Probability with Martingales のギャップとかメモ

Twitter : @skbtkey

概要

タイトル通り。ネットの海からこれを見つけ出した方は参考にしていただけると嬉しい。David Williams 著 "Probability with Martingales"。

1 Chapter 6

1.1 65 ページ 9 行目

$c_1U_1 + c_2U_2 \sim c_1V_1 + c_2V_2$ について。

$$\begin{aligned} &\{x \mid (c_1U_1 + c_2U_2)(x) \neq (c_1V_1 + c_2V_2)(x)\} \\ \Leftrightarrow &\{x \mid (c_1U_1 + c_2U_2)(x) - (c_1V_1 + c_2V_2)(x) \neq 0\} =: A \end{aligned}$$

の測度が 0 であることを示せばよい。明らかに、

$$c_1U_1 \sim c_1V_1, c_2U_2 \sim c_2V_2$$

である。したがって、

$$\begin{aligned} A_1 &:= \{x \mid c_1U_1(x) \neq c_1V_1(x)\} \\ A_2 &:= \{x \mid c_2U_2(x) \neq c_2V_2(x)\} \end{aligned}$$

の測度はそれぞれ 0 である[*1]。このとき、次が成立。

$$A \subset A_1 \cup A_2$$

背理法で示す。上式の左辺から任意に x を取る。$x \notin A_1 \cup A_2$、つまり、

$$x \in A_1^c \cap A_2^c$$

と仮定する。このとき、$c_1U_1(x) = c_1V_1(x), c_2U_2(x) = c_2V_2(x)$ であるから、

$$(c_1U_1 + c_2U_2)(x) - (c_1V_1 + c_2V_2)(x) = 0$$

[*1] 別に A_1 などと名前を付けなくてもいいが (むしろ名前を付けない方がわかりやすい。)、紙面のスペースの都合上名前を付けている。

より、矛盾する。また、

$$A = \{x \mid (c_1U_1 + c_2U_2)(x) - (c_1V_1 + c_2V_2)(x) < 0\} \\ \cup \{x \mid (c_1U_1 + c_2U_2)(x) - (c_1V_1 + c_2V_2)(x) > 0\}$$

であるから、A は可測集合。したがって、測度の単調性より、A の測度は 0。

1.2　65 ページ 10 行目

$U_n \to U$ より、$N \in \mathbb{N}$ が存在して、$n \geq N \Rightarrow ||U_n - U|| < \varepsilon \quad (\forall \varepsilon > 0)$ である。$U_n \sim V_n, U \sim V$ より、$||V_N - V|| < \varepsilon$ が分かる。

1.3　67 ページ 10 行目から 12 行目

1.3.1　(ii) $\Rightarrow$ (i) について。

(逆向きは本の中で証明されています。)

任意の $Z \in \mathcal{K}$ を取ると、$||X - Z|| \geq ||X - Y|| + \alpha$ となることを示すとよい。ここで、$\alpha > 0$ である。しかし、Z として $Y + tZ \quad (t \in \mathbb{R})$ とすればよい。これは $\mathcal{L}^p$ の線形性のためである。なんとなれば、$Z = \frac{-Y+W}{t} \quad (\forall W \in \mathcal{K})$ とすればよい。$\langle X - Y, Z \rangle = \mathbf{E}[(X - Y)Z] = 0$ 即ち、$\mathbf{E}[XY] = \mathbf{E}[YZ]$ が成り立つことに注意して、$||X - Y - tZ||^2$ を計算すると分かる[*2]。

