Kolmogorof's 0-1 law

かわわん

概要

コロモゴロフの 0-1 法則は確率論で登場する変わった定理です．とある確率空間ではすべての事象が 1 もしくは 0 の必ずどちらかの確率で起こるというものです．しかしながら，その確率空間から事象を一つ取り出したときに，その確率が 1 か 0 のどちらなのかが分からないのです．この文書はわたしがセミナーで行った内容の焼き直しです．主に Williams [1] の第四章の内容です．証明で飛んでいると思われる部分を埋めたもになります．今後 Williams [1] を読む方でこの記事を見つけた方がいらっしゃる場合は参考にしていただければ幸いです．また，この記事の前半の内容は現代確率論のモチベーションと，それに必要な測度論と呼ばれるものをかいつまんで説明したものになります．低回生の方もここは読んで欲しいと思います．

目次

1　準備

この節は準備の節です．しかしながら, 基本的な集合演算や $\forall, \exists$ のような数学でよく見かける基本的な記号の意味[*1] は知っているものと仮定します.

1.1　測度論

現代確率論は「測度論」と呼ばれる分野の上に立っています．高校までの確率論とはかなり雰囲気が異なっています．わざわざ新しい理論を導入する理由ですが, 高校までの確率論の知識だと, 起こりえる結果が無限にある場合には太刀打ちできないからです．「起こりえる結果が無限にある場合」とは現実世界にありふれていて, 例えば株価がそうです．ある時に株価が α ドルであったとすると, その次の瞬間に株価がどうなるでしょうか．無限個のパターンがあると思いませんか. そのような問題を次のようなアイデアである意味克服します.

事象を集合と捉えて, その集合の「大きさ」を確率とみなす

もちろん, 「大きさ」の最大値は 1 であり, 最小値は 0 です (「大きさ」を確率とみなしているのですからね).

「大きさ」の測り方はいろいろあります．例えば, $A = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1, 0 \leq y \leq 1\}$ という集合だとどうなるでしょうか．「大きさ」の意味が「(通常の意味での) 面積の大きさ」であったとすると, A の大きさは $1 \times 1 = 1$ となります．「大きさ」の意味が「その図形がなす辺の周の総和」であるとしたら A の大きさは $1 + 1 + 1 + 1 = 4$ となるでしょう．「大きさ」の意味が「その図形に内接する円の面積」であるとするならば, A の大きさは $\frac{1}{2} \times \frac{1}{2} \times \pi = \frac{1}{4}\pi$ となります[*2]. このように集合の「大きさ」を測る” きまり” のことを**測度**といいます.

さて, 測度を一つ定めます[*3]．次に一つの集合を取ってきます．さっき定めた測度でその集合の「大きさ」を測ることができるのでした．少し図にあらわすならば, 以下のようになります.

$$\text{集合} \xrightarrow{\text{測度}} \text{集合の大きさ}$$

これは一種の関数ですね．そう, 測度とは集合の「大きさ」を測る関数のことです．つまり, 定義域が集合族で値域が実数の関数です！

さて, 測度論で何をしているのかということの直観的な説明はこのくらいにして, ここらで物事をきちんと定義していきましょう.

[*1] 一回生の数学序論で学ぶような内容

[*2] まるで集合はユークリッド空間の部分集合であるものしか認めないみたいな書き方してるけどそんなことないよ

[*3] 先ほど述べた” きまり” を一つ定めるということです.

1.2 用語の定義

1.2.1 集合関数

定義 1.1. $\mathcal{F}$ を空でない集合族としたとき, 関数 m

$$m : \mathcal{F} \to \mathbb{R}$$

を集合関数という

この集合関数の中でも集合の大きさを測る写像 (現在の目標) のことを測度といいます.

1.2.2 σ-algebra

「事象を集合と捉える」と先ほど述べました. 「事象をすべて集めてきた集合」は **σ-algebra** とよばれ, 次の性質を満たしています.

定義 1.2. Ω を集合,$\mathcal{F}$ を Ω の部分集合族とする.$\mathcal{F}$ が σ-algebra であるとは次を満たすことである.

1. $X, \emptyset \in \mathcal{F}$
2. $A \in \mathcal{F} \Rightarrow A^c := \Omega - A \in \mathcal{F}$
3. $A_1, A_2, \cdots \in \mathcal{F} \Rightarrow \displaystyle\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}$

σ-algebra はよく, σ-alg. と略記します. この記事でもそのように略記することにします.

1.2.3 σ-algebra と呼ばれるゆえん

この節は補足的なので飛ばして構いません

σ-alg. は日本語訳すると「シグマ代数」となります. なんで突然「代数」というワードが出てくるのでしょうか. それは, 集合演算で閉じているからです. $\cup$ を環でいうところの $+$, $\cap$ を $*$ と思いましょう. いろいろ試してみてください. 可算回集合演算をしてもきちんと閉じています. たとえば, $A, B, C, D, F_1, F_2, F_3, \cdots \in \mathcal{F}$ として,

$$\bigcap_{i \in \mathbb{N}} F_i \cup ((A \cap (B \cup C)^c)^c \cup D)$$

が $\mathcal{F}$ の元であることを確かめてみてください[*4].

1.2.4 可測空間

定義 1.3. 集合 Ω と Ω 上の σ-alg. $\mathcal{F}$ があるとき, 組 $(\Omega, \mathcal{F})$ を可測空間という.

