確率論基礎 2022 まとめ

安田

概要

確率論基礎がそこそこ難解なため, PDF としてまとめました. スライドを読んで理解しながら書きますので, 間違った表現があったら申し訳ないです. 基本的には, スライドと同じことを書いて, 詳しい説明が欲しいと感じたところには詳しい説明を入れるという方式をとっていますが, 結構主観も入ってます. あと, 多少書きやすいように時系列をいじるかもしれません. そこそこ分量が多いので, 目次や最後のページにある索引を参照すると良いと思います.

目次

1 Introduction

1.1 標本空間

定義 1.1.1 (標本空間)**.** 確率的事象の起こり得る個々の (単一の) 結果のことを**標本**, 標本全体の集合を**標本空間**という. □

スライドには載っていなかったが, 定義のない言葉を使いたくないので, 事象についても定義をしておく.

定義 1.1.2 (事象)**.** 標本空間の部分集合のうち, その確率を計算できるものを**事象**という. □

曖昧な定義に見えるが, 計算可能な部分集合を決めてからそれを事象と呼ぶため, おそらく問題ないと思われる.

1.2 組合せ確率

組合せ確率は, 高校数学でやるような確率のことである.

定義 1.2.1 (組合せ確率)**.** 有限な標本空間 Ω に対して, 任意の部分集合 A を事象とし, その**組合せ確率** $P(A)$ を

$$P(A) := \frac{|A|}{|\Omega|}$$

と定義する. □

このとき, 各標本 $\omega \in \Omega$ に対して, $P(\omega)$ は同じ値を取る. (あらゆる標本の確率は等しい.)

1.3 離散確率

有限な標本空間で, 全ての標本の確率が等しいような状況は具体的すぎて, 流石に全てのケースに対応するのは無理がある. そこで, 可算まで増やし, 確率にもそこそこ自由度を持たせたものを考える.

定義 1.3.1 (離散確率)**.** 高々可算な標本空間 Ω に対して, **確率質量関数** $p : \Omega \to [0, 1]$ を次を満たすように定める.

$$\sum_{\omega \in \Omega} p(\omega) = 1$$

ここで, 任意の Ω の部分集合 A を事象とし, その確率を

$$P(A) := \sum_{\omega \in A} p(\omega)$$

と定義する. このとき, この組 (Ω, p) を**離散確率空間**と呼ぶ. □

スライドには $p(\omega) \geq 0$ が条件とあるが, $p : \Omega \to [0, 1]$ の定義から既に満たされている.

筆者は理解力が乏しいため, この定義において "A が無限集合だったら計算不可能では" と考えてしまった (当然そんなことはない). 一応, 同じ勘違いをしている方向けに説明を書いておこうと思う. p が具体的に定められていて, Ω は離散的な集合であるから, 極限的な操作が入らない. したがって, p が限りなく小さいような

ω の無限集合の確率というのは考える必要がない. p が限りなく小さくなることはあり得ない. A に順序を持たせれば, 通常の無限級数のように計算できる.

例 1.3.2. コインを表が出るまで表を投げることを考える. 標本空間を Ω とする. ω_i を i 回目に初めて表が出ることを表す標本, また ω_∞ を永遠に表が出ないことを表す標本とする. ここで, $p : \Omega \to [0,1]$ を $p(\omega_i) = \frac{1}{2^i}, p(\omega_\infty) = 0$ とする。

$$\sum_{p \in \Omega} p(\omega) = \sum_{i=1}^{\infty} \frac{1}{2^i} = \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{2}} = 1$$

より (Ω, p) は離散確率空間である. 例えばこのとき, 偶数回目に初めて表が出る事象 A の確率は

$$P(A) = \sum_{i=1}^{\infty} p(\omega_{2i}) = \sum_{i=1}^{\infty} \frac{1}{4^i} = \frac{1}{4} \cdot \frac{1}{1 - \frac{1}{4}} = \frac{1}{3}$$

となる. □

例 1.3.3. また, よく出現する確率として, 次のようなものがある.

- 標本空間 $\Omega = \{1, 2, \cdots, n\}$ に対して, $p_{\mathrm{du}}(i) = \frac{1}{n}$ とすると, $(\Omega, p_{\mathrm{du}})$ は離散確率空間であり, **離散一様分布**と言われる.
- $\alpha \in (0,1)$ とする. 標本空間 $\Omega = \{0, 1, \cdots, n\}$ に対して, $p_{\mathrm{bin}} = \binom{n}{i} \alpha^i (1-\alpha)^{n-i}$ とすると, $(\Omega, p_{\mathrm{bin}})$ は離散確率空間であり, **二項分布**と言われる.
- $\alpha \in (0,1)$ とする. 標本空間 $\Omega = \{1, 2, \cdots\}$ に対して, $p_{\mathrm{geo}}(i) = (1-\alpha)^{i-1}\alpha$ とすると, $(\Omega, p_{\mathrm{geo}})$ は離散確率空間であり, **幾何分布**と言われる.
- $\lambda > 0$ とする. 標本空間 $\Omega = \{0, 1, \cdots\}$ に対して, $p_{\mathrm{poi}}(i) = \frac{\lambda^i}{i!} e^{-\lambda}$ とすると, $(\Omega, p_{\mathrm{poi}})$ は離散確率空間であり, **ポアソン分布**と言われる.

なお, $\alpha n = \lambda$ の関係式のもと $n \to \infty$ とすると, 二項分布はポアソン分布へ収束する.

$$p_{\mathrm{bin}}(i) = \binom{n}{i} \left(\frac{\lambda}{n}\right)^i \left(1 - \frac{\lambda}{n}\right)^{n-i} = \frac{\lambda^i}{i!} \frac{n \cdot (n-1) \cdots (n-i+1)}{n \cdot n \cdots n} \left(1 - \frac{\lambda}{n}\right)^{\lambda \cdot \frac{n}{\lambda}} \left(1 - \frac{\lambda}{n}\right)^{-i}$$

ここで, $\dfrac{n-k}{n} \xrightarrow{n \to \infty} 1, \left(1 - \dfrac{\lambda}{n}\right)^{\lambda \cdot \frac{n}{\lambda}} \xrightarrow{n \to \infty} e^{-\lambda}, (1 - \dfrac{\lambda}{n})^{-i} \xrightarrow{n \to \infty} 1$ より, 主張が従う. □

1.4 離散確率の限界

上のように, 離散確率空間は様々な確率へ対応する. それでもまだ, これでもうまく定義できないような確率というものが存在する. 例えば, $[0,1]$ からランダムに 1 点を選ぶという操作において, 離散確率空間の定義では, 確率質量関数を自由に定義できない. 非可算無限個の要素に値をふるような確率質量関数を定義した時, 標本の数が大きすぎて総和が無限大となってしまうのである. しかし, なんとなくなイメージしか沸かず, 実際の証明が気になるということが筆者にはあった. 同じような人のためにしっかりとした証明をしておく. ここで, 非可算無限個の値の総和は取れないため, 可算無限個の部分集合を用意するだけで無限大へ行ってしまうことを証明する. このとき, ましてや Ω 全体での総和など取れないということである.

