proof. x_2 から x_d について同時 p.m.f. の和をとればよい. □

[10] p.d.f. p を使い

$$X_1, X_2, \ldots, X_n \sim \text{i.i.d.}\, \mathsf{p}$$

と書く. $\sim$ の右側には, 確率測度/確率分布/p.d.f./p.m.f./$\mathsf{N}(0, 1)$ など分布を特定するものを書いてよいことにする.

多変量正規分布

$$\boldsymbol{Z} := \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_d \end{pmatrix}, \quad Z_1,\, Z_2,\, \ldots,\, Z_d \sim \mathsf{N}(0,\,1)$$

とする[11]. $\boldsymbol{Z}$ の同時 p.d.f. は

$$\begin{aligned}\mathsf{p}(\boldsymbol{z}) = \mathsf{p}(z_1,\, z_2,\, \ldots,\, z_d) &= \prod_{j=1}^{d} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z_j^2\right\} = \frac{1}{(2\pi)^{d/2}} \exp\left\{-\frac{1}{2}\sum_{j=1}^{d} z_j^2\right\} \\ &= \frac{1}{(2\pi)^{d/2}} \exp\left\{-\frac{1}{2}\boldsymbol{z}^\top \boldsymbol{z}\right\} \quad (\boldsymbol{z} \in \mathbb{R}^d)\end{aligned}$$

ただし $\boldsymbol{z} = (z_1,\, z_2,\, \ldots,\, z_d)^\top$ である. これを $\boldsymbol{Z} \sim \mathsf{N}_d(\mathbf{0},\, \boldsymbol{I}_d)$ と記す. ただし $\boldsymbol{I}_d$ は d 次単位行列である. さらに定義より

$$\int \cdots \int \mathsf{p}(z_1,\, z_2,\, \ldots,\, z_d)\, dz_1 dz_2 \cdots dz_d = 1$$

となっていることもわかる.

次に

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_d \end{pmatrix} \in \mathbb{R}^d, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1d} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{d1} & \sigma_{d2} & \cdots & \sigma_{dd} \end{pmatrix}, \quad \begin{array}{l} \sigma_{ij} = \sigma_{ji} \\ (i,\, j = 1,\, 2, \ldots,\, d) \end{array}$$

とする. ただし $\boldsymbol{\Sigma}$ は正定値対称行列とする.　このときある $d \times d$ の正則行列 $\mathbf{A}$ が存在して

$$(1)\ \mathbf{A}\ \text{は対称行列},\ (2)\ \boldsymbol{\Sigma} = \mathbf{A}^2$$

と取れる. これを $\boldsymbol{\Sigma}$ の平方根といい, $\boldsymbol{\Sigma}^{1/2}$ と書くことにする. これを用いて

$$\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\boldsymbol{Z}$$

と定めたとき, $\boldsymbol{X}$ の分布を $\mathsf{N}_d(\boldsymbol{\mu},\, \boldsymbol{\Sigma})$ と記す. このとき $\boldsymbol{X}$ の同時 p.d.f. は

$$\mathsf{p}_{\boldsymbol{X}}(\boldsymbol{x}|\, \boldsymbol{\mu},\, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{d/2}\mathrm{Det}\,(\boldsymbol{\Sigma})^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right\} \quad (\boldsymbol{x} \in \mathbb{R}^d) \tag{1.11}$$

で与えられる. この分布を**平均ベクトル $\boldsymbol{\mu}$, 分散共分散行列 Σ の** d **変量正規分布**という.

[11] この節ではベクトルは縦とする. 下の式では縦ベクトルと横ベクトルを混ぜて表現している. これは記号の乱用であるが, $f((z_1,\, z_2,\, \ldots,\, z_d)^\top)$ などと書くのは煩わしい.

注意 1.60. (1.11) の導出は以下のように行うことができる．一般に $\boldsymbol{Z} = (Z_1, Z_2, \ldots, Z_d)^\top$ を d 次元確率ベクトルとし，その同時 p.d.f. を $p_{\boldsymbol{Z}}$ と書くことにする．さらに　$\mathbb{X} := \{\boldsymbol{z} \in \mathbb{R}^d; \mathsf{p}_{\boldsymbol{Z}}(\boldsymbol{z}) > 0)\}$, 関数

$$\boldsymbol{h}(\cdot) = (h_1(\cdot), h_2(\cdot), \ldots, h_d(\cdot))^\top : \mathbb{X} \longrightarrow \boldsymbol{h}(\mathbb{X})$$

