第 A 章　補遺

A.1　大数の強法則の証明

A.1.1　証明のための準備の補題

大数の強法則を証明するための Borel-Cantelli の補題を述べる.

補題 A.1 (Borel-Cantelli の補題)**.** $\sum_{n=1}^{\infty} \mathsf{Pr}(E_n) < \infty$ なる任意の事象列 $\{E_n\}_{n=1}^{\infty}$ に対して

$$\mathsf{Pr}\left(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} E_n\right) = \mathsf{Pr}(E_n,\ \text{i.o.}) = 0$$

となる.

Proof. $G_k := \bigcup_{n=k}^{\infty} E_n$, $G := \limsup_{n\to\infty} E_n = \bigcap_{k=1}^{\infty} G_k$ とおく. すると $G_k \subset G_{k+1}\,(k = 1, 2, \ldots)$ で $G = \bigcap_{k=1}^{\infty} G_k$ である. 任意の $k \in \mathbb{N}$ に対して

$$\begin{aligned}\mathsf{Pr}(G) &\le \mathsf{Pr}(G_k) \qquad (\because \text{補題 } 1.7(4))\\ &\le \sum_{n=k}^{\infty} \mathsf{Pr}(E_n). \qquad (\because \text{補題 } 1.7(7))\end{aligned}$$

となる. $\sum_{n=1}^{\infty} \mathsf{Pr}(E_n) < \infty$ なので, $\sum_{n=k}^{\infty} \mathsf{Pr}(E_n) \overset{k\to\infty}{\longrightarrow} 0$ となるので, $\mathsf{Pr}(G_k) \overset{k\to\infty}{\longrightarrow} 0$ である. ここで, 補題 1.7(6) を用いると

$$\mathsf{Pr}(G) = \mathsf{Pr}\left(\bigcap_{k=1}^{\infty} G_k\right) = \lim_{k\to\infty} \mathsf{Pr}(G_k) = 0$$

がわかる. □

補題 A.2 (Borel-Cantelli の逆)**.** $\sum_{n=1}^{\infty} \mathsf{Pr}(E_n) = \infty$ なる任意の独立な事象列 $\{E_n\}_{n=1}^{\infty}$ に対して

$$\mathsf{Pr}\left(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} E_n\right) = \mathsf{Pr}(E_n,\ \text{i.o.}) = 1$$

となる.

Proof. $p_n := \mathsf{Pr}(E_n)\,(n = 1, 2, \ldots)$ とおく．独立性から

$$\mathsf{Pr}\left(\bigcap_{n=k}^{\infty} E_n^c\right) = \prod_{n=k}^{\infty}(1-p_n) \qquad (\forall k \in \mathbb{N})$$

となる．これは, $r \geq n \geq k$ に対して

$$\mathsf{Pr}\left(\bigcap_{n=k}^{r} E_n^c\right) = \prod_{n=k}^{r}(1-p_n)$$

となる．実際,

$$\mathsf{Pr}\left(\bigcap_{n=k}^{\infty} E_n^c\right) = \lim_{r\to\infty}\mathsf{Pr}\left(\bigcap_{n=k}^{r} E_n^c\right) = \lim_{r\to\infty}\prod_{n=k}^{r}(1-p_n) = \prod_{n=k}^{\infty}(1-p_n)$$

がわかる．つぎに, $1-x \leq e^{-x}\,(x \geq 0)$ に注意すると, 任意の $k \in \mathbb{N}$ に対して

$$\lim_{r\to\infty}\prod_{n=k}^{r}(1-p_n) = \lim_{r\to\infty}\prod_{n=k}^{r} e^{-p_n} = \lim_{r\to\infty} e^{-\sum_{n=k}^{r} p_n} = e^{-\sum_{n=k}^{\infty} p_n} = 0$$

となる．よって

$$\mathsf{Pr}\left(\bigcap_{n=k}^{\infty} E_n^c\right) = 0 \tag{A.1}$$

がわかる．しかし

$$\left(\limsup E_n\right)^c = \left(\bigcap_{k=1}^{\infty}\bigcup_{n=k}^{\infty} E_n\right)^c = \bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} E_n^c$$

であることと補題 1.7(5)　に注意すると

$$\begin{aligned}\mathsf{Pr}\left(\left(\limsup E_n\right)^c\right) &= \mathsf{Pr}\left(\bigcup_{k=1}^{\infty}\bigcap_{n=k}^{\infty} E_n^c\right) = \lim_{k\to\infty}\mathsf{Pr}\left(\bigcap_{n=k}^{\infty} E_n^c\right) \\ &= 0 \qquad (\because \text{(A.1)})\end{aligned}$$

が示せた．

つぎに, 大数法則を示すための補題をふたつ述べる．

補題 A.3. $n \in \mathbb{N}$ に対して $Y_n = X_n \mathbb{1}\{|X_n| \leq n\}$ とし,

$$T_n := Y_1 + Y_2 + \cdots + Y_n \tag{A.2}$$

とする．このとき

$$\frac{T_n}{n} \overset{\text{a.s.}}{\to} \mu\,(n \to \infty) \;\Rightarrow\; \frac{S_n}{n} \overset{\text{a.s.}}{\to} \mu\,(n \to \infty)$$

が成立する．

Proof. Y_n の定義から

$$\sum_{n=1}^{\infty}\mathsf{Pr}\big(X_n \neq Y_n\big) = \sum_{n=1}^{\infty}\mathsf{Pr}\big(|X_n| > n\big) \leq \int_0^{\infty}\mathsf{Pr}\big(|X_1| > t\big)\,\mathrm{d}t = \mathsf{E}\big[|X_1|\big] < \infty$$

となる．よって Borel-Cantelli の第 1 の補題 (補題 A.1) から

$$\mathsf{Pr}\bigg(X_n \neq Y_n,\ \text{i.o.}\bigg) = 0$$

がわかる．すると $\omega \in \Omega$ に対して有限個の n に対して $X_n \neq Y_n$ なので

$$\lim_{n\to\infty}|S_n(\omega) - T_n(\omega)| = 0$$

となる．よって定理の主張は証明された． □

補題 A.4. 定理 4.19 と補題 A.3 の仮定のもと

$$\sum_{n=1}^{\infty}\frac{\mathsf{Var}\big(Y_n\big)}{n^2} \leq 4\mathsf{E}\big[|X_1|\big] < \infty$$

が成立する．

Proof. Fubini の定理を用いると

$$\begin{aligned}\mathsf{Var}\big[Y_n\big] \leq \mathsf{E}\big[Y_n^2\big] &= \int_0^{\infty} 2t\,\mathsf{Pr}\big(|Y_n| > t\big)\,\mathrm{d}t = \int_0^{n} 2t\,\mathsf{Pr}\big(|Y_n| > t\big)\,\mathrm{d}t \\ &= \int_0^{\infty}\mathbb{1}\{t < n\}2t\,\mathsf{Pr}\big(|Y_n| > t\big)\,\mathrm{d}t\end{aligned}$$

となる．よって

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{\mathsf{E}\big[Y_n^2\big]}{n^2} &\leq \sum_{n=1}^{\infty}\frac{1}{n^2}\int_0^{\infty}\mathbb{1}\{t < n\}2t\,\mathsf{Pr}\big(|Y_n| > t\big)\,\mathrm{d}t \\ &= \int_0^{\infty}\left\{\sum_{n=1}^{\infty}\frac{\mathbb{1}\{t < n\}}{n^2}\right\}2t\,\mathsf{Pr}\big(|Y_n| > t\big)\,\mathrm{d}t\end{aligned}$$

となる．最後の等号は, 各項が非負なので, Fubini の定理を用いることで和と積分の記号の交換が保証されることがわかる．よって後は $\forall t \geq 0$ に対して

$$2t\sum_{n>t}\frac{1}{n^2} \leq 4 \tag{A.3}$$

を示せばよい．$m \geq 2\,(m \in \mathbb{N})$ に対して

$$\sum_{n\geq m}\frac{1}{n^2} \leq \int_{m-1}^{\infty}\frac{\mathrm{d}x}{x^2} = \frac{1}{m-1}$$

となる. $t \geq 1$ のとき, (A.3) の和は $n = [t] + 1 \geq 2$ で始まるので

$$2t \sum_{n>t} \frac{1}{n^2} \leq \frac{2t}{[t]} \leq 4 \tag{A.4}$$

となる[1]. $0 \leq t \leq 1$ の場合

$$2t \sum_{n>t} \frac{1}{n^2} \leq 2\left(1 + \sum_{n=2}^{\infty} \frac{1}{n^2}\right) \leq 4 \tag{A.5}$$

となる. よって (A.4), (A.5) から (A.3) がわかる. □

A.1.2 大数の強法則 (定理 4.19) の証明

$n \in \mathbb{N}$ に対して

$$X_n^+ := \max\{X_n, 0\}, \qquad X_n^- := \max\{-X_n, 0\}$$

とおくと X_n^+ と X_n^- は定理 4.19 の仮定をみたし, $X_n = X_n^+ - X_n^-$ である. したがって一般性を失うことなく $X_n \geq 0\,(n \in \mathbb{N})$ と仮定してよい.

$\alpha > 1$ をとり

$$k(n) = [\alpha^n]$$

と定める. $\forall\epsilon > 0$ を取る. Chebyshev の不等式から

$$\begin{aligned}\sum_{n=1}^{\infty} \mathsf{Pr}\left(\left|T_{k(n)} - \mathsf{E}\left[T_{k(n)}\right]\right| > \epsilon k(n)\right) &\leq \frac{1}{\epsilon^2}\sum_{n=1}^{\infty} \frac{\mathsf{Var}\left[T_{k(n)}\right]}{\{k(n)\}^2} \\ &= \frac{1}{\epsilon^2}\sum_{n=1}^{\infty} \frac{1}{\{k(n)\}^2}\sum_{m=1}^{k(n)} \mathsf{Var}\left[Y_m\right] \\ &= \frac{1}{\epsilon^2}\sum_{m=1}^{\infty} \mathbb{V}\left[Y_m\right] \sum_{n:\,k(n)\geq m} \frac{1}{\{k(n)\}^2}\end{aligned} \tag{A.6}$$

を得る. 最後の等号は, すべての項が非負なので Fubini の定理を用いて和の順番を交換した. $n \geq 1$ に対して

$$k(n) = [\alpha^n] \quad かつ \quad [\alpha^n] \geq \frac{\alpha^n}{2}$$

[1]なぜならば

$$\frac{t}{[t]} \leq 2 \qquad (t \geq 1)$$

であることからわかる.

なので

$$\sum_{n:\,\alpha^n \geq m} \frac{1}{[\alpha^n]^2} \leq 4 \sum_{n:\,\alpha^n \geq m} \frac{1}{\alpha^{2n}} \leq \frac{4}{(1-\alpha^{-2})m^2} \tag{A.7}$$

となる. (A.6) と (A.7) を合わせると

$$\begin{aligned}\sum_{n=1}^{\infty} \mathsf{Pr}\left(\left|T_{k(n)} - \mathsf{E}\left[T_{k(n)}\right]\right| > \epsilon k(n)\right) &\leq \frac{4}{(1-\alpha^{-2})\epsilon^2} \sum_{m=1}^{\infty} \frac{\mathsf{E}[Y_m^2]}{m^2} \\ &< \infty \qquad (\because \text{補題 A.4})\end{aligned}$$

となる. ϵ は任意だったので

$$\frac{T_{k(n)} - \mathsf{E}[T_{k(n)}]}{k(n)} \overset{\text{a.s.}}{\to} 0 \qquad (n \to \infty)$$

を得る. 有界収束定理から

$$\mathsf{E}[Y_k] \overset{n\to\infty}{\longrightarrow} \mathsf{E}[X_1]$$

である. よって

$$\frac{\mathsf{E}\left[T_{k(n)}\right]}{k(n)} \overset{n\to\infty}{\longrightarrow} \mathsf{E}[X_1]$$

を得る. これより

$$\frac{T_{k(n)}}{k(n)} \overset{\text{a.s.}}{\to} \mathsf{E}[X_1] \qquad (n \to \infty)$$

がわかる. $Y_n \geq 0\,(\forall n \in \mathbb{N})$ なので, $k(n) \leq m < k(n+1)\,(m \in \mathbb{N})$ に対して

$$\frac{T_{k(n)}}{k(n+1)} \leq \frac{T_m}{m} \leq \frac{T_{k(n+1)}}{k(n)}$$

である. $k(n) = [\alpha^n]$ だったので,

$$\frac{k(n+1)}{k(n)} \overset{n\to\infty}{\longrightarrow} \alpha$$

である. よって

$$\frac{1}{\alpha}\mathsf{E}[X_1] \leq \liminf_{m\to\infty} \frac{T_m}{m} \leq \limsup_{m\to\infty} \frac{T_m}{m} \leq \alpha \mathsf{E}[X_1]$$

となる. $\alpha > 1$ は任意だったので

$$\frac{T_m}{m} \overset{\text{a.s.}}{\to} \mathsf{E}[X_1] \qquad (n \to \infty)$$

がわかる. □

A.1.3 Weierstrass の近似定理

大数の強法則を用いて, Weierstrass の近似定理を証明する.

系 A.5 (Weierstrass の近似定理)**.** 閉区間 $[0,\,1]$ 上の任意の連続関数 f は多項式の極限として表すことができる. 特に, $f(0)=f(1)=0$ のとき [2], $n\in\mathbb{N}$ に対して

$$B_n(x)=\sum_{j=1}^{n} f\left(\frac{j}{n}\right)\begin{pmatrix} n \\ j \end{pmatrix} x^j(1-x)^{n-j}$$

とおいたとき

$$f(x)=\lim_{n\to\infty} B_n(x) \tag{A.8}$$

が成立する.