命題 1.4.1. Ω を非可算無限集合とする．$p : \Omega \to [0,1]$ を集合 $\{\omega \in \Omega \mid p(\omega) \neq 0\}$ の濃度が非可算無限となるように定めるとき，ある可算無限部分集合 $A := \{a_1, a_2, \cdots\}$ で，$\displaystyle\sum_{i=1}^{\infty} p(a_n) = \infty$ となる． □

証明．ある $\varepsilon > 0$ が存在し，$\{\omega \in \Omega \mid p(\omega) \geq \varepsilon\}$ の濃度が無限であることを示せば，これが命題の A にあたる．そこで，そのような ε が存在しないと仮定する．$p(\omega) \neq 0$ となる ω が無限個あるから，ある 0 に収束する無限列 $(\varepsilon_i)_{i\in\mathbb{N}}$ が取れて，$A_i := \{\omega \in \Omega \mid p(\omega) \geq \varepsilon_i\}$ として

$$i > j \Rightarrow |A_i \setminus A_j| > 0$$

となる．また，仮定より $|A_i| < \aleph_0$ だから，$|A_i \setminus A_j| < \aleph_0$ である．このとき $A_0 = \emptyset$ として

$$\{\omega \in \Omega \mid p(\omega) \neq 0\} = \bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} (A_i \setminus A_{i-1})$$

であるが，右辺は有限集合の可算無限和だから可算無限集合となる．しかし，左辺は非可算無限となり，矛盾である．したがって，最初に述べたような ε は存在する．

ただし，有限集合の可算無限和が可算無限となる証明は「有限集合 可算無限和」で調べると出てくるので，そこで見て欲しい． ■

このように証明できた．筆者は，この証明をする前に，なぜ可算無限だとうまくいくのかがぼんやりとしか理解ができていなかったが，上の証明を用いれば簡単にわかる．$\varepsilon > 0$ が取れなくても $\{\omega \in \Omega \mid p(\omega) \neq 0\}$ は可算無限までは到達できるということである．

2 可測空間

先ほど述べたように，離散確率空間には限界がある．そこで，もっと一般的な確率を定義するための準備として，集合に対して可測集合というものを定義していろいろと考える．

2.1 集合演算

和や積や差または補集合のような一般的な集合論の話は省く．排反は定義しておく．

定義 2.1.1. 集合 A, B について，$A \cap B = \emptyset$ となるとき A と B は**互いに排反**であるという．また，A の部分集合 $A_1, \cdots, A_n$ がそれぞれ互いに排反で $\displaystyle\bigcup_{i=1}^{n} A_i = A$ となるとき，$A_1, \cdots, A_n$ は A の**分割**であるという． □

無限個のド・モルガンの法則の証明だけは，一応ぱっと見れたら筆者的には嬉しいので載せておく．(スライドにも載っている.)

補題 2.1.2 (ド・モルガンの法則)**.**

$$\left(\bigcup_{i=1}^{\infty} A_i\right)^c = \bigcap_{i=1}^{\infty} {A_i}^c, \quad \left(\bigcap_{i=1}^{\infty} A_i\right)^c = \bigcup_{i=1}^{\infty} {A_i}^c$$

□

証明.

$$\omega \in \left(\bigcup_{i=1}^{\infty} A_i\right)^c \iff \omega \notin \bigcup_{i=1}^{\infty} A_i \iff \forall i \geq 1, \omega \notin A_i$$
$$\iff \forall i \geq 1, \omega \in {A_i}^c \iff \omega \in \bigcap_{i=1}^{\infty} {A_i}^c$$

$$\omega \in \left(\bigcap_{i=1}^{\infty} A_i\right)^c \iff \omega \notin \bigcap_{i=1}^{\infty} A_i \iff \exists i \geq 1, \omega \notin A_i$$
$$\iff \exists i \geq 1, \omega \in {A_i}^c \iff \omega \in \bigcup_{i=1}^{\infty} {A_i}^c$$

■

2.2 σ-集合体と可測空間

ここからが測度論の本題である.

定義 2.2.1 (可測空間)**.** 集合 Ω の部分集合族 $\mathcal{F}$ が Ω 上の σ-集合体であるとは, 次のことを満たすことをいう.

(1) $\Omega \in \mathcal{F}$
(2) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
(3) $A_1, A_2, \cdots \in \mathcal{F} \Rightarrow \displaystyle\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

このとき, 組 $(\Omega, \mathcal{F})$ を**可測空間**という. また, $\mathcal{F}$ の元を**可測集合**ともいう. □

σ-集合体の例として, 次のようなものがある.

例 2.2.2.

- $\mathcal{F}_0 = \{0, \Omega\}$
- $\mathcal{F} = 2^{\Omega}$
- $\emptyset \neq A \subsetneq \Omega$ に対して, $\mathcal{F}_A = \{\emptyset, A, A^c, \Omega\}$

□

ある集合族を含むような σ-集合体を考えたい. このとき, 次のような便利な命題がある.

定理 2.2.3. Ω を集合とする. $\mathcal{V}$ を Ω の部分集合族とすると, $\mathcal{V} \subset \mathcal{F}$ を満たし, 順序関係 $\subset$ について最小の σ-集合体となる $\mathcal{F}$ が存在する. □

証明. $\mathcal{M}$ を, $\mathcal{V}$ を含む σ-集合体全体の集合とする. このとき, 2^{Ω} は $\mathcal{V}$ を含む σ-集合体であるから, $\mathcal{M} \neq \emptyset$ である. ここで, $\mathcal{F}$ を $\mathcal{M}$ の元全ての積集合 (共通部分) とすると, $\mathcal{V}$ を包含する任意の σ -集合体に含まれている. したがって, これが σ-集合体となることを示せばよい.

任意の $\mathcal{M}$ の元は Ω を元に持つから, $\Omega \in \mathcal{F}$ である. 同様の考え方で $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$, $\quad A_1, A_2, \cdots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ も示せる. ■

これにより, ある集合族から, それを包含する最小の σ-集合体へ写す写像を定義できる.

定義 2.2.4. $\mathcal{V}$ を Ω の部分集合族とする. このとき, $\mathcal{V}$ を含む最小の σ-集合体を $\sigma(\mathcal{V})$ で表す. これを $\mathcal{V}$ から生成される σ-集合体ともいう. □

また, σ-集合体の定義は最小限の条件のみで書かれているが, 実際はもっと色々な演算について閉じていることがわかる.

命題 2.2.5. (1) $\phi \in \mathcal{F}$

(2) $A, B \in \mathcal{F} \Rightarrow A \cup B, A \cap B, A \setminus B \in \mathcal{F}$

(3) $A_1, A_2, \cdots \in \mathcal{F} \Rightarrow \displaystyle\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$

□

証明. (1) は $\Omega^c = \emptyset$ により従う. (2) は A_3 以降を $\emptyset$ とした A_i の可算無限和を考えれば $A \cup B$, またここから $(A^c \cup B^c)^c$ を考えれば $A \cap B$, $A \setminus B = A \cap B^c$ を考えれば $A \setminus B$ の証明ができる. (3) についても, 可算無限和のド・モルガンの法則を用いて $\left(\displaystyle\bigcup_{i=1}^{\infty} {A_i}^c\right)^c$ を考えれば良い. ■

また, σ-集合体同士の積集合についても次の命題がある. スライドでは 2 つの積集合についての命題であったが, もっと一般的に成立する

命題 2.2.6. $(\mathcal{F}_i)_{i \in I}$ を Ω 上の σ-集合体の族とすると, その共通部分も Ω 上の σ-集合体である. □

この証明は, 2.2.3 の証明の後半部分であるから省く. また, σ-集合体の和集合は必ずしも σ-集合体とはならないので, 注意が必要である.