は 1 対 1 とし

$$X_j = h_j(Z_1, Z_2, \ldots, Z_d) \quad (j = 1, 2, \ldots, d), \boldsymbol{X} \quad = (X_1, X_2, \ldots, X_d)^\top$$

とおく．$\boldsymbol{h}$ は 1 対 1 なので，$\boldsymbol{h}$ の逆写像を

$$\boldsymbol{h}^{-1} = (h_1^{-1}, h_2^{-1}, \ldots, h_d^{-1})^\top : \boldsymbol{h}(\mathbb{X}) \longrightarrow \mathbb{X}$$

が存在して，$\boldsymbol{Z} = \boldsymbol{h}^{-1}(\boldsymbol{X})$ となる．$\boldsymbol{X}$ の同時 p.d.f. を求めるために，$\boldsymbol{h}^{-1}(\boldsymbol{x})$ の Jacobian $\mathbf{J}_{\boldsymbol{h}^{-1}}(\boldsymbol{x})$ を

$$\mathbf{J}_{\boldsymbol{h}^{-1}}(\boldsymbol{x}) = \mathrm{Det}\begin{pmatrix} \frac{\partial h_1^{-1}(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial h_1^{-1}(\boldsymbol{x})}{\partial x_d} \\ \vdots & \ddots & \vdots \\ \frac{h_d^{-1}(\boldsymbol{x})}{\partial x_1} & \cdots & \frac{\partial h_d^{-1}(\boldsymbol{x})}{\partial x_d} \end{pmatrix}$$

で定め，$\mathbf{J}_{\boldsymbol{h}^{-1}}(\boldsymbol{x}) \neq 0 \, (x \in \mathbb{X})$ と仮定[12]する．このとき

$$\mathsf{p}_{\boldsymbol{X}}(\boldsymbol{x}) = \left|\mathbf{J}_{\boldsymbol{h}^{-1}}(\boldsymbol{x})\right| \mathsf{p}_{\boldsymbol{Z}}(h^{-1}(\boldsymbol{x})) \tag{1.12}$$

となることが補題の定理 C.84 からわかる．

$\boldsymbol{h}(\boldsymbol{z}) = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\boldsymbol{z} \Longleftrightarrow \boldsymbol{h}^{-1}(\boldsymbol{x}) = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{x} - \boldsymbol{\mu})$ より

$$\mathbf{J}_{\boldsymbol{h}^{-1}}(\boldsymbol{x}) = \mathrm{Det}\,(\boldsymbol{\Sigma})^{-1/2} = \frac{1}{\sqrt{\mathrm{Det}\,(\boldsymbol{\Sigma})}}$$

を (1.12) に代入すれば，(1.11) はわかる．以上の議論から

$$\int\int\cdots\int \mathsf{p}_{\boldsymbol{X}}(\boldsymbol{x})\,dx_1 dx_2 \cdots dx_d = 1$$

となっていることもわかる．　□

定理 1.61. $\boldsymbol{X} = (X_1, X_2, \ldots, X_d)^\top \sim \mathsf{N}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ とする．ただし $\boldsymbol{\mu} \in \mathbb{R}^d$ で，$\boldsymbol{\Sigma}$ は $d \times d$ の正定値行列である．このとき次が成立する．

(1) $X_j \, (j = 1, 2, \ldots, d) \sim \mathsf{N}(\mu_j, \sigma_{jj})$．ただし μ_j は $\boldsymbol{\mu}$ の第 j 成分，σ_{jj} は $\boldsymbol{\Sigma}$ の第 j 対角成分である．

[12] 実は，a.e$\mathbb{X}$ でよい．

(2) $X_\ell = x_\ell\,(\ell = 1, 2, \ldots, d)$ を与えたときの $X_j\,(j \neq \ell)$ の条件付き分布は

$$X_j |\, X_\ell = x_\ell \ \sim\ \mathsf{N}\left(\mu_j + \frac{\sigma_{\ell\ell}}{\sigma_{jj}}(x_\ell - \mu_\ell),\, \sigma_{jj} - \frac{\sigma_{j\ell}\sigma_{\ell j}}{\sigma_{\ell\ell}}\right).$$

(3) 定数ベクトル $\boldsymbol{c} \in \mathbb{R}^d\,(\boldsymbol{c} \neq \boldsymbol{0})$ に対して

$$\boldsymbol{c}^\top \boldsymbol{X} \ \sim\ \mathsf{N}(\boldsymbol{c}^\top \boldsymbol{\mu},\, \boldsymbol{c}^\top \boldsymbol{\Sigma} \boldsymbol{c}).$$

(4) $V := (\boldsymbol{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{X} - \boldsymbol{\mu}) \ \sim\ \chi^2_d.$