1.3.2　$||Y - Y'|| = 0$ について。

$$E[(X - Y)(Y - Y')] = 0$$
$$E[(X - Y')(Y - Y')] = 0$$

という式をそれぞれ期待値の中身を展開して引き算すると、

$$E[(Y - Y')^2] = 0$$

という式が求められるから $Y = Y' \quad (a.s.)$。

1.4　70 ページ 8,9 行目

1.4.1　$(\mathbf{P}(u))^q \leq \mathbf{P}(u^q)$ について。

記号が見慣れなくてよくわからないが、$\mathbf{P}$ は直前で確率測度であると示しているから、$\mathbf{P}(u)$ とは $\mathbf{P}$ による u の期待値である。凸関数 $c(x) = x^q$ を考えると、この不等式はイェンゼンの不等式を用いるとすぐに従うことが分かる。そのためには、イェンゼンの不等式を用いるための条件を満たしているかを確認しよう。u の定義は $f(s) = 0$ の部分で $u(s) = 0$ なので、下記の計算には本当は $1_{\{f(s)>0\}}$ を書いておくべき。

[*2] 2 乗しないと計算しにくい。

1.4.2　$\mathbf{P}(u^q) < \infty$ であること

Chapter 5 で見たように、f, h を可測関数、μ を測度とすれば、

$$(hf)\mu = h(f\mu)$$

が成立していた。これを用いる。また、$\frac{1}{p} + \frac{1}{q} = 1$ にも注意しておく。

$$\begin{aligned}
\mathbf{P}(u^q) &= \frac{h^q}{f^{q(p-1)}} \frac{f^p \mu}{\mu(f^p)} = \frac{h^q}{f^p} \frac{f^p \mu}{\mu(f^p)} = \frac{h^q}{f^p} f^p \frac{\mu}{\mu(f^p)} \\
&= h^q \frac{\mu}{\mu(f^p)} = \frac{h^q}{\mu(f^p)} \mu = \frac{1}{\mu(f^p)} \int h^q d\mu < \infty
\end{aligned}$$

1.4.3　$\mathbf{P}(u) < \infty$ について

ヤングの不等式を用いる。

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

$$\begin{aligned}
\mathbf{P}(u) = \frac{h}{f^{p-1}} \frac{f^p \mu}{\mu(f^p)} = hf \frac{\mu}{\mu(f^p)} &= \frac{hf}{\mu(f^p)} \mu \\
&\leq \frac{1}{\mu(f^p)} \left(\frac{f^p}{p} + \frac{h^q}{q} \right) \mu < \infty
\end{aligned}$$

1.4.4　ヘルダーの不等式の証明

$(\mathbf{P}(u))^q \leq \mathbf{P}(u^q)$ を展開する。ここでも $\frac{1}{p} + \frac{1}{q} = 1$ を適宜変形して利用する。

$$\begin{aligned}
\left(\frac{h}{\mu(f^{(p-1)}} 1_{\{f(s)>0\}} \frac{f^p \mu}{\mu(f^p)} \right)^q &\leq \frac{h^q}{f^{q(p-1)}} 1_{\{f(s)>0\}} \frac{f^p}{\mu(f^p)} \\
\left(\frac{1}{\mu(f^p)} \right)^q (hf\mu)^q &\leq \left(\frac{1}{\mu(f^p)} \right) \left(h^q \frac{f^p}{f^{q(p-1)}} 1_{\{f(s)>0\}} \mu \right) \\
(hf\mu)^q &\leq (\mu(f^p))^{q-1} \left(h^q \frac{f^p}{f^p} 1_{\{f(s)>0\}} \mu \right) \\
(hf\mu)^q &\leq (\mu(f^p))^{\frac{q}{p}} (h^q \mu) \\
(hf\mu) &\leq (\mu(f^p))^{\frac{1}{p}} (h^q \mu)^{\frac{1}{q}} \\
\int fh d\mu &\leq \left(\int f^p d\mu \right)^{\frac{1}{p}} \left(\int h^q d\mu \right)^{\frac{1}{q}}
\end{aligned}$$

通常の log を使う証明じゃないの面白いね。

2　Chapter 7

2.1　大数の法則を証明する前の補題

命題. S_k-値確率変数列 $(X_k)_{k=1,2,\cdots,n}$ は独立で、$g_k\colon S_k \to S'_k$ は $\mathcal{S}_k$-可測関数とする。このとき、$Y_k = g_k(X_k)$ とおけば、$(Y_k)_k$ は独立。

