みんなで確率

梅崎直也@unaoya

2020 年 5 月 9 日

UTokyoOCW にある楠岡先生の講義動画 (https://ocwx.ocw.u-tokyo.ac.jp/course_11395/) を参考にまとめたノートです。この講義を見ながら一緒に勉強するという配信をこちらの Youtube チャンネル (https://www.youtube.com/channel/UCtP2OI-4D_AHehTHSILwiSA?view_as=subscriber) で行なっています。途中からでも参加いただけますので、ご興味お持ちいただければぜひご覧ください。ここまでのアーカイブはこちらの再生リスト (https://www.youtube.com/playlist?list=PL-2TwNpShL3yvJqMmLgrE7kvv9Ripk_Lb) にまとめてあります。

文章中の内容について、動画で説明している部分にはリンクを貼っていますので、合わせてご覧いただけますと幸いです。

1 集合の言葉について

このノートを読む上で最低限必要な集合についての言葉をまとめました。冪集合と逆像について書いてあります。

1.1 冪集合

与えられた集合 Ω に対し、その部分集合を全て集めるとまた集合となる。これを Ω の冪集合といい、$P(\Omega), 2^\Omega$ などとあらわす。冪集合は $P(\Omega)$ と書くこともあるが、確率の P と紛らわしいのでこのノートでは 2^Ω で表すことにする。

$A \subset \Omega$ であるとき $A \in 2^\Omega$ であり、また $A \in 2^\Omega$ であるとき $A \subset \Omega$ である。集合を要素に持つ集合になるのに慣れておこう。

例 1.1. $\Omega = \{0\}$ のとき、

$$2^\Omega = \{\emptyset, \Omega\}$$

である。冪集合 2^Ω は要素を $2^1 = 2$ 個持つ集合。

例 1.2. $\Omega = \{0, 1\}$ のとき、

$$2^\Omega = \{\emptyset, \{0\}, \{1\}, \Omega\}$$

である。冪集合 2^Ω は要素を $2^2 = 4$ 個持つ集合。

例 1.3. $\Omega = \{0,1,2\}$ のとき、

$$2^\Omega = \{\emptyset, \{0\}, \{1\}, \{2\}, \{0,1\}, \{0,2\}, \{1,2\}, \Omega\}$$

である。冪集合 2^Ω は要素を $2^3 = 8$ 個持つ集合。

上で見たことから想像できるように、Ω の要素が有限で n 個であるとき、2^Ω は要素を 2^n 個持つ有限集合である。

1.2 逆像

集合 X, Y の間の写像 $f : X \to Y$ による Y の部分集合 B の逆像を定義する。これは A の部分集合で次のように定まる。

定義 1.4. $f : X \to Y$ を写像とし、$B \subset Y$ とする。このとき、B の f による**逆像**とは次で定まる A の部分集合のことを言う。

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\}$$

f^{-1} という記号について一つ注意しておく。f についてその逆写像が存在するとき、それを f^{-1} と表記する。この逆写像を表す f^{-1} と逆像を表す f^{-1} は別のものである。これらは、逆写像であれば $f^{-1}(y)$ のように $y \in Y$ の要素に対して定まるものであるのに対し、逆像は $f^{-1}(B)$ のように部分集合 $B \subset Y$、あるいは同じことだが $B \in 2^Y$ に対して定まるものであるということで区別がつく。数学では文字や変数がどこの集合の要素であるかを常に注意する必要がある。ここでは $y \in Y$ に対する $f^{-1}(y)$ であるのか、$B \in 2^Y$ に対する $f^{-1}(B)$ なのかをしっかりと見極めよう。

例 1.5. $X = \{a,b,c\}, Y = \{0,1,2\}$ とし、$f : X \to Y$ を

$$\begin{cases} f(a) = 0 \\ f(b) = 1 \\ f(c) = 1 \end{cases}$$

と定める。このとき、

$$\begin{gathered} f^{-1}(\{0,2\}) = \{a\} \\ f^{-1}(\{2\}) = \emptyset \\ f^{-1}(\{0,1\}) = X \end{gathered}$$

となる。

例 1.6. $\mathbb{R}$ を実数全体の集合とし、$f : \mathbb{R} \to \mathbb{R}$ を $f(x) = x^2$ により定める。$\mathbb{R}$ の部分集合である閉区間を次の記号で表す。

$$[a,b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$$

このとき、

$$
\begin{aligned}
f^{-1}([1,2]) &= [-\sqrt{2},-1] \cup [1,\sqrt{2}] \\
f^{-1}([0,1]) &= [-1,1] \\
f^{-1}(\{x \in \mathbb{R} \mid 0 < x\}) &= \mathbb{R} \setminus \{0\}
\end{aligned}
$$