次に確率に組み込むためのちょっとした定義を書いておく. スライドでは, 可測空間と事象を同時に定義していたが, ここでは離して定義しておく.

定義 2.2.7. $(\Omega, \mathcal{F})$ を可測空間とする. Ω を確率空間と見るとき, 事象全体の集合を $\mathcal{F}$ で定める. □

2.3 ボレル集合体

ボレル集合体とは, σ-集合族の一種である. 全ての開集合の族を包含する最小の σ-集合族と定義される. 確率論基礎では, 実数の場合のみ扱うようである.

定義 2.3.1 (実数のボレル集合体)**.** 実数上の**ボレル集合体** $\mathcal{B}(\mathbb{R})$ とは, 全ての半開区間を含む最小の σ-集合体である. $\mathbb{R}^n$ 上では, 開集合を含むようなものと定める. また, ボレル集合体の元をボレル集合ともいう. □

なぜ半開区間で定義したのかよくわからない. それはともかく, ボレル集合体は, 直感的に思いつくような $\mathbb{R}$ の部分集合は基本的にボレル集合となる.

命題 2.3.2.

(1) $\{a\} \in \mathcal{B}(\mathbb{R})$

(2) $[a, b], (a, b), [a, b) \in \mathcal{B}(\mathbb{R})$

(3) $(a,\infty),[a,\infty),(-\infty,a],(-\infty,a)\in\mathcal{B}(\mathbb{R})$

□

証明. (1)$\{a\}=\bigcap_{n=1}^{\infty}(a-\frac{1}{n},a]$ より従う. (2) は (1) を用いて境界点を足したり引いたりすることで示せる. (3) は $(a,\infty)=\bigcup_{n=1}^{\infty}(a,a+n]$ より (a,∞) について示され, あとは (2) と同様に示せる. ■

2.4 事象列の極限

まずは単調列の極限についてである.

定義 2.4.1. $(\Omega,\mathcal{F})$ を可測空間, (A_n) をその事象列とする.

(1) $A_1\subset A_2\subset\cdots$ のとき, これを**単調増加列**と言い, A_n の極限を次のように定義する.

$$\lim_{n\to\infty}A_n=\bigcup_{i=1}^{\infty}A_i$$

(2) $A_1\supset A_2\supset\cdots$ のとき, これを**単調減少列**と言い, A_n の極限を次のように定義する.

$$\lim_{n\to\infty}A_n=\bigcap_{i=1}^{\infty}A_i$$

□

これらは, それぞれ単調増加列 A_n, 単調減少列 B_n に対して

$$A_n=\bigcup_{i=1}^{n}A_i\ ,\quad B_n=\bigcap_{i=1}^{n}B_i$$

となっていることから, 自然な定義だと分かる. ここで, 一般の事象の列にも極限を定義したい. その準備として上極限と下極限を定義する. まず, 事象列 A_n に対して

$$B_n=\bigcup_{i=n}^{\infty}A_i,C_n=\bigcap_{i=n}^{\infty}A_i$$

とすれば B_n,C_n はそれぞれ単調減少列, 単調増加列となることがすぐにわかる. これを用いる.

定義 2.4.2. $(\Omega,\mathcal{F})$ を可測空間, (A_i) をその事象列とする. このとき

$$\limsup_{n\to\infty}A_n:=\lim_{n\to\infty}\bigcup_{i=n}^{\infty}A_i=\bigcap_{n=1}^{\infty}\bigcup_{i=n}^{\infty}A_i$$

を $(A_n)_{n\in\mathbb{N}}$ の**上極限**,

$$\liminf_{n\to\infty}A_n:=\lim_{n\to\infty}\bigcap_{i=n}^{\infty}A_i=\bigcap_{n=1}^{\infty}\bigcup_{i=n}^{\infty}A_i$$

を $(A_n)_{n\in\mathbb{N}}$ の**下極限**という. □

n が十分大きい時, 上極限はどれだけ n が大きくなっても A_n は最大そこまで大きくなるという集合になるイメージで, 下極限は n が大きくなっても A_n は少なくともそれだけは包含するような最大の集合になるイメージである. また, これらの上極限, 下極限は次のように言い換えることができる.

命題 2.4.3. $(\Omega, \mathcal{F})$ を可測空間, (A_i) をその事象列とする. このとき次のことが成り立つ.

(1) $\omega \in \limsup_{n\to\infty} A_n \iff \omega \in A_i$ となる i が無限個

(2) $\omega \in \liminf_{n\to\infty} A_n \iff \exists N \in \mathbb{N}, \forall n \geq N, \omega \in A_n$

□

証明.

(1)

$$\omega \in \limsup_{n\to\infty} A_n \iff \omega \in \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} A_n \iff \forall n \geq 1, \omega \in \bigcup_{i=n}^{\infty} A_n$$

となる. 無限個の n で成立すれば当然上の命題が成立する. また, 有限個の n でこれが成立すると仮定すると, 明らかにその最大の n を N として $n = N+1$ とすれば矛盾する. したがって, 示された.

(2)

$$\omega \in \liminf_{n\to\infty} A_n \iff \bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_n \iff \exists N \in \mathbb{N}, \omega \in \bigcap_{i=n}^{\infty} \iff \exists N \in \mathbb{N}, \forall n \geq N, \omega \in$$

■

ここから, 明らかに上極限が下極限を包含していることがわかる. また, 次のように補集合についての等式が存在する.

命題 2.4.4. $(\Omega, \mathcal{F})$ を可測空間, (A_i) をその事象列とする. このとき次のことが成り立つ.

(1) $\limsup_{n\to\infty} {A_n}^c = \left(\liminf_{n\to\infty} A_n\right)^c$

(2) $\liminf_{n\to\infty} {A_n}^c = \left(\limsup_{n\to\infty} A_n\right)^c$

□

証明. (1) のみ証明すれば, (2) は A_n を ${A_n}^c$ に置き換えて, 全体の補集合を取ればよい. (1) はド・モルガンの法則から

$$\left(\liminf_{n\to\infty}\right)^c = \left(\bigcup_{n=1}^{\infty} \bigcap_{i=n}^{\infty} A_i\right)^c = \bigcap_{n=1}^{\infty} \left(\bigcap_{i=n}^{\infty}\right)^c = \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} {A_n}^c = \limsup_{n\to\infty} {A_n}^c$$

より従う. ■

上極限と下極限を用いて, ようやく事象列の極限というのが次のように定義される.

定義 2.4.5. $(\Omega, \mathcal{F})$ を可測空間, (A_i) をその事象列とする.

$$\limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n$$

のとき**極限が存在する**といい, A_n の極限を

$$\lim_{n\to\infty} A_n := \limsup_{n\to\infty} A_n = \liminf_{n\to\infty} A_n$$

で定める. □

3 確率空間

3.1 離散確率空間から一般的な確率空間へ

一般的な確率空間を定義するにあたって, 離散確率空間と似た性質を持つように定義したい. まずは離散確率空間の次の性質を振り返る.