証明. $A'_k \in \mathcal{S}'_k$ とする。

$$\{Y_k \in A'_k\} = \{X_k \in g_k^{-1}(A'_k)\}$$

であり、g_k の可測性から、

$$g_k^{-1}(A'_k) \in \mathcal{S}_k$$

であることから、

$$\{X_k \in g_k^{-1}(A'_k)\} \in \sigma(X_k)$$

よって、

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{k=1}^{n}(Y_k \in A'_k)\right) &= \mathbb{P}\left((\bigcap_{k=1}^{n}(X_k \in g_k^{-1}(A'_k))\right) \\
&= \prod_{k=1}^{n}\mathbb{P}(X_k \in g_k^{-1}(A'_k) \\
&= \prod_{k=1}^{n}\mathbb{P}(Y_k \in A'_k)
\end{aligned}$$

□

系. $(X_k)_{k=1,2,\cdots,n}$ が $\mathbb{R}$-値独立確率変数の列とする。$g = g(x_1, x_2, \cdots, x_i)$ は $\mathbb{R}^i \to \mathbb{R}$ のボレル可測関数とする。ただし、$i < n$。このとき、$Y := g(X_1, X_2, \cdots, X_i)$ とおけば、$Y, X_{i+1}, \cdots, X_n$ は独立。

証明. $X = (X_1, \cdots, X_i)$ は $\mathbb{R}^i$-値確率変数であり、$X, X_{i+1}, \cdots, X_n$ は独立である。まずはこれを示す。$A \in \mathcal{B}(\mathbb{R}^i), A_{i+1}, \cdots, A_n \in \mathcal{B}(\mathbb{R})$ について、

$$\mathbb{P}(X \in A, X_{i+1} \in A_{i+1}, \cdots, X_n \in A_n) = \mathbb{P}(X \in A)\prod_{j=i+1}^{n}\mathbb{P}(X_j \in A_j) \tag{1}$$

が成立することを言えばよい。そこで、

$$\mathcal{L} = \{A \subset \mathbb{R}^i \mid (1) \text{ が成立 }\}$$

と定義する。$\mathcal{L}$ は λ-system である。明らかに $\Omega \in \mathcal{L}$ である。次に、$A, B \in \mathcal{L}, A \subset B$ とすると、

$\mathbb{P}(X \in B - A) = \mathbb{P}(X \in B) - \mathbb{P}(X \in A)$ であることに注意すれば、

$$
\begin{aligned}
&\mathbb{P}(X \in B - A, X_{i+1} \in A_{i+1}, \cdots, X_n \in A_n) \\
=&\mathbb{P}(X \in B, X_{i+1} \in A_{i+1}, \cdots, X_n \in A_n) - \mathbb{P}(X \in A, X_{i+1} \in A_{i+1}, \cdots, X_n \in A_n) \\
=&\mathbb{P}(X \in B) \prod_{j=i+1}^{n} \mathbb{P}(X_j \in A_j) - \mathbb{P}(X \in A) \prod_{j=i+1}^{n} \mathbb{P}(X_j \in A_j) \\
=&\left(\mathbb{P}(X \in B) - \mathbb{P}(X \in A)\right) \prod_{j=i+1}^{n} \mathbb{P}(X_j \in A_j) \\
=&\mathbb{P}(X \in B - A) \prod_{j=i+1}^{n} \mathbb{P}(X_j \in A_j)
\end{aligned}
$$