Proof. (1) $j = 1$ として示す．$\boldsymbol{\Sigma}$ の $(i, j)\,(i, j = 1, 2, \ldots, d)$ 成分を σ_{ij} と書き, $(d-1) \times (d-1)$ 行列 $\boldsymbol{\Sigma}_1$ と $(d-1) \times 1$ ベクトル $\boldsymbol{\sigma}_1$ をそれぞれ

$$\boldsymbol{\Sigma}_1 = (\sigma_{ij})_{i,\, j = 2, 3, \ldots, d}\,, \quad \boldsymbol{\sigma}_1 (\sigma_{21},\, \sigma_{31},\, \ldots,\, \sigma_{d1})^\top$$

で定める．すると

$$\boldsymbol{\Sigma} = \left(\begin{array}{c|c} \sigma_{11} & \boldsymbol{\sigma}_1^\top \\ \hline \boldsymbol{\sigma}_1 & \boldsymbol{\Sigma}_2 \end{array}\right)$$

となる．このとき

$$\boldsymbol{\Sigma} = \left(\begin{array}{c|c} 1 & \boldsymbol{0}_{d-1}^\top \\ \hline \sigma_{11}^{-1}\boldsymbol{\sigma}_1 & \boldsymbol{I}_{d-1} \end{array}\right) \left(\begin{array}{c|c} \sigma_{11} & \boldsymbol{0}_{d-1}^\top \\ \hline \boldsymbol{0}_{d-1} & \boldsymbol{\Sigma}_2 - \sigma_{11}^{-1}\boldsymbol{\sigma}_1\boldsymbol{\sigma}_1^\top \end{array}\right) \left(\begin{array}{c|c} 1 & \sigma_{11}^{-1}\boldsymbol{\sigma}_1^\top \\ \hline \boldsymbol{0}_{d-1} & \boldsymbol{I}_{d-1} \end{array}\right)$$

と書ける．ただし $\boldsymbol{I}_{d-1}$ は $(d-1)$ 次の単位行列で, $\boldsymbol{0}_{d-1}$ は $(d-1) \times 1$ の零ベクトルである．$\boldsymbol{\Sigma}$ が正定値ならば, $\boldsymbol{\Sigma}_2 - \sigma_{11}^{-1}\boldsymbol{\sigma}_1\boldsymbol{\sigma}_1^\top$ も正定値であるので

$$\begin{aligned}\boldsymbol{\Sigma}^{-1} &= \left(\begin{array}{c|c} 1 & -\sigma_{11}^{-1}\boldsymbol{\sigma}_1^\top \\ \hline \boldsymbol{0}_{d-1} & \boldsymbol{I}_{d-1} \end{array}\right) \left(\begin{array}{c|c} \sigma_{11}^{-1} & \boldsymbol{0}_{d-1}^\top \\ \hline \boldsymbol{0}_{d-1} & \left\{\boldsymbol{\Sigma}_2 - \sigma_{11}^{-1}\boldsymbol{\sigma}_1\boldsymbol{\sigma}_1^\top\right\}^{-1} \end{array}\right) \\ &\quad \times \left(\begin{array}{c|c} 1 & \boldsymbol{0}_{d-1}^\top \\ \hline -\sigma_{11}^{-1}\boldsymbol{\sigma}_1 & \boldsymbol{I}_{d-1} \end{array}\right) \end{aligned} \tag{1.13}$$

となる．$\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_d)^\top$ と $\boldsymbol{x} = (x_1, x_2, \ldots, x_d)^\top$ と書いたときに, $\boldsymbol{\mu}_1 = (\mu_2, \mu_3, \ldots, \mu_d)^\top$ と $\boldsymbol{x}_1 = (x_2, x_3, \ldots, x_d)^\top$ とおく．(1.13) から

$$\begin{aligned}\operatorname{tr}(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) &= \frac{1}{\sigma_{11}}(x_1 - \mu_1)^2 \\ &\quad + \operatorname{tr}\left[\left(\boldsymbol{x}_1 - \boldsymbol{\mu}_1 - \frac{x_1 - \mu_1}{\sigma_{11}}\boldsymbol{\sigma}_1\right)^\top \left\{\boldsymbol{\Sigma}_2 - \sigma_{11}^{-1}\boldsymbol{\sigma}_1\boldsymbol{\sigma}_1^\top\right\}^{-1} \right. \qquad (1.14) \\ &\qquad \left. \times \left(\boldsymbol{x}_1 - \boldsymbol{\mu}_1 - \frac{x_1 - \mu_1}{\sigma_{11}}\boldsymbol{\sigma}_1\right)\right] \\ &= \frac{1}{\sigma_{11}}(x_1 - \mu_1)^2 + g_{d-1}(\boldsymbol{x}_1) \qquad (1.15)\end{aligned}$$