補題 3.1.1. (Ω, p) を離散確率空間, P をその離散確率関数とする. このとき次が成り立つ.

(1) $\forall A \in 2^{\Omega}, P(A) \geq 0$
(2) $P(\Omega) = 1$
(3) (A_i) を 2^{Ω} の元の列とする. 各 i, j で A_i, A_j が排反のとき

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

である.

□

証明. (1), (2) はほぼ定義そのものである. (3) については,

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{\omega \in \bigcup_{i=1}^{\infty} A_i} p(\omega) = \sum_{i=1}^{\infty} \sum_{\omega \in A_i} p(\omega) = \sum_{i=1}^{\infty} P(A_i)$$

より従う. ■

これらの性質を参考に一般的な確率空間を定義する.

定義 3.1.2 (確率空間)**.** $(\Omega, \mathcal{F})$ を可測空間とする. ここで, 写像 $P : \mathcal{F} \to [0, 1]$ は次の条件を満たす時, P を**確率測度**という.

(1) $P(\Omega) = 1$
(2) (A_i) を $\mathcal{F}$ 上の列とする. 各 i, j で A_i, A_j を排反のとき,

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

である. これを σ**-加法性**という.

事象の集合は前も述べた通り $\mathcal{F}$ である. このとき, 組 $(\Omega, \mathcal{F}, P)$ を**確率空間**という. □

授業では $\forall A \in \mathcal{F}, P(A) \geq 0$ を定義に入れていたが, P の値域からすでに条件は満たされている. また, 離散確率空間は明らかにこの定義を満たしているから, 当然確率空間となる.

3.2 確率の性質

まずは, 離散確率空間と同様に次の性質が成り立つ.

命題 3.2.1. $(\Omega, \mathcal{F}, P)$ を確率空間とし, A, B をその事象とする. このとき次が成立する.

(1) $P(\phi) = 0$

(2) A, B が排反ならば $P(A \cup B) = P(A) + P(B)$ である. これを**加法性**という.

(3) $P(A^c) = 1 - P(A)$

□

証明は離散確率空間と同じように証明できるため省く. また, 次のような性質も成り立つ.

命題 3.2.2. $(\Omega, \mathcal{F}, P)$ を確率空間とし, A, B をその事象とする. このとき次が成立する.

(1) $A \subset B \Rightarrow P(A) \leq P(B)$ これを**単調性**という.

(2) $P(A \cup B) \leq P(A) + P(B)$ これを**劣加法性**という.

(3) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ これを**加法法則**という.

□

証明. (1) 加法性を用いれば, $P(B) = P(A \cup (B \setminus A)) = P(A) + P(B \setminus A) \geq P(A)$ である.

(2) (3) から従う.

(3) 加法性を用いる. $P(A \cup B) = P(A \cup (B \setminus A)) = P(A) + P(B \setminus A)$ であり, $P(B) = P((B \cap A) \cup (B \setminus A)) = P(B \cap A) + P(B \setminus A)$ であるから, これらを比較すればよい.

■

加法法則, 劣加法性をたくさん使うと, 次のような性質も導ける.

定理 3.2.3. $(\Omega, \mathcal{F}, P)$ を確率空間とし, $A_1, \cdots, A_n$ をその事象とする. このとき次のことが成立する.

(1) $P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$ これも劣加法性という.

(2) $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 \leq \cdots \leq i_k \leq n} P\left(\bigcap_{j=1}^{k} A_{i_j}\right)$ これを**包除原理**という.

□

証明.

(1) $n = 1$ では当然成立する. ある n で成立したとすると, 劣加法性より $n + 1$ でも成立する. したがって, 任意の n で成立する.

(2) $n=1$ では当然成立する．ある n で成立したとする．このとき

$$\begin{aligned}
P\left(\bigcup_{i=1}^{n+1} A_i\right) =& \ P\left(\bigcup_{i=1}^{n} A_i\right) + P(A_{n+1}) - P\left(\bigcup_{i=1}^{n}(A_i \cap A_{n+1})\right) \\
=& \ \sum_{i=1}^{n+1} P(A_i) + \sum_{k=2}^{n}(-1)^{k-1} \sum_{1 \le i_1 \le \cdots \le i_k \le n} P\left(\bigcap_{j=1}^{k} A_{i_j}\right) \\
& - \sum_{k=1}^{n}(-1)^{k-1} \sum_{1 \le i_1 \le \cdots \le i_k \le n} P\left(\bigcap_{j=1}^{k}(A_{i_j} \cap A_{n+1})\right) \\
=& \ \sum_{i=1}^{n+1} P(A_i) + \sum_{k=2}^{n}(-1)^{k-1} \sum_{1 \le i_1 \le \cdots \le i_k < n+1} P\left(\bigcap_{j=1}^{k} A_{i_j}\right) \\
& + \sum_{k=2}^{n}(-1)^{k-1} \sum_{1 \le i_1 \le \cdots \le i_k = n+1} P\left(\bigcap_{j=1}^{k}(A_{i_j} \cap A_{n+1})\right) \\
=& \ \sum_{k=1}^{n+1}(-1)^{k-1} \sum_{1 \le i_1 \le \cdots \le i_k \le n+1} P\left(\bigcap_{j=1}^{k} A_{i_j}\right)
\end{aligned}$$

より $n+1$ でも成立する．したがって, 任意の n で成立する．

■

包除原理の使用例は次のようなものである．

例 3.2.4. 組合せ確率で考える．n 人が 1 人ずつプレゼントを持ち寄り, 一旦回収しまた 1 人 1 つずつプレゼントを受けとるとする．このとき, 少なくとも 1 人は自分のプレゼントを手にする確率は次のように計算できる．n 個から k 個選ぶ組合せは $\binom{n}{k}$ であり, k 人が自分のプレゼントを受け取る確率は $\frac{(n-k)!}{n!}$ である．A_i を i 人目が自分のプレゼントを受け取る事象とすれば, 少なくとも 1 人は自分のプレゼントを手にする確率は,

$$\begin{aligned}
P\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{k=1}^{n}(-1)^{k-1} \sum_{1 \le i_1 \le \cdots \le i_k \le n} P\left(\bigcap_{j=1}^{k} A_{i_j}\right) \\
&= \sum_{k=1}^{n}(-1)^{k-1}\binom{n}{k}\frac{(n-k)!}{n!} = \sum_{k=1}^{n}\frac{(-1)^{k-1}}{k!}
\end{aligned}$$

となる． □

3.3 確率の連続性

確率にも, 実数の連続性のような性質がある．すなわち, 確率は極限の入れ替えが可能である．この命題を示すためにはまずは単調列の場合から証明する必要がある．

補題 3.3.1 (確率の連続性 (単調列))**.** $(\Omega, \mathcal{F}, P)$ を確率空間とする．単調な事象列 $A_1, A_2, \cdots \in \mathcal{F}$ について

$$\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n)$$

である． □

証明. (A_n) が単調増加列の場合, $B_1 = A_1, B_n = A_n \setminus A_{n-1}$ $(n \geq 2)$ とおけば, これは $\lim_{n\to\infty} A_n$ の分割であり, 各 B_n は排反だから

$$\begin{aligned}
P\left(\lim_{n\to\infty} A_n\right) &= P\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(\bigcup_{n=1}^{\infty} B_n\right) \\
&= \sum_{n=1}^{\infty} P(B_n) = \lim_{n\to\infty} \sum_{i=1}^{n} P(B_n) \\
&= \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} B_n\right) = \lim_{n\to\infty} P\left(A_n\right)
\end{aligned}$$