[*4] この式に特に意味はありません

1.2.5 「組」ということばについて

数学でお馴染みの言葉ですが, 聞きなれない人もいるかと思われるので一応説明しておきます. 位相空間や群論を学んだことがある人にとって今更なことなので飛ばしてください. 一回生だと聞きなれない言葉と思われるため説明しています.

「組」とは「いくつかの数学的な対象をまとめたもの」のことです. 例えば, 先ほどの定義した「可測空間」は, 集合 Ω と Ω 上の σ-alg. $\mathcal{F}$ の組 $(\Omega, \mathcal{F})$ のことです.

これは僕が好きなたとえなのですが, マクドナルドのハッピーセットを思い浮かべてください. 何を思い浮かべますか? きっと,「ハンバーガーとポテトとジュースとおもちゃ」のイメージが頭の中に浮かんだと思います. ハンバーガーはハンバーガーでしかないし, ポテトもポテトでしかないし, 残り二つもそれはそれでしかありません. しかし, この四つを一緒に考えると「ハッピーセット」という別名が付きます. これは,

四つ組 (ハンバーガー, ポテト, ジュース, おもちゃ) のことを「ハッピーセット」と呼ぶ.

ということです. イメージがついたでしょうか? まぁ, おまけ程度の話なのでよくわからなかったら流してもらえばいいんですけれど...

1.2.6 測度

定義 1.4. 集合 Ω と Ω 上の σ-alg. $\mathcal{F}$ がある. $\mathcal{F}$ 上の集合関数 $\mathbb{P}$ が測度であるとは次を満たすことである.

1. $0 \leq \mathbb{P}(A) \leq \infty$ 特に,$\mathbb{P}(\emptyset) = 0$
2. $A_n \in \mathcal{F}(n = 1, 2, \cdots)$ が互いに交わらない $\Rightarrow \mathbb{P}(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$

測度によって集合の「大きさ」を測ることができるわけですが, この「大きさ」のことを**集合の測度**とよびます.

1.2.7 測度の性質

測度の性質として以下のようなものがあります. 集合の大きさを測るものなので定義より以下の性質を持つことが示されます.

命題 1.1.

$$\begin{cases} A, B \in \mathcal{F}, A \supset B \text{ ならば } \mathbb{P}(A) \geq \mathbb{P}(B) \\ \text{特に } \mathbb{P}(B) < +\infty \text{ ならば } \mathbb{P}(A - B) = \mathbb{P}(A) - \mathbb{P}(B) \quad (\text{単調性}) \end{cases}$$

$$\begin{cases} A_1, A_2, \cdots \in \mathcal{F} \text{ ならば } \mathbb{P}(\bigcup_{j \in \mathbb{N}} A_j) \leq \sum_{j=1}^{\infty} \mathbb{P}(A_j) \quad (\text{劣加法性}) \end{cases}$$

2 確率空間

さて, 確率論の話へと進みましょう. まだ準備段階です.

2.1 測度空間と確率空間

定義 2.1. 集合 Ω と Ω 上の σ-alg. $\mathcal{F}$, そして $\mathcal{F}$ 上の測度 $\mathbb{P}$ があるとき, 三つ組 $(\Omega, \mathcal{F}, \mathbb{P})$ を測度空間という．特に, $\mathbb{P}(\Omega) = 1$ であるとき, 確率空間という．

Ω は標本空間と呼ばれます．これは結果を全て集めた集合です．何度も述べたように, $\mathcal{F}$ は事象全体を集めた集合になります．また, $\mathbb{P}$ は確率測度と呼ばれます．

2.1.1 例

高校までの確率の話を今の話に沿って考えてみましょう．

問題. 1,2,3 と書かれた 3 枚のカードが箱の中にある．それらを一枚ずつ取り出して順番に並べ, 3 桁の数を作る．出来上がった数が 200 以上になる確率はいくらか.

まずは高校までの通りに解いてみましょう．

解答. 全部で $3! = 6$ 通りの数が作れて, その内 200 以上になるのは $123, 132$ を除くもの全てなので, 求める確率は $(\frac{4}{6} =)\frac{2}{3}$ と求められます．

次に, 確率空間の言葉を使って書いてみましょう．

解答. この問題だと, $\Omega = \{123, 132, 213, 231, 312, 321\}$ です．$\mathcal{F} = 2^{\Omega}$ です．問題の確率は $\{213, 231, 312, 321\}$ の測度です．確率測度は次のようなものになります．

$$\mathcal{F} \ni A \mapsto \mathbb{P}(A) := \frac{card(A)}{6} \in \mathbb{R}$$

ここで $card(A)$ とは A の濃度のことを指します．つまり, A の元の個数のことです．

3 Generated σ-alg.

ここからは確率論で重要になる概念を定義します．位相空間はまぁなんか頑張って (爆).