となる。

定義や上の例からわかるように、逆像により写像 $f^{-1} : 2^Y \to 2^X$ が定まる。改めて注意するが、これは逆写像 $f^{-1} : Y \to X$ とは**異なるものである**。

逆像についてはこの辺り `https://youtu.be/jtXipcBYQN0?t=481` を参照してください。

1.3 逆像の性質

テキストの演習問題にあったが、逆像という操作は集合に対する演算との相性がよい。

命題 1.7. 写像 $f : A \to B$ と $C, D \subset B$ に対して次が成り立つ。

1. $f^{-1}(B) = A$
2. $f^{-1}(\emptyset) = \emptyset$
3. $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$
4. $f^{-1}(C \cup D) = f^{-1}(C) \cup f^{-1}(D)$
5. $f^{-1}(C \setminus D) = f^{-1}(C) \setminus f^{-1}(D)$

証明についてはこちらの動画 `https://youtu.be/a_6uwLXVelc?t=997` を参照してください。

2 確率の定義

2.1 確率空間

定義 2.1 (確率空間)**.** 有限集合 Ω と $F \subset 2^\Omega$ がある条件を満たし[*1]、さらに写像 $P : F \to [0,1] \subset \mathbb{R}$ を合わせた三つ組 (Ω, F, P) が以下を満たすとき、この (Ω, F, P) を**確率空間**という。

1. $P(\Omega) = 1$
2. $P(\emptyset) = 0$
3. $A, B \in F$ に対し $A \cap B = \emptyset$ なら $P(A \cup B) = P(A) + P(B)$

Ω が有限集合でなくても確率空間を定義することはできるが少々ややこしくなるので、この講義ではしばらく Ω を有限集合とし、$F = P(\Omega)$ とする。

例 2.2. Ω をトランプのカード 52 枚の集合としよう。この集合を記号 $S = \{\clubsuit, \diamondsuit, \heartsuit, \spadesuit\}$ と数 $N = \{1,2,3,4,5,6,7,8,9,10,11,12,13\}$ の直積集合として表現しておく。つまり $\Omega = S \times N$ であり、Ω の要素 $\omega \in \Omega$ は $\omega = (\heartsuit, 3), (\clubsuit, 12)$ などと表示される。

[*1] とりあえずは $F = 2^\Omega$ でやるのでこの条件については気にしないでよい

$F = 2^\Omega$ は例えばハート全体の集合 $\{(\heartsuit, 1), (\heartsuit, 2), \ldots, (\heartsuit, 13)\}$ とか、3 全体の集合 $\{(\clubsuit, 3), (\diamondsuit, 3), (\heartsuit, 3), (\spadesuit, 3)\}$ とか、適当に選んだ集合 $\{(\diamondsuit, 2), (\heartsuit, 5), (\clubsuit, 9), (\clubsuit, 12)\}$ などを要素に持つ集合[*2]。

$P : F \to \mathbb{R}$ を $P(A) = \dfrac{|A|}{52}$ で定める。ここで $|A|$ を A の要素の個数とする。つまり、どのカードを引く確率も同様に確からしいとする。

これは確率空間になる。定義を確かめよう。まず $P(\Omega) = \dfrac{52}{52} = 1, P(\Omega) = \dfrac{0}{52} = 0$ である。また、$A \cap B = \emptyset$ なとき、

$$|A \cup B| = |A| + |B| - |A \cap B| = |A| + |B|$$

なので $P(A \cup B) = P(A) + P(B)$ である。

確率空間という概念が導入された経緯については `https://youtu.be/4wj6eGz5E5s` をご覧ください。

2.2 確率変数

定義 2.3 (確率変数)**.** $(\Omega, F = 2^\Omega, P)$ を確率空間とする。写像 $X : \Omega \to \mathbb{R}$ を**確率変数**という。[*3]

例 2.4. 前の例と同様トランプ 52 枚の集合 Ω、$F = 2^\Omega$ と、同様に確からしい確率 P を用いて考える。数字、色、記号、偶奇によって次のような確率変数 X_n, X_c, X_s, X_p を定める。$X_n : \Omega \to \mathbb{R}$ を

$$X_n(s, n) = n$$

とする。$X_s : \Omega \to \mathbb{R}$ を

$$X_s(s, n) = \begin{cases} 1 & s = \clubsuit \\ 2 & s = \diamondsuit \\ 3 & s = \heartsuit \\ 4 & s = \spadesuit \end{cases}$$