である. (A_n) が単調減少列の場合, $({A_n}^c)$ は単調増加列だから,

$$\begin{aligned}
P\left(\lim_{n\to\infty} A_n\right) &= P\left(\bigcap_{n=1}^{\infty} A_n\right) = P\left(\left(\bigcup_{n=1}^{\infty} {A_n}^c\right)^c\right) \\
&= 1 - P\left(\bigcup_{n=1}^{\infty} {A_n}^c\right) = 1 - P\left(\lim_{n\to\infty} {A_n}^c\right) \\
&= 1 - \lim_{n\to\infty} P\left({A_n}^c\right) = 1 - \lim_{n\to\infty} (1 - P(A_n)) = \lim_{n\to\infty} P(A_n)
\end{aligned}$$

である. ■

すでに有用な命題であるが, さらにこれを用いて, 極限を持つすべての事象列について連続性が成立することが示される.

定理 3.3.2 (確率の連続性)**.** $(\Omega, \mathcal{F}, P)$ を確率空間とする. 事象列 $A_1, A_2, \cdots \in \mathcal{F}$ が極限を持つとき

$$\lim_{n\to\infty} P(A_n) = P\left(\lim_{n\to\infty} A_n\right)$$

である. □

証明.

$$\bigcap_{i=n}^{\infty} A_i \subset A_n \subset \bigcup_{i=n}^{\infty} A_i$$

より,

$$P\left(\bigcap_{i=n}^{\infty} A_i\right) \leq P(A_n) \leq P\left(\bigcup_{i=n}^{\infty} A_n\right)$$

である. ここで, 補題 3.3.1 と事象列の極限の存在より,

$$\begin{aligned}
\lim_{n\to\infty} P\left(\bigcap_{i=n}^{\infty} A_i\right) &= P\left(\lim_{n\to\infty} \bigcap_{i=n}^{\infty} A_i\right) = P\left(\liminf_{n\to\infty} A_i\right) = P\left(\limsup_{n\to\infty} A_i\right) \\
&= P\left(\lim_{n\to\infty} \bigcup_{i=n}^{\infty} A_i\right) = P\left(\lim_{n\to\infty} \bigcup_{i=n}^{\infty} A_i\right) = \lim_{n\to\infty} P\left(\bigcup_{i=n}^{\infty} A_i\right)
\end{aligned}$$

だから, はさみうちの原理と事象列の極限の定義から

$$\lim_{n\to\infty} P(A_n) = P\left(\liminf_{n\to\infty} A_n\right) = P\left(\lim_{n\to\infty} A_n\right)$$

である. ■

次のような場合に使用することができる.

例 3.3.3. 表と裏の確率の等しいコインを投げ続ける試行について, 事象 A_n を 1 回目から n 回目まで常に表が出る事象とする. このとき $P(A_n) = \frac{1}{2^n}$ であり, $A = \lim_{n\to\infty} A_n$ は任意の N について N 回目まで裏が一度も出ない, すなわち永久に表のみが出る事象である. 確率の連続性を用いれば, この事象の確率は

$$P(A) = P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n) = 0$$

である. □

また, 確率の連続性を用いれば劣加法性を無限和の場合まで拡張できる.

定理 3.3.4 (劣加法性)**.** $(\Omega, \mathcal{F}, P)$ を確率空間とする. 事象列 $A_1, A_2, \cdots \in \mathcal{F}$ に対して

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

である. □

証明. 定理 3.2.3 と補題 3.3.2 より,

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(\lim_{n\to\infty} \bigcup_{i=1}^{n} A_n\right) = \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} A_n\right) \leq \lim_{n\to\infty} \sum_{i=1}^{n} P(A_n) = \sum_{n=1}^{\infty} P(A_n)$$

である. ■

劣加法性は次のような補題に応用できる. こちらも事象列の極限に関する命題である.

補題 3.3.5 (ボレル・カンテリの第 1 補題)**.** $(\Omega, \mathcal{F}, P)$ を確率空間とする. 事象列 $A_1, A_2, \cdots \in \mathcal{F}$ が $\displaystyle\sum_{n=1}^{\infty} P(A_n) < \infty$ を満たすならば, $P\left(\limsup_{n\to\infty} A_n\right) = 0$ が成立する. □

証明. 定理 3.3.2 と定理 3.2.3 から

$$\begin{aligned} P\left(\limsup_{nt\to\infty} A_n\right) &= P\left(\lim_{n\to\infty} \bigcup_{i=n}^{\infty} A_i\right) = \lim_{n\to\infty} P\left(\bigcup_{i=n}^{\infty} A_i\right) \\ &\leq \lim_{n\to\infty} \sum_{i=n}^{\infty} P(A_i) = \sum_{i=1}^{\infty} P(A_i) - \lim_{n\to\infty} \sum_{i=1}^{n-1} P(A_i) = 0 \end{aligned}$$

より従う. ■

3.4 条件付き確率

極限の次の話題は条件付き確率である.

定義 3.4.1. $(\Omega, \mathcal{F}, P)$ を確率空間とする. 事象 $A, B \in \mathcal{F}$ が $P(A) > 0$ を満たす時, **事象 A が与えられた時の事象 B の条件付き確率** $P(B \mid A)$ は次のように定義される.

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

□

$P(A \cap B) = P(A)P(B \mid A)$ なので, A が起こった後に B が起こる確率と考えれば, 妥当な定義だとわかる. また, 写像 $P(\cdot \mid A)$ は P の A への射影であると考えることができる. もう少し正しく言うべきだと考えたので, 筆者の考えを述べておく.

$$\mathcal{F}_A := \{A \cap B \mid B \in \mathcal{F}\}$$

と定義した時に, $P : \mathcal{F} \to [0,1]$ は, $P : \mathcal{F}_A \times \mathcal{F}_{A^c} \to [0,1] \times [0,1]$ と考えることができる. そして, このときの P の $\mathcal{F}_A$ への射影を $P_{\mathcal{F}_A} = P(\cdot \mid A)$ 定義すると. 2 変数版の値は, $P(B_1, B_2) = (x_1, x_2) \Rightarrow P(B_1 \cup B_2) = x_1 P(A) + x_2 P(A^c)$ となる. そしておそらく A への射影と書かれたのは, $(A, \mathcal{F}_A, P(\cdot \mid A))$ が確率空間となることを強調するためである. なお, $(\Omega, \mathcal{F}, P(\cdot \mid A))$ も確率空間になるそうである. (こちらの有用性はよくわからない.) 次の命題は, 実際にそうであることを示す命題である.