Proof. $x=0,\,1$ のとき $(A.8)$ は明らかなので, $0<x<1$ に対して, $(A.8)$ が成立すること示す. $X_j\,(j=1,\,2,\,\ldots)$ は i.i.d. 確率変数列で

$$\mathsf{Pr}(X_j=1)=x=1-\mathsf{Pr}(X_j=0)$$

をみたすとする. いま

$$\mathsf{E}\left[\frac{X_1+X_2+\cdots+X_n}{n}\right]=x$$

に注意して, Chebyshev の不等式 (系 3.3) を用いると, 任意の $\epsilon>0$ に対して

$$\begin{aligned}
&\mathsf{Pr}\left(\left|\frac{X_1+X_2+\cdots+X_n}{n}-x\right|\geq\epsilon\right)\\
&\qquad\leq\frac{\mathsf{Var}\left(\dfrac{X_1+X_2+\cdots+X_n}{n}\right)}{\epsilon^2}\\
&\qquad=\frac{\dfrac{1}{n^2}\big(\mathsf{Var}(X_1)+\mathsf{Var}(X_2)+\cdots+\mathsf{Var}(X_n)\big)}{\epsilon^2}\\
&\qquad=\frac{x(1-x)}{n\epsilon^2}
\end{aligned}$$

となる. したがって

$$\frac{X_1+X_2+\cdots+X_n}{n}\xrightarrow{P}x$$

[2] $f(0)=a,\ f(1)=b$ に対して, $\widetilde{f}(x)=f(x)+\big\{f(b)-f(a)\big\}x+f(a)$ とおくと $\widetilde{f}$ は $[0,\,1]$ 上の連続関数で $\widetilde{f}(0)=\widetilde{f}(1)=0$ となるので, 左記の設定は一般性を失わない仮定である.

がわかる. さらに, 定理 4.9(2) より

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \rightsquigarrow x$$

を得る. f は有界連続関数なので, 定理 4.12(2) から

$$\lim_{n\to\infty} \mathsf{E}\left[f\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)\right] = f(x)$$

となる. しかし, $S := X_1 + X_2 + \cdots + X_n \sim \mathsf{Bino}(n,\, x)$ なので

$$\begin{aligned}\mathsf{E}\left[f\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right)\right] &= \mathsf{E}\left[f\left(\frac{S}{n}\right)\right]\\ &= \sum_{j=0}^{n} f\left(\frac{j}{n}\right)\binom{n}{j} x^j (1-x)^{n-j}\\ &= B_n(x)\end{aligned}$$

がわかる. よって, 主張は証明された. □

A.2 中心極限定理の証明

まず $\mathsf{E}[X_1^3] < \infty$ を仮定して証明をする.

$$Y_j = \frac{X_j - \mu}{\sigma}, \qquad (j = 1,\, 2,\, \ldots,\, n)$$

とおくと $\mathsf{E}[Y_1] = 0$, $\mathsf{E}[Y_1^2] = 1$, $\mathsf{E}[|Y_1|^3] < \infty$ となることに注意する. $Z_1,\, Z_2,\, \ldots,\, Z_n \overset{\text{i.i.d.}}{\sim} \mathsf{N}(0,\, 1)$ とし

$$S_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} Y_j, \quad S = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} Z_j \sim \mathsf{N}(0,\, 1)$$

とおく.

$\forall x \in \mathbb{R}$ に対して

$$\mathsf{Pr}\big(S_n \le x\big) \overset{n\to\infty}{\longrightarrow} \mathsf{Pr}\big(S \le x\big) \tag{A.9}$$

を示せばよい. ここで

$$g(y) = \mathbb{1}_{(-\infty,\, x]}(y) \qquad (y \in \mathbb{R})$$

とおくと (A.9) を

$$\mathsf{E}\big[g(S_n)\big] \overset{n\to\infty}{\longrightarrow} \mathsf{E}\big[g(S)\big] \tag{A.10}$$

と書きかえることができる. さらに $\forall \delta > 0$ に対して関数 $g_{\delta,\, x} : \mathbb{R} \longrightarrow [0,\, 1]$ を以下のような性質をもつ関数とする.

- $g_{\delta,x}(y) = 1 \, (y \le x)$.
- $g_{\delta,x}(y) = 0 \, (y \le x + \delta)$.
- $g_{\delta,x}$ は有界な 3 次の導関数をもつ．3 次の導関数の絶対値の上界を C と書くことにする．

さらに $g_{\delta,x-\delta}(y) = g_{\delta,x}(y+\delta)$ と定める．すると

$$g_{\delta,x-\delta}(y) \le g(y) \le h_{\delta,x}(y), \qquad (y \in \mathbb{R})$$

となる．このことから $\forall x \in \mathbb{R}$ に対して

$$\begin{aligned}
\mathsf{Pr}\big(S_n \le x\big) &= \mathsf{E}\big[g_{\delta,x}(S_n)\big] \\
&= \mathsf{E}\big[g_{\delta,x}(S_n)\big] - \mathsf{E}\big[g_{\delta,x}(S)\big] + \mathsf{E}\big[g_{\delta,x}(S)\big] \\
&\le \mathsf{Pr}\big(Z \le x + \delta\big) + \big|\mathsf{E}\big[g_{\delta,x}(S_n)\big] - \mathsf{E}\big[g_{\delta,x}(S)\big]\big|
\end{aligned}$$

を得る．同様に

$$\mathsf{Pr}\big(S_n \le x\big) \ge \mathsf{Pr}\big(S \le x - \delta\big) - \big|\mathsf{E}\big[g_{\delta,x-\delta}(S_n)\big] - \mathsf{E}\big[g_{\delta,x-\delta}(S)\big]\big|$$

を得る．これらの不等式を合わせると

$$\begin{aligned}
&\mathsf{Pr}\big(S \le x + \delta\big) + \big|\mathsf{E}\big[g_{\delta,x}(S_n)\big] - \mathsf{E}\big[g_{\delta,x}(S)\big]\big| \\
&\quad \ge \mathsf{Pr}\big(S_n \le x\big) \\
&\quad \ge \mathsf{Pr}\big(S \le x - \delta\big) - \big|\mathsf{E}\big[g_{\delta,x-\delta}(S_n)\big] - \mathsf{E}\big[g_{\delta,x-\delta}(S)\big]\big|
\end{aligned}$$

となる．δ を 0 に近づけると $\mathsf{Pr}\big(S \le x \pm \delta\big)$ は $\mathsf{Pr}\big(S \le x\big)$ に近づく．したがって

$$\mathsf{E}\big[g_{\delta,x}(S_n)\big] \overset{n\to\infty}{\longrightarrow} \mathsf{E}\big[g_{\delta,x}(S)\big] \quad と \quad \mathsf{E}\big[g_{\delta,x-\delta}(S_n)\big] \overset{n\to\infty}{\longrightarrow} \mathsf{E}\big[g_{\delta,x-\delta}(S)\big] \tag{A.11}$$

を示せばよい．よって $h : \mathbb{R} \to [0,\,1]$ を有界な 3 次の導関数をもつ関数として (A.11) を示す．

$1 \le m \le n+1$ に対して

$$T_m = \frac{1}{\sqrt{n}}\big(Y_1 + \cdots + Y_{m-1} + Z_m + \cdots + Z_n\big)$$

とおく．このとき

$$\begin{aligned}
\big|\mathsf{E}\big[h(S_n)\big] - \mathsf{E}\big[h(S)\big]\big| &= \left|\sum_{m=1}^{n+1}\big\{\mathsf{E}\big[h(T_m)\big] - \mathsf{E}\big[h(T_{m+1})\big]\big\}\right| \\
&\le \sum_{m=1}^{‘} n+1 \big|\mathsf{E}\big[h(T_m) - h(T_{m+1})\big]\big|
\end{aligned} \tag{A.12}$$

と書きかえる. 次に $m=1,2,\ldots,n$ に対して

$$U_m = \frac{1}{\sqrt{n}}\bigl(Y_1+\cdots+Y_{m-1}+Z_{m+1}+\cdots+Z_n\bigr)$$

とおくと

$$T_m = U_m + \frac{1}{\sqrt{n}}Z_m, \qquad T_{m+1} = U_m + \frac{1}{\sqrt{n}}Y_m$$

となる. Taylor 展開を用いると

$$\begin{aligned} h(T_m)-h(U_m) &= (T_m-U_m)\,\dot{h}\,(U_m) + \frac{1}{2}(T_m-U_m)^2\,\ddot{h}\,(U_m) \\ &\quad + \frac{1}{6}(T_m-U_m)^3\,\dddot{h}\,(U_m^*), \qquad \text{(A.13)} \\ h(T_{m+1})-h(U_m) &= (T_{m+1}-U_m)\,\dot{h}\,(U_m) + \frac{1}{2}(T_{m+1}-U_m)^2\frac{1}{2}\,\dot{h}\,(U_m)+\,\ddot{h}\,(U_m) \\ &\quad + \frac{1}{6}(T_{m+1}-U_m)^3\,\dddot{h}\,(U_m^{**}), \qquad \text{(A.14)} \end{aligned}$$

を得る. ただし $\dot{h},\ \ddot{h},\ \dddot{h}$ は h の 1 次, 2 次, 3 次の導関数で, U_m^*, U_m^{**} は T_m と U_m の間と T_{m+1} と U_m の間の数である. U_m は Y_m と Z_m の両方とも独立であるので

$$\begin{aligned} &\mathsf{E}\bigl[\dot{h}\,(U_m)Z_m\bigr] = \mathsf{E}\bigl[\dot{h}\,(U_m)Y_m\bigr] = 0, \\ &\mathsf{E}\bigl[\ddot{h}\,(U_m)Z_m^2\bigr] = \mathsf{E}\bigl[\ddot{h}\,(U_m)\bigr] = \mathsf{E}\bigl[\ddot{h}\,(U_m)Y_m^2\bigr], \\ &\bigl|\mathsf{E}\bigl[\dddot{h}\,(U_m^*)\bigr]\bigr| \le C,\ \bigl|\mathsf{E}\bigl[\dddot{h}\,(U_m^{**})\bigr]\bigr| \le C \end{aligned}$$

となる. 上の期待値と (A.13) と (A.14) を合わせると

$$\begin{aligned} \Bigl|\mathsf{E}\bigl[h(T_m)-h(T_{m+1})\bigr]\Bigr| &= \Bigl|\mathsf{E}\bigl[h(T_m)-h(U_m)-\bigl(h(T_{m+1})-h(U_m)\bigr)\bigr]\Bigr| \\ &\le \left|\frac{C}{6n^{3/2}}\mathsf{E}\bigl[|Z_m|^3+|Y_m|^3\bigr]\right| \end{aligned}$$

を得る. この式を (A.11) に代入すると

$$|\mathsf{E}\bigl[h(S_n)\bigr]-\mathsf{E}\bigl[h(S)\bigr]\Bigr| \le \frac{C}{6\sqrt{n}}\mathsf{E}\bigl[|Z_1|^3+|Y_1|^3\bigr] \overset{n\to\infty}{\longrightarrow} 0$$

を得る. よって Y_1 が有界な 3 次の積率をもつとき

$$S_n \rightsquigarrow S \sim \mathsf{N}(0,1)$$

が示せた.

最後に $\mathsf{E}\left[Y_1^3\right] < \infty$ の仮定を除くための議論を行う．そのために任意の $\epsilon > 0$ とし, $\sup_{x\in\mathbb{R}} |\ddot{h}(y)| \le C'$, $\sup_{x\in\mathbb{R}} |\dddot{h}(y)| \le C$ をみたす定数 C をとる．$m = 1, 2, \ldots, n$ に対して, $|Y_m| \le \epsilon\sqrt{n}$ のとき

$$\begin{aligned}
&\left|h(T_m) - h(U_m) - \frac{\dot{h}(U_m)Y_m}{\sqrt{n}} - \frac{\ddot{h}(U_m)Y_m^2}{2n}\right| \mathbb{1}\{|Y_m| \le \epsilon\sqrt{n}\} \\
&\qquad \le C\left|\frac{\dddot{h}(U_m^*)Y_m^3}{6n^{3/2}}\right| \mathbb{1}\{|Y_m| \le \epsilon\sqrt{n}\} \\
&\qquad \le \frac{C\epsilon}{6n} Y_m^2 \mathbb{1}\{|Y_m| \le \epsilon\sqrt{n}\}
\end{aligned}$$

となる．一方 $|Y_m| > \epsilon\sqrt{n}$ のとき

$$\begin{aligned}
&\left|h(T_m) - h(U_m) - \frac{\dot{h}(U_m)Y_m}{\sqrt{n}} - \frac{\ddot{h}(U_m)Y_m^2}{2n}\right| \mathbb{1}\{|Y_m| > \epsilon\sqrt{n}\} \\
&\qquad \le \left\{\left|h(T_m) - h(U_m) - \frac{\dot{h}(U_m)Y_m}{\sqrt{n}}\right| + \left|\frac{\ddot{h}(U_m)Y_m^2}{2n}\right|\right\} \mathbb{1}\{|Y_m| > \epsilon\sqrt{n}\} \\
&\qquad \le \left\{\left|\frac{\ddot{h}(U_m^{***})Y_m^2}{2n}\right| + \left|\frac{\ddot{h}(U_m)Y_m^2}{2n}\right|\right\} \mathbb{1}\{|Y_m| > \epsilon\sqrt{n}\} \\
&\qquad \le \frac{C'Y_m^2}{n} \mathbb{1}\{|Y_m| > \epsilon\sqrt{n}\}
\end{aligned}$$