となる．このことから

$$\int_{\mathbb{R}^{d-1}} \exp\left[-\frac{1}{2}\mathrm{tr}\,(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right] dx_2\, dx_3 \,\times\, dx_d$$
$$= \exp\left[-\frac{(x_1-\mu_1)^2}{2\sigma_{11}}\right]\int_{\mathbb{R}^{d-1}} \exp\left[-\frac{1}{2}g_{d-1}\right] dx_2\, dx_3 \,\times\, dx_d$$
$$= \exp\left[-\frac{(x_1-\mu_1)^2}{2\sigma_{11}}\right]\mathrm{Det}\left(2\pi\big(\boldsymbol{\Sigma}_2 - \sigma_{11}^{-1}\boldsymbol{\sigma}_1\boldsymbol{\sigma}_1^\top\big)\right)^{1/2}$$

がわかる．よって

$$\int_{\mathbb{R}^{d-1}} \mathsf{p}_{\boldsymbol{X}}\big(\boldsymbol{x}|\,\boldsymbol{\mu},\,\boldsymbol{\Sigma}\big)\, dx_2\, dx_3,\, \cdots dx_d = \frac{1}{\sqrt{2\pi\sigma_{11}}}\exp\left[\frac{(x_1-\mu_1)^2}{2\sigma_{11}}\right]$$

がわかる．

(2) (1.15) からわかる．

(3) $\boldsymbol{c}^\top\boldsymbol{X}$ の積率を計算すればよい．

(4) (3) から

$$\boldsymbol{\Sigma}^{-1/2}(\boldsymbol{X}-\boldsymbol{\mu}) \,\sim\, \mathsf{N}_d(\mathbf{0}_d,\, \boldsymbol{I}_d)$$

となる．あとは χ^2 分布の定義からわかる．□

(1.12) を用いた例題を述べておく

例 1.62. 連続型確率ベクトル $(X_1,\,X_2)$ は同時 p.d.f.

$$\mathsf{p}(x_1,\,x_2) = \begin{cases} 1 & (0\le x_1\le 1,\, 0\le x_2\le 1), \\ 0 & (\text{その他の場合}) \end{cases}$$

を持つとする．

$$y_1 = h_1(x_1,\,x_2) = x_1 + x_2,\, y_2 = h_2(x_1,\,x_2) = x_2$$
$$\Longleftrightarrow x_1 = h_1^{-1}(y_1,\,y_2),\, x_2 = h_2^{-1}(y_1,\,y_2)$$

とし，$\boldsymbol{Y} = (h_1(X_1,\,X_2),\, h_2(X_1,\,X_2))^\top$, $\boldsymbol{y} = (y_1,\,y_2)^\top$ とする．このとき

$$J_{\boldsymbol{h}^{-1}}(\boldsymbol{y}) = \mathrm{Det}\begin{pmatrix} \frac{\partial h_1^{-1}(\boldsymbol{y})}{\partial y_1} & \frac{\partial h_1^{-1}(\boldsymbol{y})}{\partial y_2} \\ \frac{\partial h_2^{-1}(\boldsymbol{y})}{\partial y_1} & \frac{\partial h_2^{-1}(\boldsymbol{y})}{\partial y_2} \end{pmatrix} = \mathrm{Det}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = 1$$

より

$$\mathsf{p}_{\boldsymbol{Y}}(\boldsymbol{y}) = \begin{cases} 1 & (0\le y_1-y_2\le 1,\, 0\le y_2\le 1), \\ 0 & (\text{その他の場合}) \end{cases}$$
$$= \begin{cases} 1 & (y_2\le y_1\le 1+y_2,\, 0\le y_2\le 1), \\ 0 & (\text{その他の場合}) \end{cases}$$

を得る.

$$\int\int \mathsf{p}_{\boldsymbol{Y}}(\boldsymbol{y})\,dy_1dy_2 = 1$$

となっていることに注意せよ. □

1.7　正規分布に関連した分布

この節では, 正規分布に関連した分布であるガンマ分布, χ^2 分布, F 分布, t 分布を定義し, これらの基本的な性質を述べる. これらの分布は正規母集団から標本に基づく統計量の分布 (いわゆる標本分布) の議論で重要な役割を担う.