命題 3.4.2. $(\Omega, \mathcal{F}, P)$ を確率空間とする. このとき, $A \in \mathcal{F}$ に対して, $(\Omega, \mathcal{F}, P(\cdot \mid A)), (A, \mathcal{F}_A, P(\cdot \mid A))$ はそれぞれ確率空間となる. □

証明. $(\Omega, \mathcal{F}, P(\cdot \mid A))$ の証明. $P(\Omega \mid A) = \frac{P(A)}{P(A)} = 1$ かつ ${}^\forall B \in \mathcal{F}, P(B \mid A) = \frac{P(B \cap A)}{P(A)} \leq 1$ である. また, $B_1, B_2, \cdots \in \mathcal{F}$ を排反とすると, $B_n \cap A$ も排反であり,

$$P\left(\bigcup_{n=1}^{\infty} B_n \,\middle|\, A\right) = \frac{P(\bigcup_{n=1}^{\infty}(B_n \cap A))}{P(A)} = \sum_{n=1}^{\infty} \frac{P(B_n \cap A)}{P(A)} = \sum_{n=1}^{\infty} P(B_n \mid A)$$

となり, $P(\cdot \mid A)$ は $(\Omega, \mathcal{F})$ 上の確率測度である.

$(A, \mathcal{F}_A, P(\cdot \mid A))$ の証明. $P(A \mid A) = \frac{P(A)}{P(A)} = 1$ である. また, $\mathcal{F}_A \subset \mathcal{F}$ だから, 先ほどの証明はこちらの場合も適応できて $P(\cdot \mid A)$ は $(A, \mathcal{F}_A)$ 上の確率測度である. ■

次に条件付き確率の基本的な性質群を示す. まず 1 つ目は条件付き確率の乗法法則である.

補題 3.4.3 (乗法法則). $(\Omega, \mathcal{F}, P)$ を確率空間とする. $A_1, A_2, \cdots, A_n \in \mathcal{F}$ とすると, ($A_0 = \Omega$ とすれば)

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P\left(A_i \,\middle|\, \bigcap_{j=0}^{i-1} A_j\right) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \cdots P\left(A_n \,\middle|\, \prod_{i=1}^{n-1} A_i\right)$$

□

証明. 帰納法により示される. $n = 1$ の場合は自明である. $P(A \cap B) = P(A)P(B \mid A)$ だから, ある n で成立すると仮定すると,

$$P\left(\bigcap_{i=1}^{n+1} A_i\right) = P\left(A_{n+1} \,\middle|\, \bigcap_{j=0}^{n} A_j\right) P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n+1} P\left(A_i \,\middle|\, \bigcap_{j=0}^{i-1} A_j\right)$$

より示された. ■

$A_1, \cdots, A_n$ が同時に起こる確率は A_1 が起こって, その上で A_2 が起こって, ... と考えれば, これも直感的に妥当な命題である. また, 次のような命題もある.

補題 3.4.4. $(\Omega, \mathcal{F}, P)$ を確率空間とする. $A_1, A_2, \cdots \in \mathcal{F}$ を Ω の分割, $B \in \mathcal{F}$ とすると, 次の 2 式が成立する.

分割公式

$$P(B) = \sum_{i=1}^{\infty} P(A_i \cap B)$$

全確率の公式 $P(A_i) > 0$ $(i = 1, 2, \cdots)$ ならば,

$$P(B) = \sum_{i=1}^{\infty} P(A_i) P(B \mid A_i)$$

□

証明. 各 A_i は排反だから, $A_i \cap B$ は排反より

$$P(B) = P\left(\left(\bigcup_{i=1}^{\infty} A_i\right) \cap B\right) = P\left(\bigcup_{i=1}^{\infty} (A_i \cap B)\right) = \sum_{i=1}^{\infty} P(A_i \cap B)$$

となり, 分割公式が成立する. また, $P(A_i \cap B) = P(B)P(A_i \mid B)$ より全確率の公式が従う. ■

また, $A_i = \emptyset (i > n)$ とすれば, n までの有限和の場合でも成立することがわかる.

条件付き確率は, 次のような場面で役立つ.

例 3.4.5 (ポリアの壺)**.** a 個の赤玉と b 個の白玉の入った壺がある. 玉を 1 つランダムに取り出し, 取り出した玉と同じ色のたまを c 個加えて壺に戻す. この試行を繰り返すとき, n 回目に赤玉を取り出す確率を求める. $P_n(a, b)$ を赤玉 a 個白玉 b 個で開始して n 回目に赤玉を取り出す確率とする. まず, $P_1(a, b) = \frac{a}{a+b}$ である. 実は, 任意の n で $P_n(a, b) = \frac{a}{a+b}$ となるのだが, これを帰納法で示す. ある n で $P_n(a, b) = \frac{a}{a+b}$ と仮定する. $n + 1$ 回目に赤を引く確率を, 1 回目に赤を引いたのち, 赤, 白を $a + c, b$ で開始した時の n 回目に赤を引く確率と, 白の同様な場合で分ける. これらは排反であるから,

$$P_{n+1}(a, b) = P_1(a, b)P_n(a + c, b) + P_2(a, b)P_n(a, b + c) = \frac{a}{a+b} \cdot \frac{a+c}{a+b+c} + \frac{b}{a+b} \cdot \frac{a}{a+b+c} = \frac{a}{a+b}$$

となる. したがって示された. □

また, 条件付き確率について他にも有用な命題が存在する.

補題 3.4.6 (ベイズの公式)**.** $(\Omega, \mathcal{F}, P)$ を確率空間とする. $A_1, A_2, \cdots \in \mathcal{F}$ を Ω の分割で, $P(A_i) > 0$ $(i = 1, 2, \cdots)$ とする. また, $B \in \mathcal{F}$ も $P(B > 0)$ とすると, 各 i で

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{\sum_{j=1}^{\infty} P(A_j)P(B \mid A_j)}$$

である. □

証明. 全確率の公式を用いれば

$$P(A_i \mid B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B \mid A_i)}{\sum_{j=1}^{\infty} P(A_j)P(B \mid A_j)}$$

である. ■

当然有限の場合も成立する. これについても面白い例がある.

例 3.4.7. 次の感染症の検査を考える

- 感染者の 99% は陽性と診断され, 1% は陰性と診断される.
- 非感染者の 2% は陽性と診断され, 98% は陰性と診断される.

1000 人に 1 人が感染している状況で, 陽性と診断された人が感染者である確率を考える. A を感染している事象, B を陽性が診断される事象とすると, ベイズの公式から,

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(A)P(B \mid A) + P(A^c)P(B \mid A^c)} = \frac{0.001 \times 0.99}{0.001 \times 0.99 + 0.999 \times 0.02} < 0.05$$

となる. これだけ正確性があっても, 陽性の診断が正しい確率は 5% 未満となってしまうわけである. □

3.5 確率の独立性

最後に紹介される確率の性質は, 独立性である.

定義 3.5.1 (事象の独立性)**.** $(\Omega, \mathcal{F}, P)$ を確率空間として, $A_1, \cdots, A_n \in \mathcal{F}$ とする. $\{1, \cdots, n\}$ の空でない任意の有限部分集合 I で

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)$$

が成立するとき, $A_1, \cdots, A_n$ は**互いに独立**であるという. また, 無限列の場合も $\emptyset \neq {}^{\forall} I \subset \mathbb{N}$ として同様に定義される. □

なお定義から明らかであるが, 違いに独立である事象列の部分列も違いに独立となる. また, 独立性は余事象を取るなどしても変化しない.