となる．ただし U_m^{***} は T_m と U_m の間にあるものである．これらの不等式を合わせると

$$\begin{aligned}
&\left|h(T_m) - h(U_m) - \frac{\dot{h}(U_m)Y_m}{\sqrt{n}} - \frac{\ddot{h}(U_m)Y_m^2}{2n}\right| \\
&\qquad \le \frac{C\epsilon}{6n} Y_m^2 \mathbb{1}\{|Y_m| \le \epsilon\sqrt{n}\} + \frac{C'Y_m^2}{n} \mathbb{1}\{|Y_m| > \epsilon\sqrt{n}\}
\end{aligned}$$

をなる．この式と (A.14) を合わせると

$$\begin{aligned}
\left|\mathsf{E}\left[h(S_n)\right] - \mathsf{E}\left[h(S)\right]\right| &\le \frac{C}{\sqrt{n}}\mathsf{E}\left[Z_1^3\right] + \frac{C}{6}\epsilon\mathsf{E}\left[Y_1^2\mathbb{1}\{|Y_1| \le \epsilon\sqrt{n}\}\right] \\
&\qquad + C'\mathsf{E}\left[Y_1^2\{Y_1 > \epsilon\sqrt{n}\}\right] \\
&\le \frac{C}{\sqrt{n}}\mathsf{E}\left[Z_1^3\right] + \frac{C\epsilon}{6} + \frac{C'}{6}\mathsf{E}\left[Y_1^2\{Y_1 > \epsilon\sqrt{n}\}\right] \\
&\xrightarrow{n\to\infty} \frac{C'}{6}\epsilon
\end{aligned}$$

を得る. 上の式で $\epsilon \to 0$ とすれば

$$\mathsf{E}\big[h(S_n)\big] \overset{n\to\infty}{\longrightarrow} \mathsf{E}\big[h(S)\big]$$

が証明できた. □

A.3 Laplace 変換の一意性

定義 A.6. $[0,\,\infty)$ 上に集中した分布 P について

$$f(s) = \int_{[0,\,\infty)} e^{-sx}\,\mathrm{d}\mathsf{P}(x)$$

を P の Laplace **変換**という. また, 非負値確率変数 X の Laplace 変換とは

$$f(s) = \mathsf{E}\big[e^{-sX}\big] = \int_{[0,\,\infty)} e^{-sx}\,\mathrm{d}\mathsf{P}(x) \qquad (s \geq 0)$$

のことである.

補題 A.7. 区間 $[0,\,1]$ に集中した分布の分布関数 G_1, G_2 に対して

$$\int_0^1 x^n\,\mathrm{d}\mathsf{G}_1(x) = \int_0^1 x^n\,\mathrm{d}\mathsf{G}_2(x) \qquad (n = 0,\,1,\,2,\ldots) \tag{A.15}$$

が成り立てば

$$\mathsf{G}_1 = \mathsf{G}_2$$

である.

Proof. Weierstrass の多項式近似定理により, $(A.15)$ が成り立てば, 任意の $[0,\,1]$ 上の連続関数 f に対して

$$\int_0^1 f(x)\,\mathrm{d}\mathsf{G}_1(x) = \int_0^1 f(x)\,\mathrm{d}\mathsf{G}_2(x)$$

が成り立つ.

いま, $0 \leq a < 1$ と $n \geq 1$ に対して

$$f_n(x) = \begin{cases} 1 & (0 \leq x \leq a), \\ n(x-a) & (a \leq x \leq a + 1/n), \\ 0 & (n + 1/n \leq x \leq 1) \end{cases}$$

と定める. すると $\mathsf{G}_1(0) = \mathsf{G}_2(0) = 0$ より

$$\begin{aligned} G_1(a) &= \int_0^a \mathrm{d}\mathsf{G}_1 = \lim_{n\to\infty}\int_0^1 f_n(x)\,\mathrm{d}\mathsf{G}_1(x) = \lim_{n\to\infty}\int_0^1 f_n(x)\,\mathrm{d}\mathsf{G}_2(x) \\ &= \int_0^a \mathrm{d}\mathsf{G}_2 = G_2(a) \end{aligned}$$

を得る．また $\mathsf{G}_1(1)=\mathsf{G}_2(1)$ だから，すてべの $a\in[0,\,1]$ に対して $\mathsf{G}_1(a)=\mathsf{G}_2(a)$ がわかる． □

定理 A.8. $[0,\,\infty)$ に集中した確率をもつ 2 つの分布関数 $\mathsf{F}_1,\,\mathsf{F}_2$ に対して，それぞれの Laplace 変換を $f_1,\,f_2$ とする．このとき，f_1 と f_2 が一致すれば，$\mathsf{F}_1=\mathsf{F}_2$ である．

Proof. $\mathsf{F}:\,[0,\,\infty)\longrightarrow\mathbb{R}$ を確率分布関数とする．$x=e^{-t}$, $\mathsf{G}(x)=1-\mathsf{F}(t)$ とすると

$$\int_0^\infty e^{-\lambda t}\,\mathrm{d}\mathsf{F}(t)=\int_0^1 x^\lambda\,\mathrm{d}\mathsf{G}(x)\qquad(\lambda\geq 0)$$

となる．よって $f_1=f_2$ ならば，

$$\int_0^1 x^n\,\mathrm{d}\mathsf{G}_1(x)=\int_0^1 x^n\,\mathrm{d}\mathsf{G}_2(x)\qquad(n=0,\,1,\,2,\ldots)$$

となる．あとは補題 A.7 を用いれば，定理は証明される． □

A.4 Dynkin 族の議論

定義 A.9. Ω を空でない集合とし，$\mathcal{L}$ と $\mathcal{P}$ を Ω の部分集合族とする．
(1) 集合族 $\mathcal{P}$ が π **システム**であるとは

$$A,\,B\in\mathcal{P}\implies A\cap B\in\mathcal{P}$$

が成立するときをいう．
(2) 集合族 $\mathcal{L}$ が λ **システム**であるとは，つぎの条件 (i) ～ (iii) が成り立つときをいう．

(i) $\Omega\in\mathcal{L}$.

(ii) $A,\,B\in\mathcal{L}$ かつ $A\subset B\implies B\setminus A\in\mathcal{L}$.

(iii) $A_j\in\mathcal{L}$ かつ $A_j\subset A_{j+1}\,(j=1,\,2,\,\ldots)\implies\bigcup_{j=1}^\infty A_j\in\mathcal{L}$.

命題 A.10. 空でない集合 Ω の部分集合族 $\mathcal{L}$ が λ システムであるために必要十分条件は次の条件 (i) ～ (iii) をすべてみたすことである．

(i)′ $\Omega\in\mathcal{L}$.

(ii)′ $A\in\mathcal{L}\implies A^c\in\mathcal{L}$.

(iii)′ $A_j\in\mathcal{L}\,(j=1,\,2,\,\ldots)$ かつ $A_i\cap A_j\neq\varnothing\,(i\neq j)\implies\bigcup_{j=1}^\infty A_j\in\mathcal{L}$.

Proof. $(\Rightarrow)$: (ii)$'$ の確認. $A \in \mathcal{L}$ のとき, $A \subset \Omega$ かつ $\Omega \in \mathcal{L}$ なので, 定義 $A.9$(ii) から

$$A^c = \Omega \setminus A \in \mathcal{L}$$

となる.
(iii)$'$ の確認. まず, $A,\,B \in \mathcal{L}$ かつ $A \cap B \neq \varnothing$ のとき

$$A \cup B \in \mathcal{L} \tag{A.16}$$

を示す.
仮定と (ii)$'$ から $A^c \in \mathcal{L}$ かつ $B \subset A^c$ である. なぜならば, $\forall \omega \in B$ のとき, $A \cap B = \varnothing$ なので, $\omega \notin A$ となる. よって $\omega \in A^c$ がわかるので, $B \subset A^c$ が示せた.
このことと (ii) から $A^c \setminus B \in \mathcal{L}$ となる. さらに

$$A^c \cap B^c = A^c \setminus B \in \mathcal{L} \tag{A.17}$$

となることと (ii)$'$ より

$$A \cup B = \big(A^c \cap B^c\big)^c \in \mathcal{L}$$

がわかる. よって

$$A,\,B \in \mathcal{L} \implies A \cup B \in \mathcal{L} \tag{A.18}$$

がわかる.
$A_j \in \mathcal{L}\,(j = 1,\,2,\ldots)$ と $A_i \cap A_j = \varnothing\,(i \neq j)$ とする. このとき

$$B_n := \bigcup_{j=1}^{n} A_j \quad (n \in \mathbb{N})$$

とおく. $(A.18)$ より $B_n \in \mathcal{L}$ となる. さらに, B_n の定め方から

$$B_n \subset B_{n+1}$$

となる. よって定理 $A.9$(iii) から

$$\bigcup_{j=1}^{\infty} A_j = \bigcup_{n=1}^{\infty} B_n \in \mathcal{L}$$

がわかる. よって (i)$'$ $\sim$ (iii)$'$ が確認できた.
$(\Leftarrow)$: (ii) の確認. $A,\,B \in \mathcal{L}$ かつ $A \subset B$ とする. すると (ii)$'$ から $A^c,\,B^c \in \mathcal{L}$ となる. さらに $A \subset B$ から $A \cap B^c = \varnothing$ となる. よって (iii)$'$ から $A \cup B^c \in \mathcal{L}$ となる. 再度, (ii)$'$ を用いると

$$B \setminus A = B \cap A^c = \big(A \cup B^c\big)^c \in \mathcal{L}$$

がわかる.
(iii) の確認. $A_n \in \mathcal{L}$ かつ $A_n \subset A_{n+1}\,(n = 1, 2, \ldots)$ とする. ここで

$$B_n := A_{n+1} \setminus A_n\,(n = 1, 2, \ldots)$$

とおくと (ii) から $B_n \in \mathcal{L}$ となる. さらに B_n の定め方から

$$B_n \cap B_m = \varnothing\,(m \neq n)$$

となる. よって (iii)$'$ から

$$\bigcup_{j=1}^{\infty} A_j = \bigcup_{n=1}^{\infty} B_n \in \mathcal{L}$$

がわかる. 以上から (i) $\sim$ (iii) が確認できた. □

命題 A.11. Ω を空でない集合とし, $\mathcal{C}$ を Ω の部分集合族とする. 部分集合族 $\mathcal{C}$ が π システムかつ λ システムならば, $\mathcal{C}$ は σ 集合族となる.

Proof. まず, 部分集合族 $\mathcal{C}$ は体であることを示す.
(i) $\mathcal{C}$ は λ システムなので, 命題 A.10(ii)$'$ から $\Omega \in \mathcal{C}$ である.
(ii) $\mathcal{C}$ は π システムである. $A \in \mathcal{C}$ のとき, $\Omega \in \mathcal{C}$ なので, 命題 A.10(ii)$'$ から

$$A^c = \Omega \setminus A \in \mathcal{C}$$

となる.
(iii) $A_j \in \mathcal{C}\,(j = 1, 2, \ldots)$ とする. $\mathcal{C}$ は π システムなので, 定義 A.9(1) から

$$\bigcap_{j=1}^{n} A_j \in \mathcal{C}\,(n \in \mathbb{N})$$

となる. よって, (i) $\sim$ (iii) から $\mathcal{C}$ は体となることがわかる.
$\mathcal{C}$ は体であるので, あとは

$$\bigcup_{j=1}^{n} A_i \in \mathcal{C}$$

を示せばよい. $\mathcal{C}$ は体なので

$$B_n := \bigcup_{j=1}^{n} A_j \in \mathcal{C} \qquad (n \in \mathbb{N})$$

となる. さらに B_n の定め方から $B_n \subset B_{n+1}\,(n = 1, 2, \ldots)$ となる. よって, $\mathcal{C}$ は λ システムなので, 定義 $A.9$(iii) から

$$\bigcup_{j=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n \in \mathcal{C}$$

がわかる. 以上から $\mathcal{C}$ は σ 集合族であることがわかる. □

命題 A.12. Ω を空でない集合とする. Ω の部分集合族 $\mathcal{P}$ を π システムとし, $\mathcal{L}[\mathcal{P}]$ を $\mathcal{P}$ を含む最小の λ システムとする. このとき, $\mathcal{L}[\mathcal{P}]$ は π システムとなる.

Proof. $\sigma[\mathcal{P}]$ を $\mathcal{P}$ を含む最小の σ 加法族とする. $A \in \mathcal{L}[\mathcal{P}]$ を固定し, この A に対して

$$\mathcal{G}_A := \big\{B \in \sigma[\mathcal{P}];\, A \cap B \in \mathcal{L}[\mathcal{P}]\big\}$$

と定める. $\mathcal{G}_A$ に関する以下の主張を順に証明していく.
主張 1: $A \in \sigma[\mathcal{P}]$ のとき, $\mathcal{G}_A$ は λ システムである.