なので、$B - A \in \mathcal{L}$。最後に、$A_n \in \mathcal{L} \quad (n = 1, 2, \cdots)$ で、$A_n \uparrow A$ とする。$X \in A$ とすれば、$\exists m \in \mathbb{N} \quad s.t. \quad X \in A_m$ なので、そのような自然数で最小のものを m とする[*3]。集合列 (A_n) は包含関係に関して単調増加であるため、すべての $N \geq m$ に対して、

$$\{X \in A_m\} = \{X \in A_N\}$$

が成立する。したがって当然

$$\{X \in A_m\} = \bigcup_{k=N}^{\infty} \{X \in A_k\} = \{X \in A\}$$

よって、

$$\mathbb{P}(X \in A_m) = \mathbb{P}(X \in A)$$

よって次のように計算ができる。

$$
\begin{aligned}
&\mathbb{P}(X \in A, X_{i+1} \in A_{i+1}, \cdots, X_n \in A_n) \\
=&\mathbb{P}(X \in A_m, X_{i+1} \in A_{i+1}, \cdots, X_n \in A_n) \\
=&\mathbb{P}(X \in A_m) \prod_{j=i+1}^{n} \mathbb{P}(X_j \in A_j) \\
=&\mathbb{P}(X \in A) \prod_{j=i+1}^{n} \mathbb{P}(X_j \in A_j)
\end{aligned}
$$

これにより、$A \in \mathcal{L}$ であることが分かる。以上より $\mathcal{L}$ は λ-system。

さて、$\mathcal{P} = \{A_1 \times \cdots \times A_i \mid A_k \in \mathcal{B}(\mathbb{R}), k = 1, 2, \cdots, i\}$ は π-system であり、$A \in \mathcal{P}$ とすれば、直積の定義より

$$X \in A \Leftrightarrow X_1 \in A_1, \cdots, X_i \in A_i$$

より、

$$\{X \in A\} = \{X_1 \in A_1\} \cap \cdots \cap \{X_i \in A_i\}$$

[*3] 自然数全体の部分集合は必ず最小限を持つ

仮定から $X_1, \cdots, X_n$ は独立なので

$$\begin{aligned}
&\mathbb{P}(X \in A, X_{i+1} \in A_{i+1}, \cdots, X_n \in A_n) \\
=&\prod_{k=1}^{n} \mathbb{P}(X_k) \\
=&\mathbb{P}(X_1 \in A_1, \cdots, X_i \in A_i) \prod_{k=i+1}^{n} \mathbb{P}(X_k \in A_k) \\
=&\mathbb{P}(X \in A) \prod_{k=i+1}^{n} \mathbb{P}(X_k)
\end{aligned}$$

と計算できるから、$A \in \mathcal{L}$ がわかる。したがって $\mathcal{P} \subset \mathcal{L}$ である。よって、π-λ 定理より、

$$\mathcal{L} \supset \sigma(\mathcal{P}) = \mathcal{B}(\mathbb{R}^i)$$

以上により、$X, X_{i+1}, \cdots, X_n$ は独立であることが分かった。

あとは先の命題を適用すればよい。 □

補題. X, Y, Z, W は独立な r.v. とする。このとき、$(X, Y^n), (XY^2, Z), (XY, ZW), (X^2, Y^2)$ はそれぞれ独立。

証明. 確率変数の独立性は合成によって保たれることに注意する。

(X, Y^n) に関しては、

$$\begin{aligned}
f &\colon x \mapsto x \\
g &\colon x \mapsto x^n
\end{aligned}$$

として、$X = f(X), Y^n = g(Y)$ と見る。

(XY^2, Z) に関しては、

$$\begin{aligned}
f &\colon (x, y) \mapsto xy^2 \\
g &\colon x \mapsto x
\end{aligned}$$