まず, ガンマ関数を導入する. $\alpha > 0$ に対してガンマ関数を

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx$$

で定義する. $\alpha > 1$ のとき, 部分積分により

$$\begin{aligned}\Gamma(\alpha) &= \int_0^\infty x^{\alpha-1}e^{-x}\,dx = -x^{\alpha-1}e^{-x}\Big|_0^\infty + (\alpha-1)\int_0^\infty x^{\alpha-2}e^{-x}\,dx \\ &= \Gamma(\alpha-1).\end{aligned}$$

$\Gamma(1) = \int_0^\infty e^{-x}\,dx = 1$ なので, 帰納法により

$$\Gamma(n) = (n-1)! \qquad (n = 1, 2, \ldots)$$

となる.

1.7.1　ガンマ分布

定義 1.63. (ガンマ分布) $\alpha, \beta > 0$ とする. 連続型確率変数 X が母数 (α, β) のガンマ分布に従うとは, X の p.d.f. が

$$\mathsf{p}(x|\,\alpha, \beta) = \begin{cases} \dfrac{\beta^\alpha}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x} & (x > 0) \\ 0 & (\text{その他の場合}) \end{cases}$$

で与えられるときをいう. これを $X \sim \mathsf{Ga}(\alpha, \beta)$ と記す.

補題 1.64. $n \geq 2$ は自然数とする. $X_1, X_2, \ldots, X_n$ は独立な確率変数, 各 $X_j \sim \mathsf{Ga}(\alpha_j, \beta)\,(j = 1, 2, \ldots, n,\ \alpha_j > 0,\ \beta > 0$ としたとき

$$X_1 + X_2 + \cdots + X_n \sim \mathsf{Ga}(\alpha_1 + \alpha_2 + \cdots + \alpha_n, \beta).$$

Proof. 証明は次の補題を用いてする. □

補題 1.65. 連続型確率変数 X_1, X_2 は独立とし, それぞれの p.d.f. を $\mathsf{f}(x), \mathsf{g}(x)$ とする. このとき $X_1 + X_2$ の p.d.f. は

$$\mathsf{p}_{X_1+X_2}(x) = \int_{-\infty}^{\infty} \mathsf{f}(x-y)\mathsf{g}(y)\,dy$$

で与えられる.

Proof. X_1, X_2 の同時 p.d.f. は $\mathsf{f}(x)\mathsf{g}(y)$ になることに注意する. $t \in \mathbb{R}$ に対して

$$\begin{aligned}
\mathsf{Pr}\bigl(X_1 + X_2 \le t\bigr) &= \int\int_{x+y\le t} \mathsf{f}(x)\mathsf{g}(y)\,dx\,dy \\
&= \int_{-\infty}^{\infty} \mathsf{g}(y)\left\{\int_{-\infty}^{t-y} \mathsf{f}(x)\,dx\right\}dy \\
&= \int_{-\infty}^{\infty} \mathsf{g}(y)\left\{\int_{-\infty}^{t} \mathsf{f}(z-y)\,dz\right\}dy \qquad \left(x\,\Big|_{-\infty}^{t-y} = z - y\,\Big|_{-\infty}^{t} \text{ と変換}\right) \\
&= \int_{-\infty}^{t}\left\{\int_{-\infty}^{\infty} \mathsf{f}(z-y)\mathsf{g}(y)\,dy\right\}dz.
\end{aligned}$$

よって微積分の基本定理より

$$\mathsf{p}_{X_1+X_2}(t) = \frac{d}{dt}\mathsf{Pr}\bigl(X_1 + X_2 \le t\bigr) = \int_{-\infty}^{\infty} \mathsf{f}(t-y)\mathsf{g}(y)\,dy$$

を得る. □

補題 1.64 の証明. $X_1 \sim \mathsf{Ga}(\alpha_1, \beta)$, $X_2 \sim \mathsf{Ga}(\alpha_2, \beta)$ とし, X_1 と X_2 は独立とする. $\mathsf{Ga}(\alpha, \beta)$ の p.d.f. を $\mathsf{p}(\cdot\,|\,\alpha, \beta)$ と書く. 補題 1.65 より

$$\begin{aligned}
\mathsf{p}_{X_1+X_2}(x) &= \int_{-\infty}^{\infty} p(x-y|\,\alpha_1, \beta)p(y|\,\alpha_2, \beta)\,dy \\
&= \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} e^{-\beta x}\int_0^x (x-y)^{\alpha_1-1}y^{\alpha_2-1}\,dy \\
&\quad \bigl(x - y > 0 \text{ かつ } y > 0 \text{ より } 0 < y < x \text{ となる.}\bigr) \\
&= \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1)\Gamma(\alpha_2)} x^{\alpha_1+\alpha_2-1}e^{-\beta x}\int_0^1 (1-z)^{\alpha_1-1}z^{\alpha_2-1}\,dz \qquad (y = xz \text{ と変換})
\end{aligned}$$