主張 1 の証明: 命題 $A.10$ の条件 (i)$'$ $\sim$ (iii)$'$ を確認する.
(i)$'$ $A \cap \Omega = A \in \mathcal{L}[\mathcal{P}]$ なので, $\mathcal{G}_A$ の定義から $\Omega \in \mathcal{G}_A$ がわかる.
(ii)$'$ $B \in \mathcal{G}_A$ とする. このとき

$$B^c \cap A = A \setminus (A \cap B)$$

に注意する. $B \in \mathcal{G}_A$ なので

$$A \cap B \in \mathcal{L}[\mathcal{P}] \tag{A.19}$$

である. 仮定と $(A.19)$ から A, $A \cap B \in \mathcal{L}[\mathcal{P}]$ である. $\mathcal{L}[\mathcal{P}]$ は λ システムなので, 定義 $A.9$(2)(ii) から

$$B^c \cap A = A \setminus (A \cap B) \in \mathcal{L}[\mathcal{P}]$$

となる. よって $B^c \in \mathcal{L}[\mathcal{P}]$ がわかる.
(iii)$'$ $\{B_n\}_{n=1}^{\infty}$ は互いに排反な列で $B_n \in \mathcal{G}_A\,(n = 1, 2, \ldots)$ とする. このとき

$$A \cap \left(\bigcup_{j=1}^{\infty} B_j\right) = \bigcup_{j=1}^{\infty} \big(A \cap B_j\big) \tag{A.20}$$

となる. したがって $\{A \cap B_n\}_{n=1}^{\infty}$ は互いに排反な列で, $(A.19)$ から $A \cap B_n \in \mathcal{G}_A$ となることに注意する. よって, $A \cap B_n \in \mathcal{L}[\mathcal{P}]$ かつ $\mathcal{L}[\mathcal{P}]$ は λ システムなので, 命題 $A.10$(ii)$'$ から

$$\bigcup_{n=1}^{\infty} (A \cap B_n) \in \mathcal{L}[\mathcal{P}]$$

となる．このことと $(A.20)$ から

$$\bigcup_{n=1}^{\infty} B_n \in \mathcal{G}_A$$

がわかる．

よって (i)$'$ $\sim$ (iii)$'$ から $\mathcal{G}_A$ は λ システムであることがわかる． $\square$

主張 2: $A \in \mathcal{P}$ ならば $\mathcal{L}[\mathcal{P}] \subset \mathcal{G}_A$ である．

主張 2 の証明: $A \in \mathcal{P} \subset \mathcal{L}[\mathcal{P}]$ を取る．主張 1 から $\mathcal{G}_A$ は λ システムとなる．$B \in \mathcal{P}$ とする．$\mathcal{P}$ は π システムなので, 定義 $A.9(1)$ から

$$A \cap B \in \mathcal{P} \subset \mathcal{L}[\mathcal{P}]$$

となるので, $\mathcal{G}_A$ は λ システムで $\mathcal{P}$ を含むので, $\mathcal{L}[\mathcal{P}]$ は $\mathcal{P}$ を含むの最小性の λ システムなので

$$\mathcal{G}_A \supset \mathcal{L}[\mathcal{P}]$$

が示せた． $\square$

主張 3: $A \in \mathcal{P}$ かつ $B \in \mathcal{L}[\mathcal{P}]$ ならば $A \cap B \in \mathcal{L}[\mathcal{P}]$ となる．

主張 3 の証明: 主張 2 から

$$B \in \mathcal{L}[\mathcal{P}] \subset \mathcal{G}_A$$

となる．よって, $\mathcal{G}_A$ の定義から

$$A \cap B \in \mathcal{L}[\mathcal{P}]$$

がわかる． $\square$

主張 4: $A \in \mathcal{L}[\mathcal{P}]$ ならば $\mathcal{L}[\mathcal{P}] \subset \mathcal{G}_A$ となる．

主張 4 の証明: $B \in \mathcal{P}$ と $A \in \mathcal{L}[\mathcal{P}]$ を取る．主張 3 から

$$A \cap B \in \mathcal{L}[\mathcal{P}] \tag{A.21}$$

となる．よって, $A \in \mathcal{L}[\mathcal{P}]$ で $(A.21)$ であることに注意して, $\mathcal{G}_A$ の定義を思い出すと $B \in \mathcal{G}_A$ となる． 以上から

$$B \in \mathcal{P} \Longrightarrow B \in \mathcal{G}_A$$

がわかる. すなわち, $\mathcal{P} \subset \mathcal{G}_A$ である. さらに, 主張 1 から $\mathcal{G}_A$ は λ システムである. 以上から, $\mathcal{G}_A$ は $\mathcal{P}$ を含む λ システムとでであることがわかる. よって, $\mathcal{L}[\mathcal{P}]$ は $\mathcal{P}$ を含む最小の λ システムなので

$$\mathcal{L}[\mathcal{P}] \subset \mathcal{G}_A$$

がわかる. □

主張 5: $\mathcal{L}[\mathcal{P}]$ は π システムである.

主張 5 の証明: $A, B \in \mathcal{L}[\mathcal{P}]$ とする. すると主張 4 から

$$B \in \mathcal{L}[\mathcal{P}] \implies B \in \mathcal{G}_A$$

が成り立つ. $\mathcal{G}_A$ の定義から

$$A \cap B \in \mathcal{L}[\mathcal{P}]$$

となり, $\mathcal{L}[\mathcal{P}]$ は π システムであることがわかる. □

定理 A.13. Ω を空でない集合とする. Ω の部分集合族 $\mathcal{P}$ を π システムとし, $\mathcal{L}$ を λ システムかつ $\mathcal{P} \subset \mathcal{L}$ とする. このとき

$$\sigma[\mathcal{P}] \subset \mathcal{L} \tag{A.22}$$

が成立する.

Proof. 命題 *A*.12 から $\mathcal{L}[\mathcal{P}]$ は π システムである. よって命題 *A*.11 より $\mathcal{L}[\mathcal{P}]$ は σ 加法族である. よって $\sigma[\mathcal{P}]$ の最小性から

$$\mathcal{L}[\mathcal{P}] \supset \sigma[\mathcal{P}]$$

となる. さらに, $\mathcal{L}$ は $\mathcal{P}$ を含む λ システムである. 一方, $\mathcal{L}[\mathcal{P}]$ は $\mathcal{P}$ を含む最初の λ システムある. この二つのことから

$$\mathcal{L} \supset \mathcal{L}[\mathcal{P}] \supset \mathcal{P} \tag{A.23}$$

となる. (*A*.22) と (*A*.23) を合わせると

$$\mathcal{L}{=}\sigma[\mathcal{P}]$$

がわかる. □

A.5 Dynkin の定理の応用例

A.5.1 その 1

命題 A.14. P_1 と P_2 を標本空間 $\bigl(\mathbb{X}, \mathcal{B}\bigr)$ 上の確率測度とする. 集合族

$$\mathcal{L} := \bigl\{B \in \mathcal{B};\, \mathsf{P}_1(B) = \mathsf{P}_2(B)\bigr\}$$

は λ システムである.

Proof. 命題 A.10 の条件 (i)$'$ ～ (iii)$'$ を順に確認していく.

(i)$'$ $\mathsf{P}_1(\mathbb{X}) = \mathsf{P}_2(\mathbb{X})$ なので, $\mathbb{X} \in \mathcal{L}$ がわかる.

(ii)$'$ $A \in \mathcal{L}$ とする. すなわち

$$\mathsf{P}_1(A) = \mathsf{P}_2(A)$$

である. よって

$$\mathsf{P}_1(A^c) = 1 - \mathsf{P}_1(A) = 1 - \mathsf{P}_2(A) = \mathsf{P}_2(A^c)$$

であるので, $A^c \in \mathcal{L}$ がわかる.

(iii)$'$ $\{A_j\}_{j=1}^{\infty}$ は $\mathcal{L}$ の互いに排反な列とする. このとき $j = 1, 2, \ldots$ に対して

$$\mathsf{P}_1(A_j) = \mathsf{P}_2(A_j)$$

となる. よって

$$\mathsf{P}_1\biggl(\bigcup_{j=1}^{\infty} A_j\biggr) = \sum_{j=1}^{\infty} \mathsf{P}_1(A_j) = \sum_{j=1}^{\infty} \mathsf{P}_2(A_j) = \mathsf{P}_2\biggl(\bigcup_{j=1}^{\infty} A_j\biggr)$$

となるので

$$\bigcup_{j=1}^{\infty} A_j \in \mathcal{L}$$

がわかる. □

系 A.15. P_1 と P_2 は標本空間 $\bigl(\mathbb{X}, \mathcal{B}\bigr)$ 上の確率測度とする. $\mathcal{P}$ は π システムで

$$\forall A \in \mathcal{P} \implies \mathsf{P}_1(A) = \mathsf{P}_2(A)$$

ならば

$$\forall B \in \mathcal{B} \implies \mathsf{P}_1(B) = \mathsf{P}_2(B)$$

が成り立つ.

Proof. 命題 $A.14$ から

$$\mathcal{L} := \big\{A \in \mathcal{B};\, \mathsf{P}_1(A) = \mathsf{P}_2(A)\big\}$$

は λ システムである. しかし $\mathcal{L} \supset \mathcal{P}$ なので, 定理 $A.13$ から

$$\mathcal{L} \supset \sigma[\mathcal{P}]$$

となる. □

系 A.16. $\mathbb{X} = \mathbb{R}$ とし, P_1 と P_2 を $\big(\mathbb{R}, \mathcal{B}(\mathbb{R})\big)$ 上の確率測度とし, 対応する分布関数を F_1 と F_2 は等しいとする. すなわち

$$\mathsf{F}_1(x) = \mathsf{P}_1\big((-\infty, x]\big) = \mathsf{F}_2(x) = \mathsf{P}_2\big((-\infty, x]\big)$$

である. このとき

$$\mathsf{P}_1(B) = \mathsf{P}_2(B) \qquad \big(\forall B \in \mathcal{B}(\mathbb{R})\big)$$

が成り立つ.

Proof.

$$\mathcal{P} := \big\{(-\infty, x];\, x \in \mathbb{R}\big\}$$

とおく. $\forall x, y \in \mathbb{R}$ に対して

$$(-\infty, x] \bigcap (-\infty, y] = (-\infty, x \wedge y] \in \mathcal{P}$$

となる. ここで $x \wedge y$ は x と y の大きくない方である. 以上から $\mathcal{P}$ は π システムである.

さらに $\sigma[\mathcal{P}] = \mathcal{B}(\mathbb{R})$ であるので

$$\mathsf{F}_1(x) = \mathsf{F}_2(x)\,(\forall x \in \mathbb{R}) \implies \mathsf{P}_1(A) = \mathsf{P}_2(A)\,(\forall A \in \mathcal{P})$$

である. したがって, $\mathcal{L} = \{B \in \mathcal{B}(\mathbb{R});\, \mathsf{P}_1(B) = \mathsf{P}_2(B)\}$ とすると系 $A.15$ から

$$\mathcal{L} \supset \sigma[\mathcal{P}]$$

となる. 一方, $\mathcal{L}$ の定義から $\mathcal{L} \subset \mathcal{B}(\mathbb{R})$ である. さらに, 定理 $A.19$ から $\sigma[\mathcal{P}] = \mathcal{B}(\mathbb{R})$ である. よって, $\mathcal{L} = \sigma[\mathcal{P}] = \mathcal{B}(\mathbb{R})$ となることがわかる. 以上から

$$\mathsf{P}_1(A) = \mathsf{P}_2(A) \qquad (\forall A \in \mathcal{B}(\mathbb{R}))$$

がわかる. □

A.5.2 その２

$(\mathbb{X}, \mathcal{B})$ と $(\mathbb{Y}, \mathcal{C})$ を可測空間とし, μ を可測空間 $(\mathbb{X}, \mathcal{B})$ 上の σ 有限測度とする. $\mathcal{B}\otimes\mathcal{C}$ 可測関数 $f : \mathbb{X}\times\mathbb{Y} \longrightarrow \mathbb{R}$ に対して, 写像

$$y \mapsto \int_{\mathbb{X}} f(x, y)\, d\mu(x) \tag{A.24}$$

は $\mathcal{C}$ 可測であることを Dynkin の補題と標準機械を落ちいて示そう.

補題 A.17. $(\mathbb{X}, \mathcal{B})$ と $(\mathbb{Y}, \mathcal{C})$ を可測空間とし, μ を可測空間 $(\mathbb{X}, \mathcal{B})$ 上の σ 有限測度とする. 各 $L \in \mathcal{B}\otimes\mathcal{C}$ に対して, 写像

$$y \mapsto \int_{\mathbb{X}} \mathbb{1}_L(x, y)\, d\mu(x) \tag{A.25}$$

は $\mathcal{C}$ 可測である.

Proof. 各 $y \in \mathbb{Y}$ に対して, 写像

$$x \mapsto 1_L(x, y) \tag{A.26}$$

は $\mathcal{B}$ 可測となる. なぜならば, 写像

$$i_y : x \mapsto (x, y)$$

は $\mathcal{B}\setminus\mathcal{B}\otimes\mathcal{C}$ 可測となること [3] がわかる. よって, $(A.25)$ は $\mathcal{B}$ 可測となる. このことから $(A.24)$ の右辺の積分は well-defined である.

つぎに

$$\mathcal{L} := \left\{ L \in \mathcal{B}\otimes\mathcal{C};\ y \mapsto \textstyle\int_{\mathbb{X}} \mathbb{1}_L(x, y)\, d\mu(x) \text{ は } \mathcal{C} \text{ 可測} \right\}$$

とおいたとき, $\mathcal{L} = \mathcal{B}\otimes\mathcal{C}$ を示す.