3.1 Generated σ-alg.

定義 3.1. 集合族 $\mathcal{A}$ が与えらている．$\mathcal{A}$ から生成される σ-alg. とは, $\mathcal{A}$ を含む最小の σ-alg. のことである．$\sigma(\mathcal{A})$ と書く．

Generated σ-alg. はその最小性が重要になります．コロモゴロフの 0-1 定理の証明でも威力を発揮します．

3.1.1 Generated σ-alg. に関する証明について

$\mathcal{A}, \mathcal{F}$ を集合族とします．しばしば, 以下の等式を証明する場面に遭遇します．

$$\sigma(\mathcal{A}) \subset \mathcal{F}$$

通常ならば,

$$A \in \sigma(\mathcal{A}) \Rightarrow A \in \mathcal{F}$$

を証明すればよいです．しかし，このように定義に沿った証明をしようとすると困難が伴います[*5]．そこで，Generated σ-alg. の最小性に注意すれば，以下を示せばよいことがわかります．

$$\mathcal{F} \text{が} \sigma - alg. \quad \text{かつ} \quad \mathcal{A} \subset \mathcal{F} \tag{$*$}$$

この証明方法だと苦労は少ないです．実際，定理 6.2 の証明の $(step1)$ においてこの手法を用いています．

3.2 Borel σ-alg.

定義 3.2. $\mathbb{R}$ 上の通常の位相[a]を $\mathcal{O}$ とする．$\sigma(\mathcal{O})$ のことを Borel σ-alg. といい，$\mathcal{B}(\mathbb{R})$，もしくは単に，$\mathcal{B}$ と表わす．

[a] 開区間全体の集合を開基とする位相空間

後に述べますが，確率変数を考える際に $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ という可測空間をよく考えます．

3.3 重要な集合族：π-system

定義 3.3. 集合族 $\mathcal{I}$ が π-system であるとは次が成立すること．

$$I_1, I_2 \in \mathcal{I} \Rightarrow I_1 \cap I_2 \in \mathcal{I}$$

π-system は構造が簡単で扱いやすいです．π-system で生成される σ-alg. についてよく考察します．以下に二つの重要な結果を述べます．今回の目標とはあまり関係が無いので証明は書かないでおきます[*6]．

3.3.1 重要な結果その一

補題 3.1. $\pi(\mathbb{R}) = \{(-\infty, a] \mid a \in \mathbb{R}\}$ とおくと，$\pi(\mathbb{R})$ は π-system であり，しかも $\sigma(\pi(\mathbb{R})) = \mathcal{B}(\mathbb{R})$ が成立する．

この命題は大変便利です．というのも，普通 $\mathcal{B}(\mathbb{R})$ の元というのは具体的に考察するのは困難を極めます．しかし，この補題のおかげで $\pi(\mathbb{R})$ という分かりやすい集合に着目すればよいことになるからです．

[*5] A は σ-alg. の元であるため，具体的な形を書くことができない．

[*6] 証明は長いし易しくないし

3.3.2 重要な結果その二

補題 3.2. Ω を集合とする．$\mathcal{I}$ を Ω 上の π-system とする．$\sum = \sigma(\mathcal{I})$ と定める．可測空間 $(\Omega, \sum)$ 上の二つの測度 μ_1, μ_2 があり，$\mu_1(\Omega) = \mu_2(\Omega) < \infty$ かつ，$\mathcal{I}$ 上で $\mu_1 = \mu_2$ とする．このとき，$\sum$ 上で $\mu_1 = \mu_2$ である．

この補題の証明は少し長いうえに今回の目標とは異なるため証明は省略します．この補題のおかげで，σ-alg. 上全ての元について $\mu_1 = \mu_2$ であるかどうかを調べる必要がなくなるため便利です．

4 可測関数と確率変数

4.1 可測関数

測度空間上の関数で，とある性質を満たすもののことを可測関数といいます．位相空間上の連続関数のように，可測関数とは可測空間上での「行儀のよい」関数のことです．

定義 4.1. 可測空間 $(\Omega, \mathcal{F})$ から $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ への関数 X が可測関数とは，次を満たすことである．

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

以下は定義の言い換えです．とはいえ，写像の言葉に直しただけです．ただ，読みやすさのためにすこしインフォーマルに (形式に沿っていない書き方) しています．

定義 4.2. 可測空間 $(\Omega, \mathcal{F})$ から $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ への関数 X が可測関数とは，次を満たすことである．即ち，対応

$$X^{-1} : \mathcal{B} \to \mathcal{F}$$

が写像になっていることである．ただし，X^{-1} という対応は

$$B \in \mathcal{B}(\mathbb{R}) \text{ に対して } \{\omega \in \Omega \mid X(\omega) \in B\} \ (B \text{ による } X \text{ の逆像})$$

を対応させるものである．

可測関数はそれ自体考察しがいのある重要な概念ですが，確率論では次の定義が大切です．

4.2 確率変数

定義 4.3. 確率空間 $(\Omega, \mathcal{F}, \mathbb{P})$ から $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ への関数 X が可測関数のとき，X を確率変数という．

確率変数は次の同値な条件に言い換えることができます．

命題 4.1. 次は同値.

1. X が確率変数
2. $\forall a \in \mathbb{R}, \{X \le a\} \in \mathcal{F}$

ただし, $\{X \le a\} := \{\omega \in \Omega \mid X(\omega) \le a\}$.

証明. (1. $\Rightarrow$ 2.) の証明 : $(-\infty, a] \in \mathcal{B}(\mathbb{R})$ だから仮定より, $\{X \le a\} = X^{-1}((-\infty, a]) \in \mathcal{F}$.