を得る．ここで $\mathsf{p}_{X_1+X_2}$ は p.d.f. であることに注意すれば

$$
\begin{aligned}
1 &= \int_0^\infty \mathsf{p}_{X_1+X_2}(x)\,dx \\
&= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\left(\int_0^1 (1-z)^{\alpha_1-1} z^{\alpha_2-1}\,dz\right)\beta^{\alpha_1+\alpha_2}\int_0^\infty x^{\alpha_1+\alpha_2-1}e^{-\beta x}\,dx \\
&= \frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\left(\int_0^1 (1-z)^{\alpha_1-1} z^{\alpha_2-1}\,dz\right)\int_0^\infty w^{\alpha_1+\alpha_2-1}e^{-w}\,dw \\
&\quad (w=\beta z \text{ と変換}) \\
&= \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^1 (1-z)^{\alpha_1-1} z^{\alpha_2-1}\,dz \qquad (\text{ガンマ関数の定義より})
\end{aligned}
$$

より

$$
\int_0^1 (1-z)^{\alpha_1-1} z^{\alpha_2-1}\,dz = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)}
$$

を得る．よって

$$
X_1+X_2 \sim \mathsf{Ga}(\alpha_1+\alpha_2,\,\beta)
$$

がわかる．あとはこのことを繰り返せばよい．

1.7.2　χ^2 分布

定義 1.66. $Z_1, Z_2, \ldots, Z_n \sim$ i.i.d. $\mathsf{N}(0,1)$ とする．

$$
S = Z_1^2 + Z_2^2 + \cdots + Z_n^2
$$

の分布を**自由度 n の χ^2 分布**といい，$S \sim \chi_n^2$ と記す．

$Z \sim \mathsf{N}(0,1)$ のとき $Y = Z^2$ の p.d.f. は

$$
\mathsf{p}_Y(y) = \begin{cases} \dfrac{1}{\sqrt{2\pi}} y^{(1/2)-1} e^{-y/2} & (y>0) \\ 0 & (\text{その他の場合}) \end{cases} \tag{1.16}
$$

となる．演習問題 1.1 を参照せよ．一方，$\mathsf{Ga}(1/2,\,1/2)$ の p.d.f. は

$$
\mathsf{p}\left(y\,\middle|\,\frac{1}{2},\,\frac{1}{2}\right) = \begin{cases} \dfrac{1}{2\Gamma\left(\frac{1}{2}\right)} y^{(1/2)-1} e^{-y/2} & (y>0) \\ 0 & (\text{その他の場合}) \end{cases}
$$

であった．これらはともに p.d.f. なので，積分をすれば 1 となる．このことより

$$
\sqrt{2}\Gamma\left(\frac{1}{2}\right) = \sqrt{2\pi} \Longleftrightarrow \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}
$$

を得る. さらに補題 1.64 より

$$\chi_n^2 = \mathsf{Ga}\left(\frac{n}{2}, \frac{1}{2}\right)$$

がわかる. よって χ_n^2 の p.d.f. は

$$\mathsf{p}(x|\, n) = \begin{cases} \dfrac{\left(\dfrac{1}{2}\right)^n}{\Gamma\left(\dfrac{n}{2}\right)} x^{(n/2)-1} e^{-x/2} & (x > 0) \\ 0 & (\text{その他の場合}) \end{cases}$$

である. 演習問題 2.1 より $X \sim \chi_n^2$ のとき

$$\mathsf{E}[X] = n, \quad \mathsf{E}[X] = 2n$$

となるがわかる.

1.7.3 F 分布

定義 1.67. $k, m \in \mathbb{N}$ とする. X と Y は独立で $X \sim \chi_k^2$ と $Y \sim \chi_m^2$ とする. このとき

$$\mathsf{F} = \frac{X/k}{Y/m}$$

の分布を**自由度** (k, m) **の** F **分布**といい. $\mathsf{F} \sim \mathsf{F}(k, m)$ と記す.

補題 1.68. $k, m \in \mathbb{N}$ とする. $F \sim \mathsf{F}(k, m)$ のとき, F の p.d.f. $\mathsf{p}(x|\, k, m)$ は

$$\mathsf{p}(x|\, k, m) = \begin{cases} \dfrac{\Gamma\left(\dfrac{k+m}{2}\right)}{\Gamma\left(\dfrac{k}{2}\right)\Gamma\left(\dfrac{m}{2}\right)} \dfrac{k^{k/2} m^{m/2} x^{(k/2)-1}}{(m+kx)^{(k+m)/2}} & (x > 0) \\ 0 & (\text{その他の場合}). \end{cases}$$

で与えられる.