命題 3.5.2. $(\Omega, \mathcal{F}, P)$ を確率空間, $A_1, \cdots, A_n \in \mathcal{F}$ を互いに独立な事象とする. 任意の $I \subset \{1, \cdots, n\}$ に対して, $B_i := \begin{cases} A_i, & (i \in I) \\ {A_i}^c, & (i \notin I) \end{cases}$ としたとき, $B_1, \cdots, B_n$ は独立である. 無限列の場合も成立する. □

証明. $\{B_i\}$ の任意の有限な部分集合族で独立性が示されればよい. $|I| = 1$ の場合のみを考えれば帰納的に示されるから, その場合のみ証明する. $A_1, \cdots, A_k, {A_{k+1}}^c$ が独立であることを示す. 分割公式と $A_1, \cdots, A_{k+1}$ の独立性を用いれば,

$$\begin{aligned} P\left({A_{k+1}}^c \cap \bigcap_{i=1}^{k} A_i\right) &= P\left(\bigcap_{i=1}^{k} A_i\right) - P\left(\bigcap_{i=1}^{k+1} A_i\right) \\ &= \prod_{i=1}^{k} P(A_i) - \prod_{i=1}^{k+1} P(A_i) \\ &= (1 - P(A_{k+1})) \prod_{i=1}^{k} P(A_i) \\ &= P\left({A_{k+1}}^c\right) \prod_{i=1}^{k} P(A_i) \end{aligned}$$

である. ■

互いに独立な無限事象列については, 次の直感的な式が成立する.

補題 3.5.3. $(\Omega, \mathcal{F}, P)$ を確率空間とする. $A_1, A_2, \cdots \in \mathcal{F}$ が違いに独立の時,

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \prod_{n=1}^{\infty} P(A_n)$$

である. □

証明.

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i)$$

だから, 確率の連続性より明らかである. ■

この補題を用いて, ボレル・カンテリの補題の発散バージョンが示される.

補題 3.5.4 (ボレル・カンテリの第 2 補題)**.** $(\Omega, \mathcal{F}, P)$ を確率空間とする. 互いに独立な事象列 $A_1, A_2, \cdots \in \mathcal{F}$ に対して,

$$\sum_{i=1}^{\infty} P(A_i) = \infty \Rightarrow P\left(\limsup_{n\to\infty} A_n\right) = 1$$

である. □

証明. 確率の連続性と, $1 - x \geq e^{-x}$, ${A_i}^c$ の独立性から,

$$\begin{aligned} P\left(\limsup_{n\to\infty} A_n\right) &= P\left(\lim_{n\to\infty}\bigcup_{i=n}^{\infty} A_i\right) = \lim_{n\to\infty} P\left(\bigcup_{i=n}^{\infty} A_i\right) \\ &= \lim_{n\to\infty} P\left(\left(\bigcap_{i=n}^{\infty} {A_i}^c\right)^c\right) = 1 - \lim_{n\to\infty} P\left(\bigcap_{i=n}^{\infty} {A_i}^c\right) \\ &= 1 - \lim_{n\to\infty}\prod_{i=n}^{\infty}(1 - P(A_i)) = 1 - \lim_{n\to\infty}\prod_{i=n}^{\infty} e^{-P(A_i)} \\ &\geq 1 - \lim_{n\to\infty}\exp\left(-\sum_{i=n}^{\infty} P(A_i)\right) = 1 \end{aligned}$$

なお最後の等号は, 仮定より, $\sum_{i=n}^{\infty} P(A_i) = \infty$ となることによる. ■

これは次のような使用例がある.

例 3.5.5. 裏と表が出る確率が共に 0 でないコインを無限回投げる時, i 回目が表となる事象 A_i を考える。まず, これらは独立であり, かつ $0 < P(A_i)$ は定数であるから $\displaystyle\sum_{i=1}^{\infty} A_i = \infty$ となる. ゆえに, ボレル・カンテリの第 2 補題より,

$$P\left(\limsup_{n\to\infty} A_n\right) = 1$$

である. したがって, 表が無限回出る確率は 1 であり, 同様の議論により, 裏が無限回出る確率も 1 となる. □

また, 独立な事象に関連して, 直積確率空間というものがある. 独立な 2 つの試行を同時に行うイメージである.

定義 3.5.6 (直積確率空間)**.** $(\Omega_1, \mathcal{F}_1, P_1), (\Omega_2, \mathcal{F}_2, P_2)$ をそれぞれ確率空間とする．それぞれ $\Omega, \mathcal{F}, P$ を

$$\begin{cases} \Omega := \Omega_1 \times \Omega_2 \\ \mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2 := \sigma(\{A_1 \times A_2 \mid A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2\}) \\ P : \Omega \to [0,1], \quad P(A_1 \times A_2) \mapsto P_1(A_1)P_2(A_2) \end{cases}$$

としたときの $(\Omega, \mathcal{F}, P)$ を $\mathcal{F}_1, \mathcal{F}_2$ の**直積確率空間**という．n 個の確率空間に対する直積確率空間も同様に定義される． □

なお, 直積確率空間は確率空間となるが, この証明は測度論の知識が必要そうなので, 今回は省略する．また, 直積確率空間が独立性に関係しているというのは, 次のようなことである．2 つのサイコロを振る状況において, この 2 つの確率空間の直積を考える．1 つ目のサイコロについての事象 A_1 と 2 つ目のサイコロについての事象 A_2 は直積確率空間上では $A_1 \times \Omega_2, \Omega_1 \times A_2$ となるが, このときの積事象の確率は

$$P((A_1 \times \Omega_2) \cap (\Omega_1 \times A_2)) = P(A_1 \times A_2) = P_1(A_1)P_2(A_2) = P(A_1 \times \Omega_2)P(\Omega_1 \times A_2)$$

となり, これらは独立であることがわかる．このように別の確率空間に由来する 2 事象は常に独立となるということである．

4 分布関数

4.1 絶対連続分布

離散確率空間における確率質量関数を一般的な確率空間に拡張したようなものを考えることができる．

定義 4.1.1 (絶対連続分布)**.**

(1) $f : \mathbb{R} \to \mathbb{R}$ が

$$\forall x \in \mathbb{R}, f(x) \geq 0 \land \int_{-\infty}^{\infty} f(x)dx = 1$$

を満たす時, f を**確率密度関数**という．

(2) f を確率密度関数とする．$\mathcal{P}_f : \mathcal{B}(\mathbb{R}) \to [0,1]$ を,

$$\mathcal{P}_f(A) := \int_A f(x)dx$$

と定義する．このとき, 積分の線形性から $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{P}_f)$ は確率空間となる．

(3) 確率密度関数によって定めることのできる確率測度は, **絶対連続分布**を持つという．

□

下は有名な確率密度関数の例である．

例 4.1.2.

(1) $a, b \in \mathbb{R}(a < b)$ に対して

$$f_{\mathrm{uni}}(x) = \begin{cases} \dfrac{1}{b-a}, & x \in [a,b] \\ 0 \end{cases}$$

で定義される分布を**一様分布**という．

(2) $\mu > 0$ に対して,

$$f_{\exp}(x) = \begin{cases} \mu e^{-\mu x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

で定義される分布を**指数分布**という.