として、$XY^2 = f(X, Y), Z = g(Z)$ として見る。

(XY, ZW) に関しては、

$$f \colon (x, y) \mapsto xy$$

として $XY = f(X, Y)$ とみれば、XY, Z, W は独立。もう一度 $ZW = f(Z, W)$ とみれば、XY, ZW は独立。

(X^2, Y^2) に関しては、

$$f \colon x \mapsto x^2$$

として $X^2 = f(X), Y^2 = f(Y)$ として見る。 □

2.2　73 ページ下から 3 行目

まあ明らかなんだけど、さすがにそのまま引き算して証明できるほど明らかではない。
$1_{\{|X-\mu|>c\}}$ を考えよう。

$$\mathbb{P}(|X-\mu|>c)=\mathbf{E}[1_{\{|X-\mu|>c\}}]$$

であり、

$$1_{\{|X-\mu|>c\}}<\frac{(X-\mu)^2}{c^2}$$

が成立することに注意すると、両辺を積分して、

$$\begin{aligned}\int 1_{\{|X-\mu|>c\}}d\mathbb{P} &\leq \int \frac{(X-\mu)^2}{c^2}d\mathbb{P}\\ \mathbb{P}(|X-\mu|>c) &\leq \frac{1}{c^2}\mathbf{E}[(X-\mu)^2]\\ c^2\mathbb{P}(|X-\mu|>c) &\leq \mathbf{E}[(X-\mu)^2]=\mathrm{Var}(X)\end{aligned}$$

2.3　74 ページ (a) の 2 行上

f が有界なのは f が閉区間上の連続関数だから。位相空間でよくあるお話だけどただ単に「f は有界である」と書かれるのはちょっと...。

3　Chapter 8

眠いし疲れたので明日。なんか書かなきゃいけないことあったはずなんだけど、ちょっと忘れたよ。

4　Chapter 9

4.1　$E[X;G]$ という書き方について

Chapter 5 で実は定義されています。

$$E[X;G]:=\int X1_G d\mathbb{P}=\int_G X d\mathbb{P}$$

4.2　条件付き期待値の別の書き方について

積分の形での定義を、指示関数 (定義関数) を使って書き換えたものが次の書き方。

定義 4.1. Y が X の $\mathcal{G}$ での条件付き期待値であるとは、任意の $G \in \mathcal{G}$ に対して次が成立すること。

$$E[X1_G] = E[Y1_G]$$

が成り立つこと。

これは次のように書き換えられる。

定義 4.2. Y が X の $\mathcal{G}$ での条件付き期待値であるとは、任意の上に有界な非負の $\mathcal{G}$-可測関数 Z に対して次が成立すること。

$$E[XZ] = E[YZ]$$

が成り立つこと。

後者の定義から前者の定義が導かれるのは自明。前者の定義から後者の定義を導く。(Z_n) は Z に各点収束する単調増加な単関数の列とする。単関数は指示関数の有限個の和で表すことができるから、

$$E[XZ_n] = E[YZ_n]$$

が成り立つ。したがって、

$$\lim_{n\to\infty} E[XZ_n] = \lim_{n\to\infty} E[YZ_n]$$

が成り立つ。単調収束定理より、

$$E[\lim_{n\to\infty} XZ_n] = E[\lim_{n\to\infty} YZ_n]$$
$$E[XZ] = E[YZ]$$

がなりたつ。$(E[XZ_n], E[YZ_n]$ がそれぞれ単調増加な可測関数列であることは後述の補題 4.1 を参照)

4.3　86 ページ 3 行目

$$\{Y > \tilde{Y} + \frac{1}{n}\} \uparrow \{Y > \tilde{Y}\}$$

とは、

$$\bigcup_n \{Y > \tilde{Y} + \frac{1}{n}\} = \{Y > \tilde{Y}\}$$

よって、

$$\sum_n \mathbb{P}(Y > \tilde{Y} + \frac{1}{n}) \geq \mathbb{P}\left(\bigcup_n \{Y > \tilde{Y} + \frac{1}{n}\}\right) \geq \mathbb{P}(Y > \tilde{Y}) > 0$$