そのために, $B \in \mathcal{B}$ と $C \in \mathcal{C}$ を取る. すると

$$\begin{aligned}\int_{\mathbb{X}} \mathbb{1}_{B\times C}(x, y)\, d\mu(x) &= \int_{\mathbb{X}} \mathbb{1}_B(x)\mathbb{1}_C(y)\, d\mu(x) = \mathbb{1}_C(y)\int_{\mathbb{X}} \mathbb{1}_B(x)\, d\mu(x)\\ &= \mathbb{1}_C(y)\mu(B)\end{aligned}$$

となる. このことから, 写像

$$y \mapsto \int_{\mathbb{X}} \mathbb{1}_{B\times C}(x, y)\, d\mu(x) = \mathbb{1}_C(y)\mu(B)$$

は, 明らかに $\mathcal{C}$ 可測となる. よって

$$B\times C \in \mathcal{L} \tag{A.27}$$

[3]Hansen (2009, Lemma 8.4) を参照のこと.

である. ここで集合族 $\mathcal{P}$ を

$$\mathcal{P} := \bigl\{B \times C;\, B \in \mathcal{B},\, C \in \mathcal{C}\bigr\}$$

で定めると (A.27) から $\mathcal{P} \subset \mathcal{L}$ であり, さらに $\mathcal{P}$ は π システムである. したがって, $\mathcal{L}$ が λ システムであることがわかると定理 A.13 から

$$\sigma[\mathcal{P}] \subset \mathcal{L}$$

がわかる. さらに, $\mathcal{B} \otimes \mathcal{C}$ の定義を思い出すと

$$\sigma[\mathcal{P}] = \mathcal{B} \otimes \mathcal{C}$$

であるので

$$\mathcal{L} = \mathcal{B} \otimes \mathcal{C}$$

がわかる. よって, (A.24) で定義された写像は $\mathcal{C}$ 可測であることがわかる. □

補題 A.18.

$$\mathcal{L} := \Bigl\{L \in \mathcal{B} \otimes \mathcal{C};\;\; y \mapsto \textstyle\int_{\mathbb{X}} \mathbb{1}_L(x,\, y)\, d\mu(x) \text{ は } \mathcal{C} \text{ 可測 }\Bigr\}$$

は λ システムである.

Proof. 定義 A.9 の (i) ～ (iii) を確認する.
(i): $\mathbb{X} \times \mathbb{Y} \in \mathcal{L}$ であることは明らか.
(ii): $L_1,\, L_2 \in \mathcal{L}$ かつ $L_1 \subset L_2$ に対して

$$\mathbb{1}_{L_1}(x,\, y) + \mathbb{1}_{L_2 \setminus L_1}(x,\, y) = \mathbb{1}_{L_2}(x,\, y) \qquad \bigl((x,\, y) \in \mathbb{X} \times \mathbb{Y}\bigr)$$

である. $L_1,\, L_2 \in \mathcal{L}$ なので, 写像

$$y \mapsto \int_{\mathbb{X}} \mathbb{1}_{L_1}(x,\, y)\, d\mu(x) \quad \text{と} \quad y \mapsto \int_{\mathbb{X}} \mathbb{1}_{L_2}(x,\, y)\, d\mu(x)$$

は $\mathcal{C}$ 可測である. 写像

$$y \mapsto \int_{\mathbb{X}} \mathbb{1}_{L_2 \setminus L_1}(x,\, y)\, d\mu(x)$$

は $\mathcal{C}$ 可測関数の差なので, $\mathcal{C}$ 可測となる. よって, $L_2 \setminus L_1 \in \mathcal{L}$ となる.
$L_n \in \mathcal{L}\,(n = 1,\, 2,\, \ldots)$ かつ $L_1 \subset L_2 \subset \cdots$ とする. $L = \bigcup_{n=1}^{\infty} L_n$ とおいたとき

$$\mathbb{1}_{L_n}(x,\, y) \nearrow \mathbb{1}_L(x,\, y) \qquad \bigl((x,\, y) \in \mathbb{X} \times \mathbb{Y}\bigr)$$

となる. よって, 単調収束定理から

$$\lim_{n\to\infty}\int_{\mathbb{X}} \mathbb{1}_{L_n}(x,\,y)\,d\mu(x) = \int_{\mathbb{X}} \mathbb{1}_L(x,\,y)\,d\mu(x)$$

となる. よって, 写像

$$y \mapsto \int_{\mathbb{X}} \mathbb{1}_L(x,\,y)\,d\mu(x)$$

も $\mathcal{C}$ 可測となる. よって, $L \in \mathcal{L}$ がわかる.

(i) $\sim$ (iii) から $\mathcal{L}$ は λ システムであることが示せた. □

(A.24) の可測性の証明: あとは標準機械を用いる. すなわち, 以下の手続きをふくことで証明をする.

① 指示関数 $\mathbb{1}_A$ に対してある性質を証明する.

② 有限個の $A_1,\,A_2,\,\ldots,\,A_n \in \mathcal{A}$ で互いに排反なものと異なる $a_1,\,a_2,\,\ldots,\,a_n$ に対して単関数 $\sum_{k=1}^n a_k \mathbb{1}_A$ に対して ① で示した性質を確認する.

③ 単関数の極限として非負の確率変数を定義し, ② で示した性質を非負値確率変数に対して証明する.

④ 一般の可測関数 f を非負の部分 f^+ と負の部分 f^- にわけたものに ③ を適用し, $f = f^+ - f^-$ にたいしてその結果を拡張して, X についての性質を確認する.

すなわち, 補題 A.18 から出発して, ① $\sim$ ④ の議論を実行することにより, (A.24) によって定義される写像は $\mathcal{C}$ であることがわかる. □

A.5.3 $\mathbb{R}$ の位相についての注意

定理 A.19. $U \subset \mathbb{R}$ が開集合ならば, 互いに排反な開区間 $\{I_j\}_{j=1}^\infty$ があって

$$U = \bigcup_{j=1}^\infty I_j$$

と書ける.

Proof. U は開集合なので, 各 $x \in U$ に対して $y > x$ なる $y \in \mathbb{R}$ が存在して,

$$(x,\,y) \subset U$$

とできる. このとき

$$a := \inf\{z;\, (z,\, x) \subset U\}, \quad b := \sup\{y;\, (x,\, y) \subset U\}$$

と定める. a と b は x に依存することに注意する. このとき

$$x \in (a,\, b) \quad \text{かつ} \quad I_x = (a,\, b) \text{ は開区間}$$

である. これらを踏まえて, 次の主張を示していく.

(i) $U = \bigcup\big\{I_x;\, x \in U\big\}$.

(ii) $\forall x,\, y \in U$ に対して

$$x \neq y \implies I_x \cap I_y = \varnothing.$$

(iii) $\#\{I_x;\, x \in U\} \leq \aleph_0$. ただし $I_x = I_y$ なるものを除いてである.

まず, 次のことを注意する. $x \in U$ に対して, $I_x = (a,\, b)$ とすると $I_x \subset U$ かつ $a,\, b \notin U$ である. なぜならば $w \in (a,\, b)$ とすると $y > w$ なる $y \in \mathbb{R}$ が存在して, $(x,\, y) \subset U$ となる. よって $w \in U$ である. いま $b \in U$ と仮定するとある $\epsilon > 0$ が存在して

$$(b - \epsilon,\, b + \epsilon) \subset U$$

とできる. ϵ を十分小さくすると

$$\left(x,\, b - \frac{\epsilon}{2}\right) \subset U$$

となる. 結局

$$(x,\, b + \epsilon) \subset U$$

となり, b の定義と矛盾する. よって $b \notin U$ が示せた.

(i) 各 $x \in U$ に対して, $x \in I_x \subset U$ なので

$$\bigcup_{x \in U} I_x \subset U$$

となる. 逆は明らかなので

$$\bigcup_{x \in U} I_x = U$$

がわかる.

(ii) $x,\, y \in U$ とし, $I_x = (a,\, b)$ と $I_y = (c,\, d)$ と定める. $I_x \cap I_y \neq \varnothing$ と仮定する. このとき $a < d$ かつ $c < b$ となる. $c \notin U$ なので $c \in (a,\, b)$ である. $c < b$ なので

$$c \leq a \tag{A.28}$$

である. $a \notin U$ なので

$$a \notin (c,\, d)$$

である. よって $a < d$ なので

$$a \leq c \tag{A.29}$$

がわかる. $(A.28)$ と $(A.29)$ から

$$a = c$$

がわかる. 同様に

$$b = d \tag{A.30}$$

を示すことができる. よって

$$I_x \cap I_y \neq \varnothing \implies I_x = I_y$$

がわかる. これの対偶を取ると

$$I_x \neq I_y \implies I_x \cap I_y = \varnothing$$

がわかる.

(iii) 各 $I_x\,(x \in U)$ に対して

$$q_x \in I_x \cap \mathbb{Q}$$

を取る. $I_x \neq I_y$ ならば $I_x \cap I_y = \varnothing$ なので

$$q_x \neq q_y$$

がわかる. よって $\#(\mathbb{Q}) \leq \aleph_0$ なので, (iii) がわかる. □

A.6 凸関数の性質

A.7 Radon-Nikodym の定理と条件付き期待値の議論

$(\Omega, \mathcal{A}, \mathsf{Pr})$ を確率空間とし, $\mathcal{F} \subset \mathcal{A}$ も σ 加法族とする. X は $(\Omega, \mathcal{A}, \mathsf{Pr})$ 上の確率変数で $\mathsf{E}\big[|X|\big] < \infty$ とする.

定義 A.20. $\mathcal{F}$ を与えたときの X の**条件付き期待値**を次の条件をみたす任意の確率変数 Y とする.

(i) Y は $\mathcal{F}$ 可測.

(ii) 任意の $A \in \mathcal{F}$ に対して

$$\int_A X(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) = \int_A Y(\omega)\,\mathrm{d}\mathsf{Pr}(\omega).$$

まず条件付き期待値の存在と一意性を確認する.

補題 A.21. Y を定義 A.20 の条件 (i)(ii) をみたすとき Y は可積である.

Proof. $A := \{\omega \in \Omega;\, Y(\omega) > 0\}$ とおくと, $A,\, A^c \in \mathcal{F}$ となる. (ii) を用いると

$$\int_A Y(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) = \int_A X(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) \le \int_A |X(\omega)|\,\mathrm{d}\mathsf{Pr}(\omega),$$
$$\int_{A^c} \big(-Y(\omega)\big)\,\mathrm{d}\mathsf{Pr}(\omega) = \int_{A^c} \big(-X(\omega)\big)\,\mathrm{d}\mathsf{Pr}(\omega) \le \int_{A^c} |X(\omega)|\,\mathrm{d}\mathsf{Pr}(\omega).$$

よって

$$\begin{aligned}
\int |Y(\omega)|\,\mathrm{d}\mathsf{Pr}(\omega) &= \int_A Y(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) + \int_{A^c} \big(-Y(\omega)\big)\,\mathrm{d}\mathsf{Pr}(\omega) \\
&\le \int_A |X(\omega)|\,\mathrm{d}\mathsf{Pr}(\omega) + \int_{A^c} |X(\omega)|\,\mathrm{d}\mathsf{Pr}(\omega) \\
&= \int |X(\omega)|\,\mathrm{d}\mathsf{Pr}(\omega)
\end{aligned}$$

からわかる. □

補題 A.22. Y は一意.

Proof. Y' も定義 A.20 の条件 (i)(ii) をみたすとする. すると (ii) より

$$\int_A Y(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) = \int_A X(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) = \int_A Y'(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) \qquad (\forall A \in \mathcal{A}).$$

$\epsilon > 0$ を取り, $A := \{\omega \in \Omega;\, Y(\omega) - Y'(\omega) \geq \epsilon\}$ とする. $A \in \mathcal{F}$ であることから (ii) から

$$\begin{aligned}
0 &= \int_A X(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) - \int_A X(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) \\
&= \int_A Y(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) - \int_A Y'(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) \\
&= \int_A \big\{Y(\omega) - Y'(\omega)\big\}\,\mathrm{d}\mathsf{Pr}(\omega) \\
&\geq \epsilon \int_A \mathrm{d}\mathsf{Pr}(\omega) = \epsilon \mathsf{Pr}(A)
\end{aligned}$$

となり, $\mathsf{Pr}(A) = 0$ がわかる.

$$\Big\{\omega \in \Omega;\, Y(\omega) - Y'(\omega) > 0\Big\} = \bigcup_{n=1}^{\infty} \Big\{\omega \in \Omega;\, Y(\omega) - Y'(\omega) \geq \frac{1}{n}\Big\}$$

と表現できることとすべての $\epsilon > 0$ に対して $\mathsf{Pr}(A) = 0$ なので

$$\begin{aligned}
\mathsf{Pr}\bigg(\Big\{\omega \in \Omega;\, Y(\omega) - Y'(\omega) > 0\Big\}\bigg) &= \mathsf{Pr}\bigg(\bigcup_{n=1}^{\infty}\Big\{\omega \in \Omega;\, Y(\omega) - Y'(\omega) \geq \frac{1}{n}\Big\}\bigg) \\
&= \lim_{n\to\infty} \mathsf{Pr}\bigg(\Big\{\omega \in \Omega;\, Y(\omega) - Y'(\omega) \geq \frac{1}{n}\Big\}\bigg) \\
&= 0
\end{aligned}$$

がわかる. よって

$$\mathsf{Pr}\big(\omega \in \Omega;\, Y(\omega) \leq Y'(\omega)\big) = 1$$

となる. 次に $A := \{\omega \in \Omega;\, Y'(\omega) - Y(\omega) \geq \epsilon\}$ として, 同じ議論を行えば

$$\mathsf{Pr}\big(\omega \in \Omega;\, Y(\omega) \geq Y'(\omega)\big) = 1.$$

よって

$$\mathsf{Pr}\big(\omega \in \Omega;\, Y(\omega) = Y'(\omega)\big) = 1$$

となる. □

定義 A.23. 定義 A.20 で定めた Y を

$$Y(\omega) = \mathsf{E}\big[X|\,\mathcal{F}\big](\omega)$$

と記すことにする.