(3) $m \in \mathbb{R}, \sigma > 0$ に対して,

$$f_{\mathrm{nor}}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$$

で定義される分布を**正規分布**という. 特に $m = 0, \sigma = 1$ のとき**標準正規分布**という.
また, これが確率密度関数となる証明は, $\sigma^{-1}(x-m)$ を x に置き換えられるので, 標準正規分布の場合のみ考えれば良い.

$$\left(\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx\right)^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x^2+y^2}{2}} dxdy = \int_0^{2\pi}\int_0^{\infty} e^{-\frac{r^2}{2}} rdrd\theta = 2\pi \int_0^{\infty} e^{-\frac{r^2}{2}} d\frac{r^2}{2} = 2\pi$$

より示される.

□

これで, 節 1.4 で離散確率空間では表しきれないと言及した, $[0,1]$ 上にランダムに 1 点を打つという試行における確率を, 一様分布により定義できるようになった.

n 次元についても同じように絶対連続分布が定義される.

定義 4.1.3 (n 次元絶対連続分布)**.**

(1) $f : \mathbb{R}^n \to \mathbb{R}$ が

$$\forall \boldsymbol{x} \in \mathbb{R}^n, f(\boldsymbol{x}) \geq 0 \wedge \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(\boldsymbol{x}) dx_1 \cdots dx_n = 1$$

を満たす時, f を n **次元確率密度関数**という.

(2) f を確率密度関数とする. $\mathcal{P}_f : \mathcal{B}(\mathbb{R}^n) \to [0,1]$ を,

$$\mathcal{P}_f(A) := \int_A f(\boldsymbol{x}) dx_1 \cdots dx_n$$

と定義する. このとき, 積分の線形性から $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \mathcal{P}_f)$ は確率空間となる.

(3) n 次元確率密度関数によって定めることのできる確率測度は, n **次元絶対連続分布**を持つという.

□

n 次元の場合も正規分布が有名である.

例 4.1.4. $\boldsymbol{m} \in \mathbb{R}^n, \Sigma \in \mathbb{R}^{n \times n}$ とし, また Σ を実対称正定値行列とする. このとき,

$$f_{\mathrm{nor}}(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}{}^t(\boldsymbol{x}-\boldsymbol{m})\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{m})\right)$$

と定義される分布を n **次元正規分布**という.

これが確率密度関数となる証明は次の通りである.

まず, Σ は実対象正定値行列であることから, 実直交行列 U を用いて, $\Sigma = UD^tU$ と対角化できる. ここで, D の各成分の平方根を取った対角行列と U の積を L とすれば, L は正則であり, $\Sigma = L^tL$ と変換できる. また, $\det \Sigma = \det L \det {}^tL = (\det L)^2$ である. ここで, $\boldsymbol{y} = L^{-1}(\boldsymbol{x} - \boldsymbol{m})$ と変換すれば, ヤコビアンは $\det L = \sqrt{\det \Sigma}$ だから,

$$\begin{aligned}\int_{\mathbb{R}^n} f_{\mathrm{nor}}(\boldsymbol{x}) dx_1 \cdots dx_n &= \int_{\mathbb{R}^n} \frac{1}{\sqrt{2\pi}^n} \exp\left(-\frac{||\boldsymbol{y}||^2}{2}\right) \\ &= \prod_{i=1}^{n} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y_i^2}{2}\right) dy_i = 1\end{aligned}$$

となる. □

先ほど離散確率空間で表せないものを表せるようになったという話があったが, これでもまだ表せないものがある. 例えば, 絶対連続分布を持つ確率において, 1 点集合の確率を計算すると,

$$\mathcal{P}_f(\{a\}) = P\left(\lim_{n\to\infty}\left(a - \frac{1}{n}, a\right]\right) = \lim_{n\to\infty} P\left(\left(a - \frac{1}{n}, a\right]\right) = \lim_{n\to\infty}\int_{a-\frac{1}{n}}^{a} f(x)dx = 0$$

となってしまう. つまり, 1 点集合に正の確率を持たせることができない. そこでさらに一般化されたものが分布間数である.

4.2 分布関数

まずは分布関数の定義から入る.

定義 4.2.1 ((1 次元) 分布関数)**.** $F : \mathrm{R} \to [0, 1]$ が次の性質を満たす時, F を **(1 次元) 確率分布関数**という.

(1) $x < y \Rightarrow F(x) \leq F(y)$ (**単調性**)
(2) $\displaystyle\lim_{h\to+0} F(x + h) = F(x)$ (**右連続性**)
(3) $\lim_{x\to-\infty} = 0, \lim_{x\to\infty} F(x) = 1$

□

$\mathcal{P}_F : \mathcal{B}(\mathbb{R}) \to [0, 1]$ を,

$$\mathcal{P}_F((-\infty, x]) = F(x)$$

として σ-加法性を満たすように $\mathcal{B}(\mathbb{R})$ へ拡張すると, $\mathcal{P}_F$ は $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度となる.

また, 確率質量関数や確率密度関数により定義される確率測度も分布関数で表すことができる.

例 4.2.2. (1) $p : \mathbb{Z} \to [0, 1]$ を確率分布関数とする. このとき,

$$F(x) = \sum_{i=-\infty}^{\lfloor x \rfloor} p(i)$$

とすると, これは分布関数となる. また, $k \in \mathbb{Z}$ に対して,

$$\mathcal{P}_F(\{k\}) = \lim_{n\to\infty} \mathcal{P}_F\left(\left(k - \frac{1}{n}, k\right]\right) = \lim_{n\to\infty}\left(F(k) - F\left(k - \frac{1}{n}\right)\right) = p(k)$$

であり, また,

$$\mathcal{P}_F((k,k+1)) = F(k) - \lim_{n\to\infty} F\left(k - \frac{1}{n}\right) = 0$$

となる. つまり, 確率質量関数 p を $\mathbb{Z}$ から $\mathbb{R}$ へと拡張したようなものとなっている.

(2) $f : \mathbb{R} \to \mathbb{R}$ を確率密度間数とする. このとき,

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

は分布関数となる. また, このとき F は連続となり, 区分的に微分可能となる. これについても, f により定義される確率測度と, F により定義される確率測度は一致する.

□

上の例で示したように, 不連続な分布の場合は, 1 点集合で正の確率測度を持つ.

命題 4.2.3. F を分布関数, それにより定義される $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度を $\mathcal{P}_F$ とする. このとき, 次の同値関係が成立する.

$$F \text{ が } a \in \mathbb{R} \text{ で不連続} \iff \mathcal{P}_F(\{a\}) > 0$$

□

証明.

$$\mathcal{P}_F(\{a\}) = \lim_{n\to\infty} \left(F(a) - F\left(a - \frac{1}{n}\right)\right)$$

であるから, F が a で不連続であれば, $F(a) > F(a - \frac{1}{n})$ が任意の n で成立することから, $\Rightarrow$ は従う. 逆に $mathcalP_F(\{a\}) > 0$ であれば, F の連続性が成立せず, $\Leftarrow$ が示された. ■

索引