とする。$X_c : \Omega \to \mathbb{R}$ を

$$X_c(s, n) = \begin{cases} 1 & s = \clubsuit, \spadesuit \\ 2 & s = \diamondsuit, \heartsuit \end{cases}$$

とする。$X_p : \Omega \to \mathbb{R}$ を

$$X_p(s, n) = (-1)^n$$

とする。[*4]

確率変数に対しては、次のようにして期待値という数を定義できる。

定義 2.5 (期待値)**.** 確率空間 (Ω, F, P) と確率変数 X に対し X の**期待値**を

$$E(X) = \sum_{\omega \in \Omega} P(\{\omega\}) X(\omega)$$

と定める。

[*2] これは「集合の集合」だが、そういうことは気にするときと気にしないときを明確にするのがよい
[*3] Ω が無限集合の場合にはより複雑な定義になる
[*4] 記号や色などを数字にするというのは、ゲームの点数をつけるようなものと思えばよい。

例 2.6. 上のトランプの確率空間 $(\Omega, 2^{\Omega}, P)$ と確率変数 X_n, X_s, X_c, X_p について、期待値は以下の通り。

$$
\begin{aligned}
E(X_n) &= \frac{4}{52}(1+2+3+4+5+6+7+8+9+10+11+12+13) = 7 \\
E(X_s) &= \frac{13}{52}(1+2+3+4) = \frac{5}{2} \\
E(X_c) &= \frac{26}{52}(1+2) = \frac{3}{2} \\
E(X_p) &= \frac{24}{52} - \frac{28}{52} = -\frac{1}{13}
\end{aligned}
$$

2.3 部分加法族

定義 2.7 (部分加法族)**.** (Ω, F, P) を確率空間とする。$G \subset F$ が**部分加法族**であるとは次を満たすこと。

1. $\emptyset \in G$
2. $A, B \in G$ ならば $A \cup B \in G$
3. $A \in G$ ならば $A^c = \Omega \setminus A \in G$

定義では $\cup, {}^c$ のみを条件にしているが、実は G の要素に対して他の操作 $\cap, \setminus$ をしてもまた G に属することを示すことができる。

命題 2.8. $G \subset F$ が部分加法族であるとする。このとき、

1. $\Omega \in G$
2. $A, B \in G$ ならば $A \cap B \in G$
3. $A, B \in G$ ならば $A \setminus B \in G$

証明はこちら `https://youtu.be/RR8Vh3qZD4A?t=2982` をご覧ください。

つまり、G の要素に対して部分集合に対する操作（これは基本的な論理操作とも言える）を行ったときにまた G の要素になるというのが部分加法族の定義である。集合演算を代数的操作と見て、sub algebra ともいう。

確率変数 $X : \Omega \to \mathbb{R}$ から $\sigma\{X\}$ という部分加法族を次のようにして定める。

定義 2.9 (確率変数から定まる部分加法族)**.** (Ω, F, P) を確率空間、X を確率変数とする。これに対し

$$\sigma\{X\} = \{X^{-1}(A) \in F | A \in P(\mathbb{R})\}$$

と定める。

命題 2.10. 上で定めた $\sigma\{X\} \subset F$ は部分加法族である。

これの証明はこちらの動画 `https://youtu.be/mgbaNADBj3M?t=1771` をご覧ください。

例 2.11. 上の例と同じくトランプの確率空間 $(\Omega = S \times N, 2^{\Omega}, P)$ を考える。確率変数 X_p に対して $\sigma\{X_p\}$ は

$$\sigma\{X_p\} = \{\emptyset, \{(s,n) \in \Omega \mid n \text{ が偶数 }\}, \{(s,n) \in \Omega \mid n \text{ が奇数 }\}, \Omega\}$$

の 4 つの要素からなる集合。

確率変数 X_s に対して $\sigma\{X_s\}$ は要素が 16 からなる集合で、例えば

$$\{(s,n)\in\Omega \mid s=\spadesuit\}\in\sigma\{X_s\}$$
$$\{(s,n)\in\Omega \mid s=\heartsuit\text{または}\diamondsuit\}\in\sigma\{X_s\}$$
$$\{(s,n)\in\Omega \mid s\neq\clubsuit\}\in\sigma\{X_s\}$$