(2. $\Rightarrow$ 1.) の証明 : $\mathcal{B}_0 := \{A \in \mathcal{B}(\mathbb{R}) \mid X^{-1}(A) \in \mathcal{F}\}$ とする. このとき, $\mathcal{B}_0$ は σ-alg である. このことを確認しておこう.

- $X^{-1}(\mathbb{R}) = \Omega$ より, $\mathbb{R} \in \mathcal{B}_0$.
- $A \in \mathcal{B}_0$ とする. このとき, $X^{-1}(A^c) = (X^{-1}(A))^c \in \mathcal{F}$. したがって, $A^c \in \mathcal{B}_0$
- $A_1, A_2, \cdots \in \mathcal{B}_0$ とする. このとき, $X^{-1}(\bigcup_{i\in\mathbb{N}} A_i) = \bigcup_{i\in\mathbb{N}} X^{-1}(A_i) \in \mathcal{F}$ であり, このとき, $\bigcup_{i\in\mathbb{N}} A_n \in \mathcal{F}$

もちろん, 仮定より, $\{(-\infty, a] \mid a \in \mathbb{R}\} \subset \mathcal{B}_0$ である. $(*)$ より次が分かる.

$$\sigma(\{(-\infty, a] \mid a \in \mathbb{R}\}) \subset \mathcal{B}_0$$

上式の左辺は補題 3.1 より $\mathcal{B}(\mathbb{R})$ に等しい. ここまでで,

$$\mathcal{B}(\mathbb{R}) \subset \mathcal{B}_0$$

が分かった. さらに $\mathcal{B}_0$ の定義より,

$$\mathcal{B}_0 \subset \mathcal{B}(\mathbb{R})$$

これで証明が終わる. □

4.3 確率変数が生成する σ-alg.

定義 4.4. X を関数とする. $\sigma(X)$ とは X が可測関数になるような最小の σ-alg. のことである.

これはコロモゴロフの 0-1 法則で登場する”tail σ-alg. ” で登場します.

5 独立

独立性は確率変数を用いた定義をよく見ますが, σ-alg. から話を進めましょう.

5.1 独立な σ-alg.

定義 5.1. $\mathcal{F}$ を σ-alg. とする．集合 Λ で添え字づけられた[a] $\mathcal{F}$ の部分 σ-alg. の族 $\{\mathcal{F}_\lambda\}_{\lambda\in\Lambda}$ が独立であるとは, 任意の Λ の有限部分集合 $\{\lambda_1,\cdots,\lambda_n\}$ と各 $F_{\lambda_i}\in\mathcal{F}_{\lambda_i}$ に関して次が成り立つこと.

$$\mathbb{P}(F_{\lambda_1}\cap F_{\lambda_2}\cap\cdots\cap F_{\lambda_n})=\prod_{k=1}^{n}\mathbb{P}(F_{\lambda_k})\qquad(\lambda_1,\cdots,\lambda_k\text{は互いに異なる})$$

ただし, $\mathcal{G}$ が $\mathcal{F}$ の部分 σ-alg. であるとは, $\mathcal{G}\subset\mathcal{F}$ かつ $\mathcal{G}$ が σ-alg. であることを言う.

[a] 一般に非可算ですが今回考えるのは $\Lambda=\mathbb{N}$ の場合です.

定義 5.2. 確率変数 $X_1,X_2,\cdots$ が独立とは, $\sigma(X_1),\sigma(X_2),\cdots$ が独立であること.

定義 5.3. 事象 $E_1,E_2,\cdots$ が独立とは σ-alg. $\mathcal{E}_1,\mathcal{E}_2,\cdots$ が独立であること．ただし, ここで $\mathcal{E}_i$ とは,

$$\mathcal{E}_i=\{\emptyset,E_n,E_n^c,\Omega\}$$

注意 5.1. $\mathcal{E}_n=\sigma(I_{E_n})$ であることに注意しましょう．I_{E_n} とは E_n の定義関数です．定義関数は可測関数なので確率変数です.

5.2 π-system と独立 σ-alg.

独立性も π-system という簡単なものから考えることができます.

補題 5.1. $\mathcal{G},\mathcal{H}$ を σ-alg. $\mathcal{F}$ の部分 σ-alg. とする．$\mathcal{I},\mathcal{J}$ は π-system であり, $\sigma(\mathcal{I})=\mathcal{G},\sigma(\mathcal{J})=\mathcal{H}$ であるとする．このとき,

$$\mathcal{G},\mathcal{H}\text{ が独立.}\ \Leftrightarrow\mathcal{I},\mathcal{J}\text{ が独立.}$$

ただし,「$\mathcal{I},\mathcal{J}$ が独立」とは

$$\mathbb{P}(I\cap J)=\mathbb{P}(I)\mathbb{P}(J)\quad(I\in\mathcal{I},J\in\mathcal{J})$$

が成り立つこと.

証明. ($\Rightarrow$ の証明)： $I\in\mathcal{I},J\in\mathcal{J}$ を取って, $\mathbb{P}(I\cap J)=\mathbb{P}(I)\mathbb{P}(J)$ を示せばよいが, $\mathcal{I}\subset\mathcal{G},\mathcal{J}\subset\mathcal{H}$ であることに注意すると, 独立の定義から明らか.