Proof. 証明は次の補題を用いて行う. □

補題 1.69. X, Y を正値の連続型確率変数とし, それぞれの p.d.f. を $\mathsf{f}(x)$ と $\mathsf{g}(x)$ とする. このとき

$$Z = \frac{X}{Y}$$

の p.d.f. は

$$\mathsf{h}(z) = \begin{cases} \int_0^\infty \mathsf{f}(zy)\mathsf{g}(y)y\,dy & (z > 0) \\ 0 & (\text{その他の場合}) \end{cases}$$

で与えられる.

Proof. 仮定から (X, Y) の同時 p.d.f. は $\mathsf{f}(x)\mathsf{g}(y)$ となる．よって $t > 0$ に対して

$$\begin{aligned} \mathsf{Pr}\big(Z \le t\big) &= \int_0^\infty \left\{ \int_0^{ty} \mathsf{f}(x)\mathsf{g}(y)\,dx \right\} dy \\ &= \int_0^\infty \left\{ \int_0^{t} \mathsf{f}(zy)\mathsf{g}(y)y\,dz \right\} dy \quad (x = zy \text{ と変換}) \\ &= \int_0^t \left\{ \int_0^{\infty} \mathsf{f}(zy)\mathsf{g}(y)y\,dy \right\} dz. \end{aligned}$$

ここで微積分の基本定理を用いると

$$\mathsf{h}(t) = \frac{d}{dt}\mathsf{Pr}\big(Z \le t\big) = \int_0^\infty \mathsf{f}(zy)\mathsf{g}(y)y\,dy.$$

□

補題 1.68 の証明　補題 1.69 を用いる．$w > 0$ のとき, $W = X/Y$ の p.d.f. は

$$\begin{aligned} \mathsf{p}_W(w) &= \int \mathsf{f}(wy)\mathsf{g}(y)y\,dy \\ &= \frac{\left(\frac{1}{2}\right)^{k/2}\left(\frac{1}{2}\right)^{m/2}}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{m}{2}\right)} \int_0^\infty (wy)^{(k/2)-1}e^{-(wy)/2}y^{(m/2)-1}e^{-y/2}y\,dy \\ &= \frac{\left(\frac{1}{2}\right)^{(k+m)/2}}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{m}{2}\right)} w^{(k/2)-1} \int_0^\infty y^{(k+m)/2-1}e^{-(w+1)y/2}\,dy \\ &= \frac{\left(\frac{1}{2}\right)^{(k+m)/2}}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{m}{2}\right)} w^{(k/2)-1} \int_0^\infty \left(\frac{2x}{w+1}\right)^{(k+m)/2-1} e^{-x}\frac{2}{w+1}\,dx \\ &\qquad \left(x = \tfrac{w+1}{2}y\right) \text{ と変換} \\ &= \frac{1}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{m}{2}\right)} w^{(k/2)-1}(1+w)^{-(k+m)/2} \int_0^\infty x^{(k+m)/2-1}e^{-x}\,dx \\ &= \frac{\Gamma\left(\frac{k+m}{2}\right)}{\Gamma\left(\frac{k}{2}\right)\Gamma\left(\frac{m}{2}\right)} w^{(k/2)-1}(1+w)^{-(k+m)/2}. \end{aligned}$$

よって

$$\mathsf{p}(w|\,k,\,m) = \mathsf{p}_W\left(\frac{k}{m}z\right)\frac{k}{m}$$

よりわかる.

1.7.4　t 分布

定義 1.70. $Z_0, Z_1, Z_2, \ldots, Z_n \sim$ i.i.d. $\mathsf{N}(0, 1)$ のとき

$$T = \frac{Z_0}{\sqrt{\frac{1}{n}\bigl(Z_1^2 + Z_2^2 + \cdots + Z_n^2\bigr)}}$$

を自由度 n の t 分布といい，これを $T \sim t_n$ と記す.