補題 A.24. $\mathsf{E}\big[X|\,\mathcal{F}\big](\omega)$ は存在する.

Proof. $\mu = \mathsf{Pr}$ と書くことにする．まず $X \geq 0$ とする．$\mathcal{F}$ 上の測度 ν を

$$\nu(A) = \int_A X(\omega)\,\mathrm{d}\mu(\omega) \qquad (A \in \mathcal{F})$$

と定める．単調収束定理を用いると ν は $\mathcal{F}$ 上の測度となる．さらに $\nu \ll \mu$ である．よって Radon-Nikodym の定理より, $\mathcal{F}$ 可測関数 $\dfrac{\mathrm{d}\nu}{\mathrm{d}\mu}$ が存在し

$$\begin{aligned}\int_A \mathsf{E}\big[X|\,\mathcal{F}\big](\omega)\,\mathrm{d}\mu(\omega) &= \int_A X(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) = \nu(A)\\ &= \int_A \frac{\mathrm{d}\nu}{\mathrm{d}\mu}(\omega)\,\mathrm{d}\mu(\omega) \qquad (A \in \mathcal{F}) \end{aligned} \tag{A.31}$$

となる．$A = \Omega$ とすると $\dfrac{\mathrm{d}\nu}{\mathrm{d}\mu} \geq 0$ は可積となる．Radon-Nikodym の一意性から

$$\mathsf{E}\big[X|\,\mathcal{F}\big](\omega) = \frac{\mathrm{d}\nu}{\mathrm{d}\mu}(\omega)$$

となる．

一般の場合については

$$\begin{aligned}&X(\omega) = X^+(\omega) - X^-(\omega),\\ &X^+(\omega) = \max\{X(\omega),\, 0\}, \quad X^-(\omega) = \max\{-X(\omega),\, 0\}\end{aligned}$$

とする．$Y_1(\omega) = \mathsf{E}\big[X^+|\,\mathcal{F}\big](\omega)$ と $Y_2(\omega) = \mathsf{E}\big[X^{-1}|\,\mathcal{F}\big](\omega)$ とおくと $Y_1 - Y_2$ は $\mathcal{F}$ 可測である．さらに $\forall A \in \mathcal{F}$ に対して

$$\begin{aligned}\int_A X(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) &= \int_A X^+(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) - \int_A X^-(\omega)\,\mathrm{d}\mathsf{Pr}(\omega)\\ &= \int_A Y_1(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) - \int_A Y_2(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) \qquad \big(\because \text{(A.31)}\big)\\ &= \int_A \left\{Y_1(\omega) - \int_A Y_2(\omega)\right\}\mathrm{d}\mathsf{Pr}(\omega)\\ &= \int_A \big(Y_1 - Y_2\big)(\omega)\,\mathrm{d}\mathsf{Pr}(\omega)\end{aligned}$$

となる．よって $\big(Y_1 - Y_2\big)(\omega) = \mathsf{E}\big[X|\,\mathcal{F}\big](\omega)$ がわかる． □

注意 A.25. 以下の記法を導入する．
(1) $A \in \mathcal{A}$ に対して

$$\mathsf{Pr}\big(A|\,\mathcal{F}\big)(\omega) := \mathsf{E}\big[\mathbb{1}_A|\,\mathcal{F}\big](\omega)$$

と定める．

(2) $A,\, B \in \mathcal{A}$ に対して

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

と定める.
(3) X と Y を $(\Omega,\, \mathcal{A},\, \mathsf{Pr})$ 上の確率変数としたとき

$$\mathsf{E}\big[X|\, Y\big](\omega) := \mathsf{E}\big[X|\, \sigma(Y)\big](\omega)$$

とする. ただし $\sigma(Y)$ は Y によって生成された σ 加法族である. すなわち

$$\sigma(Y) := \sigma\Big[\big\{\omega \in \Omega;\, Y(\omega) \le r\, (\forall r \in \mathbb{R})\big\}\Big]$$

である.
(4) X と Y を $(\Omega,\, \mathcal{A},\, \mathsf{Pr})$ 上の連続型確率変数とする. さらに $(X,\, Y)$ は同時 p.d.f. $\mathsf{p}(x,\, y)$ を持つとする. 議論を簡単にするために $\forall y \in \mathbb{R}$ に対して

$$\int \mathsf{p}(x,\, y)\, \mathrm{d}x > 0$$

とする.
いま関数 $g: \mathbb{R} \to \mathbb{R}$ で $\mathsf{E}\big[|g(X)|\big] < \infty$ なるものと考える. このとき

$$\mathsf{E}\big[g(X)|\, Y\big](\omega) = h(Y(\omega)), \qquad h(y) = \frac{\int g(x)\mathsf{p}(x,\, y)\, \mathrm{d}x}{\int \mathsf{p}(x,\, y)\, \mathrm{d}x}$$

となる. これを示すために $A \in \sigma(Y)$ を取る. このときある $B \in \mathcal{B}(\mathbb{R})$ が存在して

$$A := \big\{\omega \in \Omega;\, Y(\omega) \in B\big\}$$

と書ける. したがって Fubini の定理から

$$\begin{aligned}
\mathsf{E}\big[h(Y)\mathbb{1}_A\big] &= \mathsf{E}\big[h(Y)\mathbb{1}_B(Y)\big] \qquad \big(\because Y \in B \Leftrightarrow \omega \in A\big) \\
&= \int_B \int h(y)p(x,\, y)\, \mathrm{d}x\, \mathrm{d}y \\
&= \int_B \int \left\{\frac{\int g(x)p(x,\, y)\, \mathrm{d}x}{\int p(x,\, y)\, \mathrm{d}x}\right\} p(x,\, y)\, \mathrm{d}x\, \mathrm{d}y \qquad (\because h \text{ の定義を代入}) \\
&= \int_B \left\{\int \frac{\int g(x)p(x,\, y)\, \mathrm{d}x}{\int p(x,\, y)\, \mathrm{d}x} p(x,\, y)\, \mathrm{d}x\right\} \mathrm{d}y \qquad (\because \text{Fubini の定理}) \\
&= \int_B \frac{\int g(x)p(x,\, y)\, \mathrm{d}x}{\int p(x,\, y)\, \mathrm{d}x} \left\{\int p(x,\, y)\, \mathrm{d}x\right\} \mathrm{d}y \\
&= \int_B \int g(x)p(x,\, y)\, \mathrm{d}x\, \mathrm{d}y \\
&= \mathsf{E}\big[g(X)\mathbb{1}_A\big] \qquad \big(\because Y \in B \Leftrightarrow \omega \in A\big)
\end{aligned}$$

となる．よって

$$\int_A h\bigl(Y(\omega)\bigr)\,\mathrm{d}\mathsf{Pr}(\omega) = \int_A g\bigl(X(\omega)\bigr)\,\mathrm{d}\mathsf{Pr}(\omega)$$

なので

$$\mathsf{E}\bigl[g(X)|\,\mathcal{F}\bigr](\omega) = h\bigl(Y(\omega)\bigr)$$

がわかる． □

A.8　条件付き期待値の性質

定理 A.26. $X,\ \{X_n\}_{n=1}^{\infty},\ Y$ を確率空間 $(\Omega,\ \mathcal{A},\ \mathsf{Pr})$ 上の確率変数列とする．
(1) $\mathsf{E}\bigl[|X|\bigr] < \infty$ かつ $\mathsf{E}\bigl[|Y|\bigr] < \infty$ とする．このとき $a,\ b \in \mathbb{R}$ に対して

$$\mathsf{E}\bigl[aX + bY|\,\mathcal{F}\bigr](\omega) = a\mathsf{E}\bigl[X|\,\mathcal{F}\bigr](\omega) + b\mathsf{E}\bigl[Y|\,\mathcal{F}\bigr](\omega)$$

と成り立つ．
(2) さらに $X \le Y$ のとき

$$\mathsf{E}\bigl[X|\,\mathcal{F}\bigr](\omega) \le \mathsf{E}\bigl[Y|\,\mathcal{F}\bigr](\omega)$$

が成り立つ．
(3) $X_n \ge 0\,(n = 1,\ 2,\ \ldots)$ かつ $X_n \uparrow X\ (n \to \infty)$ で $\mathsf{E}\bigl[X\bigr] < \infty$ のとき

$$\mathsf{E}\bigl[X_n|\,\mathcal{F}\bigr](\omega) \uparrow \mathsf{E}\bigl[Y|\,\mathcal{F}\bigr](\omega) \quad (n \to \infty)$$

が成り立つ．

Proof. (1) あきらかに $a\mathsf{E}\bigl[X|\,\mathcal{F}\bigr] + b\mathsf{E}\bigl[Y|\,\mathcal{F}\bigr]$ は $\mathcal{F}$ 可測である．$\forall A \in \mathcal{F}$ に対して

$$\begin{aligned}
&\int_A \bigl\{a\mathsf{E}\bigl[X|\,\mathcal{F}\bigr](\omega) + b\mathsf{E}\bigl[Y|\,\mathcal{F}\bigr](\omega)\bigr\}\,\mathrm{d}\mathsf{Pr}(\omega)\\
&\qquad = a\int_A \mathsf{E}\bigl[X|\,\mathcal{F}\bigr](\omega)\,\mathrm{d}\mathsf{Pr}(\omega) + b\int_A \mathsf{E}\bigl[Y|\,\mathcal{F}\bigr](\omega)\,\mathrm{d}\mathsf{Pr}(\omega)\\
&\qquad = a\int_A X(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) + b\int_A Y(\omega)\,\mathrm{d}\mathsf{Pr}(\omega)\\
&\quad = \int_A \bigl(aX + bY\bigr)(\omega)\,\mathrm{d}\mathsf{Pr}(\omega)
\end{aligned}$$

を得る．

(2) 条件付き期待値の定義より $\forall A \in \mathcal{F}$ に対して

$$\begin{aligned}\int_A \mathsf{E}\big[X|\,\mathcal{F}\big](\omega)\,\mathrm{d}\mathsf{Pr}(\omega) &= \int_A X(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) \le \int_A Y(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) \\ &= \int_A \mathsf{E}\big[Y|\,\mathcal{F}\big](\omega)\,\mathrm{d}\mathsf{Pr}(\omega)\end{aligned}$$

を得る. $\forall \epsilon > 0$ に対して

$$A := \big\{\omega \in \Omega;\ \mathsf{E}\big[X|\,\mathcal{F}\big](\omega) - \mathsf{E}\big[Y|\,\mathcal{F}\big](\omega) \ge \epsilon\big\}$$

とおくと $\mathsf{Pr}(A) = 0$ となる. ϵ は任意だったので

$$\begin{aligned}&\Big\{\omega \in \Omega;\ \mathsf{E}\big[X|\,\mathcal{F}\big](\omega) > \mathsf{E}\big[Y|\,\mathcal{F}\big](\omega)\Big\} \\ &\qquad = \bigcup_{n=1}^{\infty}\Big\{\omega \in \Omega;\ \mathsf{E}\big[X|\,\mathcal{F}\big](\omega) - \mathsf{E}\big[Y|\,\mathcal{F}\big](\omega) \ge \frac{1}{n}\Big\}\end{aligned}$$

となる. よって

$$\begin{aligned}&\mathsf{Pr}\Big(\Big\{\omega \in \Omega;\ \mathsf{E}\big[X|\,\mathcal{F}\big](\omega) > \mathsf{E}\big[Y|\,\mathcal{F}\big](\omega)\Big\}\Big) \\ &\quad = \lim_{n\to\infty} \mathsf{Pr}\Big(\Big\{\omega \in \Omega;\ \mathsf{E}\big[X|\,\mathcal{F}\big](\omega) - \mathsf{E}\big[Y|\,\mathcal{F}\big](\omega) \ge \frac{1}{n}\Big\}\Big) = 0\end{aligned}$$

となるから

$$\mathsf{Pr}\Big(\omega \in \Omega;\ \mathsf{E}\big[X|\,\mathcal{F}\big](\omega) \le \mathsf{E}\big[Y|\,\mathcal{F}\big](\omega)\Big) = 1$$

となる.

(3) $Y_n := X - X_n\,(n \in \mathbb{N})$ とおくと $\mathsf{E}\big[Y_n|\,\mathcal{F}\big] \downarrow 0\,(n \to \infty)$ を示せばよい. $Y_n \downarrow 0\,(n \to \infty)$ なので (2) の結果から

$$Z_n(\omega) := \mathsf{E}\big[Y_n|\,\mathcal{F}\big](\omega)$$

はほとんど確実に単調減少列なので

$$Z_\infty(\omega) := \lim_{n\to\infty} \mathsf{E}\big[Y_n|\,\mathcal{F}\big](\omega)$$

はほとんど確実に存在する. $A \in \mathcal{F}$ に対して

$$\int_A Z_n(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) = \int_A Y_n(\omega)\,\mathrm{d}\mathsf{Pr}(\omega)$$

である. $n \to \infty$ のとき $Y_n(\omega) \downarrow 0\,(n \to \infty)$ なので, 単調収束定理から

$$\int_A Z_\infty(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) = 0 \qquad (\forall A \in \mathcal{F})$$

となる. よって $Z_\infty \equiv 0$ となる. □

定理 A.27. $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{A}$ とする. このとい

(1) $\mathsf{E}\big[\mathsf{E}\big[X|\,\mathcal{F}_1\big]|\,\mathcal{F}_2\big](\omega) = \mathsf{E}\big[X|\,\mathcal{F}_1\big](\omega)$.