($\Leftarrow$ の証明) $I\in\mathcal{I}$ を取る．可測空間 $(\Omega,\mathcal{H})$ 上の二つの写像 μ_{I1},μ_{I2} を次のように定める:

$$\mathcal{H}\ni H\mapsto\mu_{I1}(H):=\mathbb{P}(I\cap H)$$
$$\mathcal{H}\ni H\mapsto\mu_{I2}(H):=\mathbb{P}(I)\mathbb{P}(H)$$

このとき, μ_{I1}, μ_{I2} は共に測度である. このとき,

$$\mu_{I1}(\Omega) = \mathbb{P}(I \cap \Omega) = \mathbb{P}(I) = \mathbb{P}(I)\mathbb{P}(\Omega) = \mu_{I2}(\Omega)$$

そして, $\mathcal{I}, \mathcal{J}$ は独立であるから, μ_{I1}, μ_{I2} を $\mathcal{J}$ に制限した写像は等しくなる. すなわち,

$$\mu_{I1}|_{\mathcal{J}} = \mu_{I2}|_{\mathcal{J}}$$

したがって, 補題 3.2 より, μ_{I1}, μ_{I2} は $\sigma(\mathcal{J})$ 上, すなわち $\mathcal{H}$ 上で一致する.

次に $H \in \mathcal{H}$ をとり, 可測空間 $(\Omega, \mathcal{H})$ 上の二つの写像 μ_{H1}, μ_{H2} を考える. μ_{H1}, μ_{H2} は以下のとおりである：

$$\mathcal{G} \ni G \mapsto \mu_{H1}(G) := \mathbb{P}(G \cap H)$$
$$\mathcal{G} \ni G \mapsto \mu_{H2}(G) := \mathbb{P}(G)\mathbb{P}(H)$$

μ_{H1}, μ_{H2} は測度である. また, $\mu_{H1}(\Omega) = \mu_{H2}(\Omega) = \mathbb{P}(H)$ であり, $\mathcal{I}$ 上に制限すると, この二つの写像は一致する (このことは前半で示したこと) ので, 再び補題 3.2 より, $\mathcal{G}$ 上で $\mu_{H1} = \mu_{H2}$ である. したがって, $\mathcal{G}, \mathcal{H}$ は独立である. □

5.3 独立確率変数 X と π-system と $\sigma(X)$

次に, この σ-alg. による独立の定義が, 確率変数が独立であることを含意するのを確かめておきましょう.

補題 5.2. $(\Omega, \mathcal{F}, \mathbb{P})$ を確率空間とする. X をこの確率空間上の $\mathbb{R}$-値確率変数とする. $\pi(X) = \{\{X \leq a\} \mid a \in \mathbb{R}\}$ とすると, $\pi(X)$ は π-system であり, $\sigma(\pi(X)) = \sigma(X)$

証明. ($\pi(X)$ が π-system になること) 自明.

($\sigma(X) = X^{-1}(\mathcal{B}(\mathbb{R})) := \{X^{-1}(B) \mid B \in \mathcal{B}(\mathbb{R})\}$ であること) これを示す際に定義 4.2 を思い出す. さらに, Generated σ-alg. の最小性に注意しておこう.

【$\sigma(X) \subset X^{-1}(\mathcal{B}(\mathbb{R}))$ であること】$X^{-1}(\mathcal{B}(\mathbb{R}))$ は逆像の性質より σ-alg. であることが示される. 対応 X^{-1} は $\mathcal{B}(\mathbb{R})$ の元 B を $X^{-1}(\mathcal{B}(\mathbb{R}))$ の元 $X^{-1}(B)$ に対応付けるものであるとする, すなわち,

$$\text{if} \quad B \in \mathcal{B}(\mathbb{R}) \Rightarrow X^{-1}(B) = \{\omega \mid X(\omega) \in B\}$$

とすると, X^{-1} は $\mathcal{B}(\mathbb{R})$ から $X^{-1}(\mathcal{B}(\mathbb{R}))$ への写像として well-defined である. つまり,

$$X^{-1} : \mathcal{B}(\mathbb{R}) \to X^{-1}(\mathcal{B}(\mathbb{R}))$$

は写像の定義をみたす. $X^{-1}(\mathcal{B}(\mathbb{R}))$ が σ-alg. であることに注意すれば, 定義 4.2 より X は $X^{-1}(\mathcal{B}(\mathbb{R}))$ 上の可測関数である. よって, $\sigma(X)$ の最小性より,

$$\sigma(X) \subset X^{-1}(\mathcal{B}(\mathbb{R}))$$

がわかる.

【$\sigma(X) \supset X^{-1}(\mathcal{B}(\mathbb{R}))$ であること】写像 X^{-1} の終域を $\mathcal{A}$ とおく.

$$\mathcal{A} \subset X^{-1}(\mathcal{B}(\mathbb{R}))$$

と仮定する[*7]. X^{-1} は写像としては well-defined ではない. したがって, $X^{-1}(\mathcal{B}(\mathbb{R}))$ は X が可測となる最小の σ-alg. である. すなわち,

$$\sigma(X) \supset X^{-1}(\mathcal{B}(\mathbb{R}))$$

がわかる. 以上より, $\sigma(X) = X^{-1}(\mathcal{B}(\mathbb{R}))$ が示せた.