であることに注意する。

4.4 86 ページ下から 6 行目

almost surelyで Y_n が非負かつ、n に関して単調増加であることを示すことに気を付ける。

4.4.1 $0 \le Y_n$(a.s.) について

$F := \{Y_n < 0\}$ について考える。Y_n は $\mathcal{G}$ 可測なので、F は $\mathcal{G}$ 可測集合となる。Y_n が X_n の条件付き期待値であることと、非負の可測関数を確立測度で積分するとその値は非負になることに気をつければ、以下のように計算することができる。

$$\int_F Y_n d\mathbb{P} = \int_F X_n d\mathbb{P} \\ \ge 0$$

一方、

$$\int_F Y_n d\mathbb{P} = \int_\Omega Y_n 1_F d\mathbb{P}$$

だから、

$$\int_\Omega Y_n 1_F d\mathbb{P} \ge 0$$

となる。さて、

$$Y_n 1_F(\omega) \le 0 \quad (\forall \omega \in \Omega)$$

だから、

$$\int_\Omega Y_n 1_F d\mathbb{P} \le 0$$

よって、

$$\int_\Omega Y_n 1_F d\mathbb{P} = 0$$

これから、

$$\int_\Omega -Y_n 1_F d\mathbb{P} = 0$$

51 ページの LEMMA(b) を用いると

$$\mathbb{P}(F) = 0$$

が分かる。

4.4.2 $Y_n \uparrow$(a.s.) について

前小小節とほぼ同じ。$s < r$ として、$G := \{Y_s > Y_r\}$ 上で X_s, X_r をそれぞれ積分する。そしてそれらを引き算する。

4.5　86 ページ下から 3 行目 $Y_n \uparrow Y$(a.s) について

上極限の扱い分かってない人みたいになった。

各点収束することを示すから、$\omega \in \Omega$ を一つ固定する。

4.5.1　$\limsup_{n\to\infty} Y_n(\omega) = \infty$ のとき

$$\limsup_{n\to\infty} Y_n(\omega) = \inf_{n\geq 1} \sup_{\nu\geq n} Y_\nu(\omega)$$

である (定義)。見やすさのために、新たに

$$a_n = \sup_{\nu\geq n} Y_\nu(\omega)$$

とおく。すると、

$$\inf_{n\geq 1} a_n$$

となる。探索範囲が狭まっているので a_n は単調減少列であることに注意する。仮定より、

$$\inf_{n\geq 1} a_n = \infty$$

だから、

$$a_n = \infty \quad (\forall n)$$

つまり、とりわけ $n = 1$ に関して

$$\sup_{\nu\geq 1} Y_\nu(\omega) = \infty$$

これはつぎのように書ける。

$$\forall K > 0, \exists m \in \mathbb{N} \quad s.t. \quad Y_m(\omega) > K$$

これは、数列 $(Y_n(\omega))$ が正の無限大に発散することの定義そのものである。したがって、

$$\lim_{n\to\infty} Y_n(\omega) = \infty$$

4.5.2　$\limsup_{n\to\infty} Y_n(\omega) = \alpha < \infty$ のとき

定義より、

$$\limsup_{n\to\infty} Y_n(\omega) = \inf_{n\geq 1} \sup_{\nu\geq n} Y_\nu(\omega)$$

であるが、$Y_n(\omega)$ は n に関する単調増加列なので、各 b_n の値は全て等しい。ただし、

$$b_n = \sup_{\nu\geq n} Y_\nu(\omega)$$

したがって、

$$\inf_{n\geq 1}\sup_{\nu\geq n} Y_\nu(\omega) = \sup_{n\geq 1} Y_n(\omega)$$

つまり、

$$\sup_{n\geq 1} Y_n(\omega) = \alpha$$

であることが分かる。よって、$Y_n(\omega)$ は上限が α の単調増加な数列なので

$$\lim_{n\to\infty} Y_n(\omega) = \alpha$$

4.6　89 ページ (f),(g) の証明について

4.6.1　(f) について

次の補題が必要。

補題 4.1. $0 \leq X \leq Y$ なる確率変数について、

$$E[X \mid \mathcal{G}] \leq E[Y \mid \mathcal{G}]$$

証明. $Z = E[X \mid \mathcal{G}, W = E[Y \mid \mathcal{G}]$ とする。$\{W < Z\}$ の確率が 0 であることを証明すればいい。 □