(2) $\mathsf{E}\big[\mathsf{E}\big[X|\,\mathcal{F}_2\big]|\,\mathcal{F}_1\big](\omega) = \mathsf{E}\big[X|\,\mathcal{F}_1\big](\omega)$.

Proof. (1) $\mathcal{F}_1 \subset \mathcal{F}_2$ なので, $\mathsf{E}\big[X|\,\mathcal{F}_1\big]$ は $\mathcal{F}_2$ 可測でもある. よって $A \in \mathcal{F}_2$ に対して

$$\int_A \mathsf{E}\big[X|\,\mathcal{F}_1\big](\omega)\,\mathrm{d}\mathsf{Pr}(\omega) = \int_A \mathsf{E}\left[\mathsf{E}\big[X|\,\mathcal{F}_1\big]\,\middle|\,\mathcal{F}_2\right](\omega)\,\mathrm{d}\mathsf{Pr}(\omega)$$

がわかる.
(2) $\mathsf{E}\big[X|\,\mathcal{F}_1\big]$ は $\mathcal{F}_1$ 可測である. $A \in \mathcal{F}_1 \subset \mathcal{F}_2$ に対して

$$\begin{aligned}\int_A \mathsf{E}\big[X|\,\mathcal{F}_1\big](\omega)\,\mathrm{d}\mathsf{Pr}(\omega) &= \int_A X(\omega)\,\mathrm{d}\mathsf{Pr}(\omega) && \big(\because A \in \mathcal{F}_1\big)\\ &= \int_A \mathsf{E}\big[X|\,\mathcal{F}_2\big](\omega)\,\mathrm{d}\mathsf{Pr}(\omega) && \big(\because A \in \mathcal{F}_2\big)\\ &= \int_A \left\{\mathsf{E}\left[\mathsf{E}\big[X|\,\mathcal{F}_2\big]\,\middle|\,\mathcal{F}_1\right](\omega)\right\}\mathrm{d}\mathsf{Pr}(\omega) && \big(\because A \in \mathcal{F}_1\big)\end{aligned}$$

より

$$\mathsf{E}\big[X|\,\mathcal{F}_1\big](\omega) = \mathsf{E}\left[\mathsf{E}\big[X|\,\mathcal{F}_2\big]\,\middle|\,\mathcal{F}_1\right](\omega)$$

がわかる.

定理 A.28. X は $\mathcal{F}$ 可測とし, $\mathsf{E}\big[|X|\big] < \infty$ かつ $\mathsf{E}\big[|XY|\big] < \infty$ とする. このとき

$$\mathsf{E}\big[XY|\,\mathcal{F}\big](\omega) = X(\omega)\mathsf{E}\big[Y|\,\mathcal{F}\big](\omega) \tag{A.32}$$

が成立する.

Proof. $X(\omega) = \mathbb{1}_B(\omega)\,(\forall B \in \mathcal{F})$ のとき (A.32) が成り立つことを示す. $\forall A \in \mathcal{F}$ に対して

$$\begin{aligned}\int_A \mathbb{1}_B(\omega)\mathsf{E}\big[Y|\,\mathcal{F}\big](\omega)\,\mathrm{d}\mathsf{Pr}(\omega) &= \int_{A\cap B} \mathsf{E}\big[Y|\,\mathcal{F}\big](\omega)\,\mathrm{d}\mathsf{Pr}(\omega)\\ &= \int_{A\cap B} Y(\omega)\,\mathrm{d}\mathsf{Pr}(\omega)\\ &= \int_A \mathbb{1}_B(\omega)Y(\omega)\,\mathrm{d}\mathsf{Pr}(\omega)\end{aligned}$$

となる. よって $X(\omega) = \mathbb{1}_B(\omega)$ のとき (A.32) は 成立する.

次に $X, Y \geq 0$ とし, X_n は階段関数とし, $X_n \uparrow X\,(n \to \infty)$ とする. 単調収束定理から

$$\begin{aligned}\int_A X(\omega)\mathsf{E}\big[Y|\,\mathcal{F}\big](\omega)\,\mathrm{d}\mathsf{Pr}(\omega) &= \lim_{n\to\infty}\int_A X_n(\omega)\mathsf{E}\big[Y|\,\mathcal{F}\big](\omega)\,\mathrm{d}\mathsf{Pr}(\omega)\\ &= \lim_{n\to\infty}\int_A \mathsf{E}\big[X_nY|\,\mathcal{F}\big](\omega)\,\mathrm{d}\mathsf{Pr}(\omega) \qquad \big(\because (1)\text{ の結果}\big)\\ &= \int_A \lim_{n\to\infty}\mathsf{E}\big[X_nY|\,\mathcal{F}\big](\omega)\,\mathrm{d}\mathsf{Pr}(\omega) \qquad \big(\because \text{単調収束定理}\big)\\ &= \int_A \lim_{n\to\infty}\mathsf{E}\big[XY|\,\mathcal{F}\big](\omega)\,\mathrm{d}\mathsf{Pr}(\omega) \qquad \big(\because \text{定理 A.26(3)}\big)\end{aligned}$$

がわかり, $X, Y \geq 0$ のとき (A.32) は成立する.

最後に一般の X, Y に対して

$$X^+ = \max\{X,\,0\}, \quad X^{-1} = \max\{-X,\,0\}, \quad Y^+ = \max\{Y,\,0\}, \quad Y^{-1} = \max\{-Y,\,0\}$$

として, 上の結果を用いればよい. □

A.9 Neyman-Fisher の因子分解定理の証明 (Keener から?)

Radon-Nikodym の定理から導かれることを用いて Fisher-Neyman の因子分解定理の測度論的な証明を与える.

補題 A.29. μ を可測空間 $(\mathbb{X},\,\mathcal{A})$ 上の σ 有限測度とし, $\mathcal{N}$ を $(\mathbb{X},\,\mathcal{A})$ 上の測度の集合で, $\forall\nu \in \mathcal{N}$ に対して, $\nu \ll \mu$ をみたしているとする. このとき非負の数列 $\{c_n\}_{n=1}^\infty$ で $\sum_{n=1}^\infty c_n = 1$ なるものと $\{\nu_n\}_{n=1}^\infty \subset \mathcal{N}$ が存在して $\forall\nu \in \mathcal{N}$ に対して $\nu \ll \sum_{n=1}^\infty c_n\nu_n$ となる.

Proof. $\mathcal{N}$ が可算集合ならば定理の主張は自明であるので, $\mathcal{N}$ は非可算とする.

μ は有限測度の場合は

$$\lambda = \mu$$

とおく. μ が有限測度でない場合には $\mathbb{X}$ の可算個の部分集合 $\{\mathbb{X}_n\}_{n=1}^\infty$ をうまくとると

$$\mathbb{X} = \bigcup_{n=1}^\infty \mathbb{X}_n; \qquad 0 < \mu(\mathbb{X}_n) = d_n < \infty$$

とできる. $\forall B \in \mathcal{A}$ に対して

$$\lambda(B) := \sum_{n=1}^\infty \frac{\mu(B \cap \mathbb{X}_n)}{2^n d_n}$$

とおく．いずれの場合でも λ は有限測度で $\nu \ll \lambda\,(\forall \nu \in \mathcal{N})$ となる [4]．ここで

$$\mathcal{Q} := \left\{ \sum_{n=1}^{\infty} c_n \nu_n ;\ \sum_{n=1}^{\infty} c_n = 1\, c_n \in \mathbb{R},\ \text{かつ}\ \{\nu_n\}_{n=1}^{\infty} \subset \mathcal{N} \right\}$$

とする．明らかに

$$\beta \in \mathcal{Q} \Rightarrow \beta \ll \lambda$$

である．

つぎに

$$\mathcal{D} := \left\{ C \in \mathcal{A};\ \exists \mathsf{Q} \in \mathcal{Q}\ \text{s.t.}\ \lambda\left(\left\{ x \in C;\ \frac{\mathrm{d}\mathsf{Q}}{\mathrm{d}\lambda}(x) = 0 \right\} \right) = 0\ \text{かつ}\ \mathsf{Q}(C) > 0 \right\}$$

とする．まず $\mathcal{D} \neq \varnothing$ を確認する．そのために $\nu \in \mathcal{N}$ で $\nu \neq 0$ なるものを取り

$$C := \left\{ x \in \mathbb{X};\ \frac{\mathrm{d}\nu}{\mathrm{d}\lambda}(x) > 0 \right\}$$

とする．$\mathsf{Q} = \nu$ としたとき

$$\left\{ x \in C;\ \frac{\mathrm{d}\nu}{\mathrm{d}\lambda}(x) = 0 \right\} = \varnothing$$

となり

$$Q(C) = \nu(C) = \mu(\mathbb{X}) > 0$$

となる．よって $C \in \mathcal{D}$ となる．λ は有限測度なので

$$\sup_{C \in \mathcal{D}} \lambda(C) =: c < \infty$$

となる．よって $\{C_n\}_{n=1}^{\infty} \subset \mathcal{D}$ かつ $C_1 \subset C_2 \subset \cdots$ をうまくとると

$$\lim_{n \to \infty} \lambda(C_n) = c$$

とできる．ここで

$$C_0 := \bigcup_{n=1}^{\infty} C_n$$

[4] $\lambda(A) = 0$ なる $A \in \mathcal{A}$ をとる．すると $\mu(A \cap \mathbb{X}_n) = 0\,(\forall n \in \mathbb{N})$ となる．したがって

$$\mu(A) = \sum_{n=1}^{\infty} \mu(A \cap \mathbb{X}_n) = 0$$

となる．$\nu \ll \mu$ から $\nu(A) = 0$ となる．よって $\lambda \ll \mu$ がわかる．

とおき, $\mathsf{Q}_n \in \mathcal{Q}\,(n \in \mathbb{N})$ で

$$\mathsf{Q}_n(C_n) > 0 \text{ かつ } \lambda\left(\left\{x \in C_n;\ \frac{\mathrm{d}\mathsf{Q}_n}{\mathrm{d}\lambda}(x) = 0\right\}\right) = 0$$

なるもの[5]を取る. さらに

$$\mathsf{Q}_0 := \sum_{n=1}^{\infty} \frac{\mathsf{Q}_n}{2^n} \in \mathcal{Q} \qquad \left(\because \sum_{n=1}^{\infty} \frac{1}{2^n} = 1\right)$$

とおく. すると

$$\frac{\mathrm{d}\mathsf{Q}_0}{\mathrm{d}\lambda} = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{\mathrm{d}\mathsf{Q}_n}{\mathrm{d}\lambda}$$

かつ

$$\left\{x \in C_0;\ \frac{\mathrm{d}\mathsf{Q}_0}{\mathrm{d}\lambda}(x) = 0\right\} \subset \bigcup_{n=1}^{\infty} \left\{x \in C_n;\ \frac{\mathrm{d}\mathsf{Q}_n}{\mathrm{d}\lambda}(x) = 0\right\}$$

となる. このことより

$$\lambda\left(\left\{x \in C_0;\ \frac{\mathrm{d}\mathsf{Q}_0}{\mathrm{d}\lambda}(x) = 0\right\}\right) \leq \sum_{n=1}^{\infty} \lambda\left(\left\{x \in C_n;\ \frac{\mathrm{d}\mathsf{Q}_n}{\mathrm{d}\lambda}(x) = 0\right\}\right)$$

となるで, $C_0 \in \mathcal{D}$ がわかる. さらに

$$\lambda(C_0) = \lambda\left(\bigcup_{n=1}^{\infty} C_n\right) = \lim_{n\to\infty} \lambda(C_n) = c$$

がわかる.

あとは $\mathsf{Q}_0 \in \mathcal{Q}$ なので $\forall \nu \in \mathcal{N}$ に対して $\nu \ll Q_0$ を示せばよい. そのために $A \in \mathcal{A}$ で $\mathsf{Q}_0(A) = 0$ なるものと $\forall \nu \in \mathcal{N}$ を取る. $x \in C_0$ に対して

$$\mathsf{Q}_0(A \cap C_0) \leq \mathsf{Q}_0(A) = 0 \text{ かつ } \frac{\mathrm{d}\mathsf{Q}_0}{\mathrm{d}\lambda}(x) > 0\,(\forall x \in C_0)$$

となる[6]ので

$$\lambda(A \cap C_0) = 0 \Rightarrow \nu(A \cap C_0) = 0 \qquad (\because \nu \ll \lambda)$$

---

[5] $C_n \in \mathcal{D}$ なのでこのような $\mathsf{Q}_n \in \mathcal{Q}$ が取れる.

[6] $x \in C_0$ のときある $n \in \mathbb{N}$ があって $x \in C_n$ となる. すると

$$\lambda\left(\left\{x \in C_n;\ \frac{\mathrm{d}\mathsf{Q}_n}{\mathrm{d}\lambda}(x) = 0\right\}\right) = 0$$

なので

$$\frac{\mathrm{d}\mathsf{Q}_0}{\mathrm{d}\lambda}(x) > 0$$

となる.