($\sigma(\pi(X)) = \sigma(X)$ であること) 簡単のために左辺, 右辺の集合がどのようなものであるかを明示しておく.

$$\sigma(\pi(X)) = \sigma(X^{-1}(\pi(\mathbb{R}))) = \sigma(\{X^{-1}(A) \mid A \in \pi(\mathbb{R})\})$$
$$\sigma(X) = X^{-1}(\mathcal{B}(\mathbb{R})) = \{X^{-1}(A) \mid A \in \mathcal{B}(\mathbb{R})\}$$

$X^{-1}(\mathcal{B}(\mathbb{R}))$ は σ-alg だから,

$$X^{-1}(\mathcal{B}(\mathbb{R})) = \sigma(X^{-1}(\mathcal{B}(\mathbb{R})))$$

である. 明らかに,

$$\mathcal{B}(\mathbb{R}) \supset \pi(\mathbb{R})$$

であるため,

$$X^{-1}(\mathcal{B}(\mathbb{R})) \supset X^{-1}(\pi(\mathbb{R}))$$

となるから,

$$\sigma(X) \supset \sigma(\pi(X))$$

である. よって, あとは $\sigma(X) \subset \sigma(\pi(X))$ を示せばよい. これを示すために, 左辺から任意に元を取って固定する. すなわち, $A \in \sigma(X)$ を取る. これはつまり,

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

ということである. 命題 3.1 より, A は $\pi(\mathbb{R})$ の元の高々可算個の集合の高々可算回の集合演算の結果である. 逆像は $\cup, \cap$ および, 補集合を取る操作を保存する, すなわち,

$$X^{-1}(\bigcap_\alpha A_\alpha) = \bigcap_\alpha X^{-1}(A_\alpha)$$
$$X^{-1}(\bigcup_\alpha A_\alpha) = \bigcup_\alpha X^{-1}(A_\alpha)$$
$$X^{-1}(A^c) = (X^{-1}(A))^c$$

が成り立つ. したがって,

$$A \in \sigma(\{X^{-1}(A) \mid A \in \pi(\mathbb{R})\})$$

が分かり, 証明が終わる. □

$(\Omega, \mathcal{F}, \mathbb{P})$ を確率空間, X, Y を確率変数, $x, y \in \mathbb{R}$ として, 以下が成り立っているとする.

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

これは π-system である $\pi(X), \pi(Y)$ が独立であることを言っている. 補題 5.1 より, $\sigma(\pi(X)), \sigma(\pi(Y))$ が独立であることが分かり, 補題 5.2 より $\sigma(X), \sigma(Y)$ が独立であることが分かる.

[*7] ここでは真部分集合であることを仮定している.

6　コロモゴロフの 0-1 法則

コロモゴロフの 0-1 法則で重要になってくるのは独立性です．具体的な話をする前に天下り的ではありますが用語の定義をしましょう．

6.1　Tail σ-alg.；末尾 σ 加法族

定義 6.1. $X_1, X_2, \cdots$ を確率変数とする．

$$\mathcal{T}_n := \sigma(X_{n+1}, X_{n+2}, \cdots)$$
$$\mathcal{T} := \bigcap_n \mathcal{T}_n$$

このとき, $\mathcal{T}$ は σ-alg. であり (手を動かしてみると明らか), 確率変数列 $(X_n)_{n\in\mathbb{N}}$ の末尾 σ-alg. もしくは, tail σ-alg. と呼ばれる．

この σ-alg. は重要な集合が含まれます．

命題 6.1. $X_1, X_2, \cdots$ を確率変数列, $\mathcal{T}$ をこれに関する tail σ-alg. とする．このとき, 次の F_1, F_2, F_3 は $\mathcal{T}$ の元である．

$$F_1 := \{\lim_{k\to\infty} X_k \text{が存在}\} = \{\omega \mid \lim_{k\to\infty} X_k(\omega) \text{ が存在}\}$$
$$F_2 := \{\sum_k X_k \text{が収束}\} = \{\omega \mid \sum_k X_k(\omega) \text{ が収束}\}$$
$$F_3 := \{\lim_{k\to\infty} \frac{X_1 + \cdots X_k}{k} \text{が収束}\}$$

証明. (F_1 について) $n \in \mathbb{N}$ を固定する．$\mathcal{T}_n$ において, $X_{n+1}, X_{n+2}, \cdots$ は可測なので極限関数は存在すれば可測となる．したがって, $\{\lim_{k\to\infty} X_k \text{が存在}\} \in \mathcal{T}_n$ である．n は任意より, $F_1 \in \mathcal{T}$.

(F_2 について) $n \in \mathbb{N}$ を固定．このとき,

$$\sum_{k=1}^{\infty} X_k = \sum_{k=1}^{n} X_k + \sum_{k=n+1}^{\infty} X_{k+1}$$

とできる．

$$c_n = \sum_{k=1}^{n} X_k$$

とおく．関数列 $(Y_m^{(n)})$ を

$$Y_m^{(n)} = c_n + \sum_{k=1}^{m} X_{n+k}$$

を考える．$\mathcal{T}_n$ の上では $X_{n+1}, X_{n+2}, \cdots$ は可測関数なので, 関数列 $(Y_m^{(n)})$ は $\mathcal{T}_n$ 上の可測関数の列である．可測関数列の極限関数が存在する点の集合は $\mathcal{T}_n$ に属するから (F_1 のこと), これはすなわち, $F_2 \in \mathcal{T}_n$. n は任意なので $F_2 \in \mathcal{T}$

(F_3 について)

$$\lim_{k\to\infty}\frac{X_1+\cdots+X_k}{k}=\lim_{k\to\infty}\frac{1}{k}\sum_{l=1}^{k}X_l$$

より,

$$Z_k=\frac{1}{k}\sum_{l=1}^{k}X_l$$

とおけば, (Z_k) は可測関数列. F_3 はすなわち, (Z_k) が収束することであり, これは $F_1\in\mathcal{T}$ であることから, $F_3\in\mathcal{T}$ が分かる. □

6.2 コルモゴロフの 0-1 法則

さて, いよいよコロモゴロフの 0-1 法則を見ていこう.