などを要素に持つ。

これは X が与える情報と見ることができる。このことをより理解しやすくするために、次に部分加法族の原子と分割という概念を導入する。

定義 2.12 (部分加法族の原子). $G\subset F$ を部分加法族とする。$A\in G$ が G の**原子**であるとは、$A\neq\emptyset$ であり、$B\subset A$ ならば $B=A$ または $B=\emptyset$ であること。

自然数における素数の定義を思い出そう。p が素数であるとは、$p\neq 1$ であり、x が p を割り切るならば $x=p$ または $x=1$ であることである。これと似たような定義になっていることを確認しよう。

「原子」という言葉から想像できるように、これらが与えられた部分加法族の基本的な構成要素と言える。実際、部分加法族の他の要素は原子の和集合として表される。これは自然数が素数の積で表されるのと同様。

例 2.13. 前の例と同様にトランプの確率空間 $(\Omega=S\times N, 2^\Omega, P)$ と、確率変数 X_n, X_p, X_s, X_c を考える。

1. $\sigma\{X_n\}$ の原子は次の 13 個の集合。

$$\{(s,1)\mid s\in S\},\dots,\{(s,13)\mid s\in S\}$$

2. $\sigma\{X_p\}$ の原子は次の 2 個の集合。

$$\{(s,n)\mid s\in S, n\text{ は偶数 }\},\{(s,n)\mid s\in S, n\text{ は奇数 }\}$$

3. $\sigma\{X_s\}$ の原子は次の 4 個の集合。

$$\{(\heartsuit,n)\mid n\in N\},\{(\spadesuit,13)\mid n\in N\},\{(\diamondsuit,n)\mid n\in N\},\{(\clubsuit,13)\mid n\in N\}$$

4. $\sigma\{X_c\}$ の原子は次の 2 個の集合。

$$\{(s,n)\mid s=\heartsuit,\diamondsuit, n\in N\},\{(s,n)\mid s=\spadesuit,\clubsuit, n\in N\}$$

命題 2.14. 確率変数 $X:\Omega\to\mathbb{R}$ に対して、$\sigma\{X\}$ の原子はある $x\in\mathbb{R}$ に対して $X^{-1}(x)$ であって $\emptyset$ でないものである。

定義 2.15. 集合 Ω の**分割**とは F の部分集合 $\{B_1,\dots,B_n\}$ で

1. $i\neq j$ ならば $B_i\cap B_j$
2. $\bigcup_{i=1}^n B_i=\Omega$

を満たすもの。

分割と部分加法族には次のような関係がある。

命題 2.16. 部分加法族 $G \subset F$ に対し、その原子を集めると Ω の分割を与える。逆に Ω の分割から部分加法族を定める。

これらは互いに逆の対応になっている。

これについてはこちら `https://youtu.be/mgbaNADBj3M?t=84` をご覧ください。

つまり Ω が有限であれば、分割を与えることと部分加法族を与えることは等価。分割という概念の方がイメージはつきやすいが、部分加法族という概念の方が数学的な記述はスッキリする。例えば、G_1, G_2 という二つの部分加法族があったとき、$G_1 \subset G_2$ であることで G_2 の方がより豊富な情報を持つことを記述できる。分割で述べるのであれば G_2 で定まる分割の方が細かいということになるが、数学的にこれを述べるのはやや面倒。このあたりの事情はこちら `https://youtu.be/eyhPJATyN_c?t=2840` をご覧ください。

改めて $\sigma\{X\}$ が情報であるということについて考える。これの原子は $X^{-1}(x)$ であって $\emptyset$ でないもの全体である。X の値を知ることで Ω がどう分割できるかを捉えている。逆に、Ω を分割するとそれに対応する確率変数を定めることもできる。

2.4 G 可測

定義 2.17. (X, P, F) を確率空間、X を確率変数、G を部分加法族とする。X が G **可測**であるとは、$\sigma\{X\} \subset G$ であること。言い換えると $X^{-1}(A) \in G$ であること。

今 Ω が有限集合であるから、σX の原子について $X^{-1}(x) \in G$ としても同値である。

つまり、X により定まる分割が G により定まる分割より粗いということ。言い換えれば、G で与えられた情報は X の様子を全て捕まえている。

命題 2.18. 確率変数 X が G 可測であるとする。このとき、G の原子において X は定数である。

証明. B を G の原子とする。X の値が B の要素に対して 2 つ以上定まると仮定し x_1, x_2 をそのような値とする。つまり $X^{-1}(x_1) \cap B \neq \emptyset, X^{-1}(x_2) \cap B \neq \emptyset$ であるとする。これらはいずれも B とは一致しない。