となる．ここで

$$C := \left\{ x \in \mathbb{X};\ \frac{\mathrm{d}\nu}{\mathrm{d}\lambda}(x) > 0 \right\}$$

とおく．すると

$$x \in C^c \ \Rightarrow\ \frac{\mathrm{d}\nu}{\mathrm{d}\lambda}(x) = 0$$

なので

$$\nu(A \cap C_0^c \cap C^c) \leq \int_{C^c} \frac{\mathrm{d}\nu}{\mathrm{d}\lambda}(x)\,\mathrm{d}\lambda(x) = 0$$

となる．さらに $D := A \cap C_0^c \cap C$ とおくと $D \cap C_0 = \varnothing$ となる．$\lambda(D) > 0$ のとき

$$\lambda(C_0 \cap D) > \lambda(D)$$

であり

$$Q_0\left(\left\{ x \in D;\ \frac{\mathrm{d}Q_0}{\mathrm{d}\lambda}(x) = 0 \right\} \right) = 0$$

となる[7]となるので, $D \in \mathcal{D}$ となる．明らかに $C_0 \cup D \in \mathcal{D}$ であり

$$\lambda(C_0 \cup D) > \lambda(C_0) = c$$

となるので, c の定義[8]と矛盾する．しがたって

$$\lambda(D) = 0 \text{ かつ } \nu(D) = 0$$

となる．このことより

$$\begin{aligned}
\nu(A) &= \nu(A \cap C_0) + \nu(A \cap C_0^c) \\
&= \nu(A \cap C_0) + \nu(A \cap C_0^c \cap C^c) + \nu(A \cap C_0^c \cap C) \\
&= \nu(A \cap C_0) + \nu(A \cap C_0^c \cap C^c) + \nu(D) = 0
\end{aligned}$$

がわかる．よって定理が証明された． □

[7]

$$\mathsf{Q}_0 \in \mathcal{Q} \ \Rightarrow\ \mathsf{Q}_0 \ll \lambda$$

であることからわかる．

[8] $c = \sup_{C \in \mathcal{D}} \lambda(C)$ である．

定理 A.30. ν を可測空間 $(\mathbb{X}, \mathcal{A})$ 上の σ 有限測度とし,

$$\mathcal{P} = \{\mathsf{P};\ \mathsf{P} \text{ は } (\mathbb{X}, \mathcal{A}) \text{ 上の確率分布で}, \mathsf{P} \ll \nu\}$$

を統計的モデルとする. X を母集団分布 $\mathsf{P} \in \mathcal{P}$ からの標本とし, $T(X)$ を統計量とする. このとき $T(X)$ は $\mathcal{P}$ に対する十分統計量であるための必要十分条件は, $(\mathbb{X}, \mathcal{A})$ 上の可測関数 h (これは $\mathsf{P} \in \mathcal{P}$ に依存しない) と T の値域上の可測関数 g_P (これは $\mathsf{P} \in \mathcal{P}$ に依存する) が存在して

$$\frac{\mathrm{d}\mathsf{P}}{\mathrm{d}\nu}(x) = g_{\mathsf{P}}(T(x))h(x), \qquad \mathsf{P} - a.s. \tag{A.33}$$

と書けることである.

Proof. $\underline{(\Rightarrow)\text{ の証明}}$: T は $\mathcal{P}$ に対する十分統計量とする. 補題 A.29 から, $\{\mathsf{P}_n\}_{n=1}^{\infty} \subset \mathcal{P}$ と数列 $\{c_n\}_{n=1}^{\infty}$ $(\sum_{n=1}^{\infty} c_n = 1)$ をうまく取れば

$$\mathsf{P} \ll \mathsf{Q} = \sum_{n=1}^{\infty} c_n \mathsf{P}_n$$

とできる. ここで $\dfrac{d\mathsf{Q}}{d\nu}$ は $\mathsf{P} \in \mathcal{P}$ に依存しないことに注意する. $\mathsf{E}[\cdot]$ と $\mathsf{E}[\cdot|T]$ を P に関する期待値と条件付き期待値とし, $\mathsf{E}_{\mathsf{Q}}[\cdot]$, $\mathsf{E}_{\mathsf{P}_n}[\cdot]$ と $\mathsf{E}_{\mathsf{Q}}[\cdot|T]$, $\mathsf{E}_{\mathsf{P}_n}[\cdot|T]$ を Q, P_n それぞれにに関する期待値と条件付き期待値とする. T は十分統計量であるので T を与えたときの X の条件付き分布は $\mathsf{P} \in \mathcal{P}$ に依存しないので

$$\mathsf{E}[\mathbb{1}_A(X)|T] = \mathsf{E}_{\mathsf{P}_n}[\mathbb{1}_A(X)|T] \qquad (\forall A \in \mathcal{A}) \tag{A.34}$$

が成立することに注意する. $\forall A \in \mathcal{A}$ と $B \in \sigma(T)$ に対して

$$
\begin{aligned}
\mathsf{Q}(A \cap B) &= \sum_{n=1}^{\infty} c_n \mathsf{P}_n(A \cap B) = \sum_{n=1}^{\infty} c_n \mathsf{E}_{\mathsf{P}_n}[\mathbb{1}_{A\cap B}(X)] \\
&= \sum_{n=1}^{\infty} c_n \mathsf{E}_{\mathsf{P}_n}\big[\mathsf{E}_{\mathsf{P}_n}[\mathbb{1}_{A\cap B}(X)|\,T]\big] \qquad \big(\because \text{条件付き期待値の性質}\big) \\
&= \sum_{n=1}^{\infty} c_n \mathsf{E}_{\mathsf{P}_n}\big[\mathsf{E}_{\mathsf{P}_n}[\mathbb{1}_{A}(X)\mathbb{1}_{B}(X)|\,T]\big] \\
&= \sum_{n=1}^{\infty} c_n \mathsf{E}_{\mathsf{P}_n}\big[\mathbb{1}_{B}(X)\mathsf{E}_{\mathsf{P}_n}[\mathbb{1}_{A}(X)|\,T]\big] \qquad \big(\because \mathbb{1}_B(X) \text{ は } \sigma(T) \text{ 可測}\big) \\
&= \sum_{n=1}^{\infty} c_n \mathsf{E}_{\mathsf{P}_n}\big[1_{B}(X)\mathsf{E}[\mathbb{1}_{A}(X)|\,T]\big] \qquad \big(\because \text{(A.34)}\big) \\
&= \sum_{n=1}^{\infty} c_n \int_B \Big\{\mathsf{E}[\mathbb{1}_{A}(X)|\,T](x)\Big\}\,\mathrm{d}\mathsf{P}_n \\
&= \int_B \Big\{\mathsf{E}[\mathbb{1}_{A}(X)|\,T](x)\Big\} \sum_{n=1}^{\infty} c_n\,\mathrm{d}\mathsf{P}_n \qquad \big(\because \text{有界収束定理}\big) \\
&= \int_B \Big\{\mathsf{E}[\mathbb{1}_{A}(X)|\,T](x)\Big\}\,\mathrm{d}\mathsf{Q}(x)
\end{aligned}
$$

となる. 一方

$$
\mathsf{Q}(A \cap B) = \int_B \mathbb{1}_A(x)\,\mathrm{d}\mathsf{Q}(x) = \int_B \Big\{\mathsf{E}_Q[\mathbb{1}_{A}(X)|\,T](x)\Big\}\,\mathrm{d}\mathsf{Q}(x)
$$

となるので

$$
\mathsf{E}[\mathbb{1}_{A}(X)|\,T](x) = \mathsf{E}_{\mathsf{Q}}[\mathbb{1}_{A}(X)|\,T](x), \qquad \mathsf{Q}-a.s.
$$

を得る. いま

$$
\frac{\mathrm{d}\mathsf{P}}{\mathrm{d}\mathsf{Q}}(x) =: g_P\big(T(x)\big)
$$

とおくと $\forall A \in \mathcal{A}$ に対して

$$\begin{aligned}
\mathsf{P}(A) &= \mathsf{E}[\mathbb{1}_A(X)] = \mathsf{E}\big[\mathsf{E}[\mathbb{1}_A(X)|\,T]\big] = \int \left\{ \mathsf{E}[\mathbb{1}_A(X)|\,T](x) \right\} \mathrm{d}\mathsf{P}(x) \\
&= \int \left\{ \mathsf{E}_Q[\mathbb{1}_A(X)|\,T](x)\, \frac{\mathrm{d} mathsfP}{\mathrm{d}\mathsf{Q}}(x) \right\} \mathrm{d}\mathsf{Q}(x) \\
&= \int \left\{ \mathsf{E}_{\mathsf{Q}}[\mathbb{1}_A(X)|\,T](x)\, g_{\mathsf{P}}(T(x)) \right\} \mathrm{d}\mathsf{Q}(x) \\
&= \int \left\{ \mathsf{E}_{\mathsf{Q}}[\mathbb{1}_A(X) g_{\mathsf{P}}(T(X))|\,T](x) \right\} \mathrm{d}\mathsf{Q}(x) \qquad \big(\because g_{\mathsf{P}}(T(x)) \text{ は } \sigma(T) \text{ 可測}\big) \\
&= \mathsf{E}_{\mathsf{Q}}\big[\mathbb{1}_A(X) g_{\mathsf{P}}(T(X))\big] \qquad \big(\because \text{条件付き期待値の定義}\big) \\
&= \int \left\{ \mathbb{1}_A(x) g_{\mathsf{P}}(T(x)) \right\} \mathrm{d}\mathsf{Q}(x) \\
&= \int_A \left\{ g_{\mathsf{P}}(T(x)) \frac{\mathrm{d}\mathsf{Q}}{\mathrm{d}\nu}(x) \right\} \mathrm{d}\nu(x)
\end{aligned}$$

となる．よって

$$h = \frac{\mathrm{d}\mathsf{Q}}{\mathrm{d}\nu}$$

ととればよいことがわかる．

<u>($\Leftarrow$) の証明</u>: (A.33) が成り立つとする．T が $\mathcal{P}$ に対する十分統計量であることを示すために

$$\begin{aligned}
&\mathsf{P}(A|\,T) = \mathsf{E}_{\mathsf{Q}}[\mathbb{1}_A(X)|\,T] \qquad \mathsf{P}-a.s. \\
\Leftrightarrow\ &\int_B \left\{ \mathsf{P}(A|\,T)(x) \right\} \mathrm{d}\mathsf{P}(x) = \int_B \left\{ \mathsf{E}_{\mathsf{Q}}[\mathbb{1}_A(X)|\,T](x) \right\} \mathrm{d}\mathsf{P}(x) \quad \big(\forall B \in \sigma(T)\big)
\end{aligned} \tag{A.35}$$

を示せばよい．Radon-Nikodym の微分の性質と (A.33) から

$$\begin{aligned}
\frac{d\mathsf{P}}{d\mathsf{Q}}(x) &= \frac{\mathrm{d}\mathsf{P}}{\mathrm{d}\nu}(x) \Big/ \frac{\mathrm{d}\mathsf{Q}}{\mathrm{d}\nu}(x) = \frac{\mathrm{d}\mathsf{P}}{\mathrm{d}\nu}(x) \Big/ \sum_{n=1}^{\infty} c_n \frac{\mathrm{d}\mathsf{P}_n}{\mathrm{d}\nu}(x) \\
&= g_{\mathsf{P}}(T(x)) h(x) \Big/ \sum_{n=1}^{\infty} c_n g_{\mathsf{P}_n}(T(x)) h(x) \\
&= g_{\mathsf{P}}(T(x)) \Big/ \sum_{n=1}^{\infty} c_n g_{\mathsf{P}_n}(T(x))
\end{aligned}$$

となることに注意する．このことより $\dfrac{d\mathsf{P}}{d\mathsf{Q}}(x)$ は $\sigma(T)$ 可測であることが

わかる. $\forall B \in \sigma(T)$ に対して

$$\begin{aligned}
\int_B \left\{ \mathsf{E}_{\mathsf{Q}}[\mathbb{1}_A(X)|\,T](x) \right\} \mathrm{d}\mathsf{P}(x) &= \int_B \left\{ \mathsf{E}_{\mathsf{Q}}[\mathbb{1}_A(X)|\,T](x) \right\} \mathrm{d}\mathsf{Q}(x) \\
&= \int_B \left\{ \mathsf{E}_{\mathsf{Q}}\left[\mathbb{1}_A(X) \frac{\mathrm{d}\mathsf{P}}{\mathrm{d}\mathsf{Q}}(X) \middle| T \right](x) \right\} \mathrm{d}\mathsf{Q}(x) \\
&= \int_B \left\{ \mathbb{1}_A(x) \frac{\mathrm{d}\mathsf{P}}{\mathrm{d}\mathsf{Q}}(x) \right\} \mathrm{d}\mathsf{Q}(x) \qquad (\text{条件付き期待値の定義}) \\
&= \int_B \mathbb{1}_A(x)\, \mathrm{d}\mathsf{P}(x) = \int_B \left\{ \mathsf{E}[\mathbb{1}_A|\,T](x) \right\} \mathrm{d}\mathsf{P}(x) \\
&= \int_A \left\{ \mathsf{P}(A|\,T)(x) \right\} \mathrm{d}\mathsf{P}(x)
\end{aligned}$$

となる. よって (A.35) が示されたので, T は $\mathcal{P}$ に対する十分統計量となることがわかる. □

A.10 Dini の定理を加筆

A.11 最尤推定量の漸近分布の証明 (定理 12.11)