定理 6.2. $(X_n)_{n\in\mathbb{N}}$ を独立確率変数列とする. $\mathcal{T}$ を $(X_n)_{n\in\mathbb{N}}$ の tail σ-alg. とする. このとき, $\mathcal{T}$ は $\mathbb{P}$-trivial である. すなわち次の二つの条件が成立する.

1. $F\in\mathcal{T}\Rightarrow\mathbb{P}(F)=0\quad\text{or}\quad\mathbb{P}(F)=1$
2. ξ が $\mathcal{T}$ 可測確率変数なら, ξ はほぼ決定的である. すなわち, ある $c\in\mathbb{R}$ があって, $\mathbb{P}(\xi=c)=1$ である. $c=\pm\infty$ でもよい.

証明. (1. について) $\mathcal{X}_n:=\sigma(X_1,X_2,\cdots,X_n),\mathcal{T}_n:=\sigma(X_{n+1},X_{n+2},\cdots)$ とおく. 証明を 4 つのステップに分けて進める.

($step$1.)

主張 6.1. $\mathcal{X}_n,\mathcal{T}_n$ は独立.

$\mathcal{K}=\{\{X_1\le x_1,\cdots,X_n\le x_n\}\mid x_1,\cdots,x_n\in\mathbb{R}\cup\{\infty\}\}$ とする. $\mathcal{K}$ は π-system である[*8].

$\mathcal{K}':=\{\{X_1\le x_1,X_2\le x_2\}\mid x_1,x_2\in\mathbb{R}\}$ とする. $\sigma(\mathcal{K}')=\sigma(X_1,X_2)$ を示そう. そうすれば $\sigma(\mathcal{K})=\mathcal{X}_n$ は同様に証明できる. $\sigma(\mathcal{K}')$ 上で X_1,X_2 が可測であることは定義より明らか. したがって,

$$\sigma(X_1,X_2)\subset\sigma(\mathcal{K}')$$

である. 逆に, $A\in\mathcal{K}$ をとれば $\mathcal{K}$ の定義より

$$A=\{X_1\le x_1\}\cap\{X_2\le x_2\}\quad(\exists x_1,\exists x_2\in\mathbb{R}\cup\{\infty\})$$

と表わせる. 当然 X_1,X_2 が $\sigma(X_1,X_2)$ 上可測関数であることから,

$$\{X_1\le x_1\},\{X_2\le x_2\}\ \in\sigma(X_1,X_2)$$

である. σ-alg. であることから

$$A=\{X_1\le x_1\}\cap\{X_2\le x_2\}\in\sigma(X_1,X_2)$$

[*8] 定義を確かめるとそのまますぐわかる

よって,

$$\sigma(\mathcal{K}') \subset \sigma(X_1, X_2)$$

以上より,

$$\sigma(\mathcal{K}') = \sigma(X_1, X_2)$$

が分かる.

$\mathcal{J} := \{\{X_{n+1} \leq x_{n+1}, \cdots, X_{n+r} \leq x_{n+r}\} \mid r \in \mathbb{N}, x_{n+1}, \cdots, x_{n+r} \in \mathbb{R} \cup \{\infty\}\}$ とする. この時, 次が成立する.

$$\sigma(\mathcal{J}) = \mathcal{T}_n = \sigma(X_{n+1}, X_{n+2}, \cdots)$$

$X_j\ (j \geq n+1)$ を任意に選ぶと, X_j は $\sigma(\mathcal{J})$ 上で可測なので,

$$\sigma(X_{n+1}, X_{n+2}, \cdots) \subset \sigma(\mathcal{J})$$

である. 逆に $A \in \mathcal{J}$ を取ると, $r \in \mathbb{N}$ と $x_{n+1}, \cdots, x_{n+r} \in \mathbb{R} \cup \{\infty\}$ が存在して,

$$A = \{X_{n+1} \leq x_{n+1}\} \cap \cdots \cap \{X_{n+r} \leq x_{n+r}\}$$

と表わせるから, 先ほどと同様の議論により,

$$A \in \sigma(X_{n+1}, X_{n+2}, \cdots)$$

がわかる. よって,

$$\sigma(\mathcal{J}) \subset \sigma(X_{n+1}, X_{n+2}, \cdots)$$

また, 当然ながら $\mathcal{J}$ は π-system である.

ここまでで次のようなことが分かった.

$$\sigma(\mathcal{K}) = \mathcal{X}_n$$
$$\sigma(\mathcal{J}) = \mathcal{T}_n$$
$$\mathcal{K}, \mathcal{J} \text{は} \pi\text{-system}$$

仮定より, (X_n) が独立確率変数列であることから二つの π-system $\mathcal{K}, \mathcal{J}$ は独立であることが分かる. したがって補題 5.2 を適用することによって終了する.