補題 1.71. T の p.d.f. は

$$\mathsf{p}_T(x|\,n) = \begin{cases} \dfrac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n}{2}\right)}\frac{1}{\sqrt{n}}\left(1+\frac{x^2}{n}\right)^{-(n+1)/2} & (x > 0) \\ 0 & (\text{その他の場合}) \end{cases}$$

で与えられる

Proof. 定義より

$$T^2 \sim \mathsf{F}(1, n)$$

であることをまず思い出す. T の分布は対称なので

$$\mathsf{p}_T(x|\,n) = \mathsf{p}_T(-x|\,n) \qquad (x \geq 0)$$

となる. よって $x > 0$ に対して

$$2\mathsf{Pr}\bigl(0 \leq T \leq x\bigr) = \mathsf{Pr}\bigl(-x \leq T \leq x\bigr) = \mathsf{Pr}\bigl(T^2 \leq x^2\bigr).$$

上の式から $\mathsf{F}(1, n)$ の p.d.f. を $\mathsf{p}_{1,n}$ と書けば

$$2\int_0^x \mathsf{p}_T(t|\,n)\,dt = \int_0^{x^2} \mathsf{p}_{1,n}(t)\,dt.$$

$x = y^2$ と変換すると

$$\int_0^{x^2} \mathsf{p}_{1,n}(t)\,dt = \int_0^x \mathsf{p}_{1,n}(t^2)2t\,dt. \tag{1.17}$$

となる. ここで補題 1.69 から

$$\mathsf{p}_{1,n}(x) = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{1}{2}\right)\Gamma\left(\dfrac{n}{2}\right)}\frac{n^{n/2}x^{-1/2}}{(n+x)^{(n+1)/2}}$$

であることを思い出す. この式を (1.17) の左辺に代入すると

$$\mathsf{p}_T(x|\,n) = \mathsf{p}_{1,n}(x^2)x = \frac{\Gamma\left(\dfrac{n+1}{2}\right)}{\Gamma\left(\dfrac{1}{2}\right)\Gamma\left(\dfrac{n}{2}\right)}\frac{1}{\sqrt{n}}\left(1+\frac{x^2}{n}\right)^{-(n+1)/2}$$

となる. □

1.8　章末注釈と参考文献

1.9　演習問題

演習問題 1.1. $Z \sim \mathsf{N}(0,1)$ とする. Z の p.d.f. を

$$\mathsf{p}(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \qquad (-\infty < z < \infty)$$

としたとき, 以下の問いに答えよ.

(1) 任意の $y > 0$ に対して, $\mathsf{Pr}(Z^2 \leq y)$ を p も用いて表現せよ.

(2) $Y := Z^2$ とし, Y の p.d.f. を p_Y と書く. このとき, $y > 0$ に対して

$$\mathsf{p}_Y(y) = \frac{d}{dy}\mathsf{Pr}(Y \leq y)$$

となることを利用して, Y の p.d.f. が (1.16) で与えられることを示せ.

演習問題 1.2. X を確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の連続型確率変数とし, X の p.d.f. を

$$\mathsf{p}(x) = \begin{cases} 2x & (0 < x < 1) \\ 0 & (\text{その他の場合}) \end{cases}$$

とする. さらに, 離散型確率変数 Y を

$$Y := \mathbb{1}_{(1/2,\,\infty)}(X) := \begin{cases} 1 & \left(X > \frac{1}{2}\right) \\ 0 & \left(X \leq \frac{1}{2}\right) \end{cases}$$

で定める. このとき, 以下の問に答えよ.

(1) 確率変数 X の c.d.f. を求めよ. なお, X の c.d.f. を F と書くことにする.

(2) 確率変数 Y の p.m.f. を求めよ. なお, X の p.m.f. を p_Y と書くことにする.

(3) 確率 $\mathsf{Pr}\left(0 < X \leq \frac{1}{2}\right)$, $\mathsf{Pr}(Y = 0)$ を求めよ.

(4) 確率 $\mathsf{Pr}\left(0 < X \leq \frac{1}{2}, Y = 0\right)$ を求めよ.

(5) 確率変数 X と Y は独立か従属かを判定せよ.

演習問題 1.3. 大小 2 つのサイコロを投げたとき, それぞれの出る目を X と Y とおく. これらを確率空間 $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上で定義されている確率変数と考える. この空間上の確率変数 Z と W を

$$Z(\omega) = \min\{X(\omega), Y(\omega)\}, \quad W(\omega) = \max\{X(\omega), Y(\omega)\} \quad (\forall \omega \in \Omega)$$

で定める. このとき, 以下の問いに答えよ.
(1) 確率ベクトル (Z, W) の同時 p.m.f. $\mathsf{p}_{(Z,W)}$ を求めよ.
(2) 確率 $\mathsf{Pr}(Z = W)$ と条件付き確率 $\mathsf{Pr}(Z = 2 | W < 5)$ を求めよ.
(3) 確率変数 Z の周辺 p.m.f. p_Z を求めよ.
(4) 確率変数 W の周辺 p.m.f. p_W を求めよ.
(5) 確率変数 Z と W は独立であるかどうかを判定せよ.