このとき、X が G 可測であるから $X^{-1}(x_1), X^{-1}(x_2) \in G$ である。さらに $B \in G$ であることと G が部分加法族であることから、$X^{-1}(x_1) \cap B \in G, \emptyset \neq X^{-1}(x_1) \cap B \subset B$ であり $X^{-1} \cap B \neq B$ であるから B が原子であることに矛盾する。

よって X の値は B の上でただ一つである。 □

例 2.19. 前の例と同様にトランプの確率空間 $(\Omega = S \times N, 2^{\Omega}, P)$ と、確率変数 X_n, X_p, X_s, X_c を考える。

1. X_p は $\sigma\{X_n\}$ 可測である。実際、σX_p の原子は前に見たように

   $$\{(s,n) \mid s \in S, n \text{ は偶数 }\}, \{(s,n) \mid s \in S, n \text{ は奇数 }\}$$

   であるが、例えば

   $$\{(s,n) \mid s \in S, n \text{ は偶数 }\} = \bigcup_{n \in N, n \text{ は偶数}} \{(s,n) \mid s \in S\} \in \sigma\{X_n\}$$

   である。
2. X_c は $\sigma\{X_s\}$ 可測である。例えば

   $$\{(s,n) \mid s = \heartsuit, \diamondsuit, n \in N\} = \{(\heartsuit, n) \mid n \in N\} \cup \{(\diamondsuit, n) \mid n \in N\} \in \sigma\{X_s\}$$

である。

3. X_p は $\sigma\{X_s\}$ 可測ではない。もしそうなら X_p は $\sigma\{X_s\}$ の各原子において定数となるが、例えば $\{(\heartsuit, n) \mid n \in N\}$ において X_p は $1, -1$ 両方の値をとる。
4. X_c は $\sigma\{X_n\}$ 可測ではない。もしそうなら X_c は $\sigma\{X_n\}$ の各原子において定数となるが、例えば $\{(s, 1) \mid s \in S\}$ において X_c は $1, 2$ 両方の値をとる。

2.5 独立性

定義 2.20 (独立性)**.** (Ω, F, P) を確率空間とする。$n \geq 2$ とする。$A_1, \ldots, A_n$ をそれぞれ部分加法族、$X_1, \ldots, X_n$ をそれぞれ確率変数とする。

1. $A_1, \ldots, A_n$ が**独立**であるとは、任意の $B_1 \in A_1, \ldots, B_n \in A_n$ に対して
$$P(B_1 \cap \cdots \cap B_n) = P(B_1) \times \cdots \times P(B_n)$$
が成り立つこと。
2. $X_1, \ldots, X_n$ が**独立**であるとは、$\sigma\{X_1\}, \ldots, \sigma\{X_n\}$ が独立なこと。

例 2.21. 前の例と同様にトランプの確率空間 $(\Omega = S \times N, 2^\Omega, P)$ と、確率変数 X_n, X_p, X_s, X_c を考える。

1. X_s, X_n は独立である。例えば
$$P(X_s^{-1}(1) \cap X_n^{-1}(1)) = \frac{1}{52}$$
$$P(X_s^{-1}(1)) \times P(X_n^{-1}(1)) = \frac{13}{52} \times \frac{4}{52} = \frac{1}{52}$$
であり、他も同様である。
2. X_s, X_c は独立でない。例えば
$$P(X_s^{-1}(1) \cap X_c^{-1}(1)) = \frac{13}{52}$$
$$P(X_s^{-1}(1)) \times P(X_c^{-1}(1)) = \frac{13}{52} \times \frac{26}{52}$$
である。

命題 2.22. (Ω, F, P) を確率空間、X, Y を確率変数とする。X, Y が独立ならば
$$E[XY] = E[X]E[Y]$$
である。

より一般に、$X_1, \ldots, X_{n+m}$ が独立な確率変数とし、$f : \mathbb{R}^n \to \mathbb{R}, g : \mathbb{R}^m \to \mathbb{R}$ とする。このとき
$$E[f(X_1, \ldots, X_n)g(X_{n+1}, \ldots, X_{n+m})] = E[f(X_1, \ldots, X_n)]E[g(X_{n+1}, \ldots, X_{n+m})]$$
である。

証明はこちら `https://youtu.be/eF5C84fk0lM?t=1904` をご覧ください。