($step$2.)

主張 6.2. $\mathcal{X}_n$ と $\mathcal{T}$ は独立.

$step$1 で $\mathcal{X}_n$ と $\mathcal{T}_n$ が独立なことは示している. $\mathcal{T} \subset \mathcal{T}_n$ より,

$$F \in \mathcal{T} \Rightarrow F \in \mathcal{T}_n$$

であることに注意すれば, $\mathcal{X}_n$ と $\mathcal{T}$ が独立であることは定義通りに確かめることができる.

($step$3.)

主張 6.3. $\mathcal{X}_\infty := \sigma(X_n \mid x \in \mathbb{N})$ と $\mathcal{T}$ は独立.

$\mathcal{X}_n \subset \mathcal{X}_{n+1}$ より, $\mathcal{K}_\infty := \bigcup_n \mathcal{X}_n$ は π-system である. 実際, $A, B \in \mathcal{K}_\infty$ を取ると,

$$A \in \mathcal{X}_n, B \in \mathcal{X}_m$$

なる n, m が存在するが $n \leq m$ とすると,

$$A, B \in \mathcal{X}_m$$

が成立する. $\mathcal{X}_m$ は σ-alg. だから,

$$A \cap B \in \mathcal{X}_m$$

次が成立.

$$\sigma(\mathcal{K}_\infty) = \mathcal{X}_\infty$$

実際, $n \in \mathbb{N}, x_n \in \mathbb{R} \cup \{\infty\}$ に対して,

$$\{X_n \le x_n\} \in \mathcal{K}_\infty$$

であるから, X_n は $\sigma(\mathcal{K}_\infty)$ 上可測である. したがって,

$$\mathcal{X}_\infty \subset \sigma(\mathcal{K}_\infty)$$

が成立する. また, $A \in \mathcal{K}_\infty$ とすると, ある $n \in \mathbb{N}$ があって

$$A \in \mathcal{X}_n = \sigma(X_1, \cdots, X_n) \subset \mathcal{X}_\infty$$

より,

$$\mathcal{K}_\infty \subset \mathcal{X}_\infty$$

がわかるので,

$$\sigma(\mathcal{K}_\infty) \subset \mathcal{X}_\infty$$

$\mathcal{K}_\infty$ と $\mathcal{T}$ は $step2$ より独立であるから, 補題 5.2 より主張が確かめられる.

($step4.$) $\mathcal{T} \subset \mathcal{X}_\infty$ より, $\mathcal{T}$ は自分自身と独立である. したがって, $F \in \mathcal{T}$ とすると,

$$\mathbb{P}(F) = \mathbb{P}(F \cap F) = \mathbb{P}(F)\mathbb{P}(F)$$

ゆえに,

$$\mathbb{P}(F) = 0 \quad \text{or} \quad \mathbb{P}(F) = 1$$

がしたがう.

(2. について[*9]) 1. より, $x \in \mathbb{R}$ に対して

$$\mathbb{P}(\xi \le x) = 0 \, \text{or} \, 1$$

である. $c = \sup\{x \mid \mathbb{P}(\xi \le x) = 0\}$ とする.

$c = \pm\infty$ のとき, $\mathbb{P}(\xi = \pm\infty) = 1$ なので (復号同順[*10]), この場合は OK.

c が有限のとき, 任意の $n \in \mathbb{N}$ に対して $\mathbb{P}(\xi \le c - \frac{1}{n}) = 0$ より,

$$\mathbb{P}\left(\bigcup_n \{\xi \le c - \frac{1}{n}\}\right) = \mathbb{P}(\xi < c) = 0$$

一方, 任意の $n \in \mathbb{N}$ に対して $\mathbb{P}(\xi \le c + \frac{1}{n}) = 1$ より,

$$\mathbb{P}\left(\bigcap_n \{\xi \le c + \frac{1}{n}\}\right) = \mathbb{P}(\xi \le c) = 1$$

である. したがって,

$$\mathbb{P}(\xi = c) = \mathbb{P}(\{\xi \le c\} - \{\xi < c\}) = 1 - 0 = 1$$

□

[*9] ここまで来たら消化試合

[*10] 別名書くのが面倒臭い

7　あとがき

この定理の不思議なところは, 確率が 1 もしくは 0 になることが分かっているにもかかわらず, 具体的に一つ一つ事象の確率を求めようとすると 1 なのか 0 なのかわからないというところです. 不思議ちゃんもええ加減にせえよってなります.

この定理の証明はマルチンゲールを用いるとかなり簡潔になるらしいです. 実際舟木 [2] では証明に 1 ページくらいしか紙面を割いていませんでした. ぶっちゃけまだそこまで勉強できていないのでその証明は省かせていただきます. ただ, この文章の LaTeX ファイルは手元にあるので, そこまで勉強が終わったのちに知らぬ間に更新しておくかもしれません.

また, この記事の「準備」の章は, 今の僕の確率論への理解を, 他人に分かりやすくなるように自分の言葉で書き表したものです*11. 一回生の後期には集合論の知識を持っている方もいるかと思われますので, そういう方の目に入って現代確率論に興味を持っていただければなと思う次第です.

参考文献

[1] David Williams *Probability with Martingales*

[2] 舟木直久『確率論』

*11 つまりめっちゃ頑張って書いたところ