12月10日
ここから

不等式(6.2)の応用として,次を証明しよう.

Thm6.3 (Weierstrass(ヴァイエルシュトラス)の多項式近似定理)
$a,b\in\mathbb{R}$, $a<b$ とし, $f:[a,b]\to\mathbb{R}$ は連続とする.このとき
$\forall\varepsilon\in(0,\infty)$, $\exists P(x)=\sum_{k=0}^n a_k x^k$ ($n\in\mathbb{N}\cup\{0\}$, $\{a_k\}_{k=0}^n\subset\mathbb{R}$),
$$\sup_{x\in[a,b]}|f(x)-P(x)|<\varepsilon.$$

∵ $a=0$, $b=1$ と仮定して証明すればよい.
(実際,そのとき一般の $f:[a,b]\to\mathbb{R}$ の場合でも, $g:[0,1]\to\mathbb{R}$ を $g(x):=f(a+x(b-a))$ で定めることにより, $\forall\varepsilon\in(0,\infty)$, $\exists P(x)$ 多項式, $\sup_{x\in[a,b]}\left|f(x)-P\left(\frac{x-a}{b-a}\right)\right|=\sup_{x\in[0,1]}|g(x)-P(x)|<\varepsilon$.)
例4.15のように, $p\in[0,1]$ とし,また $n\in\mathbb{N}$ とし $\{X_k\}_{k=1}^n$ を独立な real r.v.s で $\forall k\in\{1,\cdots,n\}$, $X_k\sim Be(p)=B(1,p)$ となるものとする. $S_n:=\sum_{k=1}^n X_k$ とおくと, $S_n\sim B(n,p)$ なので
$$\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right]=\sum_{k=0}^n f\left(\frac{k}{n}\right)\mathbb{P}[S_n=k]=\sum_{k=0}^n f\left(\frac{k}{n}\right)\binom{n}{k}p^k(1-p)^{n-k}=:B_{f,n}(p). \tag{6.3}$$
($B_{f,n}(p)$ を f に対する n 次Bernstein多項式という.)
claim $\lim_{n\to\infty}\sup_{p\in[0,1]}|f(p)-B_{f,n}(p)|=0$.
∵ $\varepsilon\in(0,\infty)$ とする.([0,1]のコンパクト性により) f は $[0,1]$ 上で一様連続であるので, $\exists\delta\in(0,\infty)$,
$$\forall x,y\in[0,1],\ |x-y|<\delta\Longrightarrow|f(x)-f(y)|<\varepsilon. \tag{6.4}$$
また([0,1]のコンパクト性により) $M:=\sup_{x\in[0,1]}|f(x)|=\max_{x\in[0,1]}|f(x)|<\infty$.
そこで(6.3),(6.4),(6.2)を用いれば,
$$\begin{aligned}|f(p)-B_{f,n}(p)|&=\left|\mathbb{E}\left[f(p)-f\left(\tfrac{S_n}{n}\right)\right]\right|\le\mathbb{E}\left[\left|f(p)-f\left(\tfrac{S_n}{n}\right)\right|\right]\\&=\mathbb{E}\left[\left|f(p)-f\left(\tfrac{S_n}{n}\right)\right|\left(\mathbb{1}_{\{|S_n/n-p|<\delta\}}+\mathbb{1}_{\{|S_n/n-p|\ge\delta\}}\right)\right]\\&\le\varepsilon+2M\,\mathbb{P}\left[\left|\tfrac{S_n}{n}-p\right|\ge\delta\right]\\&\le\varepsilon+2M\delta^{-2}\frac{p(1-p)}{n}\le\varepsilon+2M\delta^{-2}\frac{1}{n}\end{aligned}$$
となり,これは $n\ge 2M\delta^{-2}\frac{1}{\varepsilon}$ であるような任意の $n\in\mathbb{N}$ に対し $\sup_{p\in[0,1]}|f(p)-B_{f,n}(p)|\le 2\varepsilon$ であることを意味する//(claim) ■

実は Thm6.2は次のように $\xrightarrow{P}$ を $\xrightarrow{a.s.}$ に強められる!

Thm6.4 (大数の強法則(SLLN, strong law of large numbers))
$m\in\mathbb{R}$ とし, $\{X_n\}_{n=1}^\infty\subset\mathcal{L}^2(\mathbb{P})$ は独立で $\forall n\in\mathbb{N}$, $\mathbb{E}[X_n]=m$, かつ $\sup_{n\in\mathbb{N}}var(X_n)<\infty$ と仮定する.このとき
$$\frac{1}{n}\sum_{k=1}^n X_k\xrightarrow{a.s.}m. \tag{6.5}$$
特に,任意の i.i.d. $\{X_n\}_{n=1}^\infty\subset\mathcal{L}^2(\mathbb{P})$ に対し(6.5)が成り立つ.

∵ $v:=\sup_{n\in\mathbb{N}}var(X_n)$ とおき, $n\in\mathbb{N}$ に対し $Y_n:=\frac{1}{n}\sum_{k=1}^n X_k-m$ とおく. (6.2)により $\mathbb{E}[Y_n^2]\le\frac{v}{n}$ であるので,
$$\mathbb{E}\left[\sum_{n=1}^\infty Y_{n^2}^2\right]\overset{\text{Prop1.20}}{=}\sum_{n=1}^\infty\mathbb{E}[Y_{n^2}^2]\le\sum_{n=1}^\infty\frac{v}{n^2}<\infty.$$
従ってProp1.27-(2)により a.s.に $\sum_{n=1}^\infty Y_{n^2}^2<\infty$ であり,特に $Y_{n^2}^2\xrightarrow{a.s.}0$, すなわち $Y_{n^2}\xrightarrow{a.s.}0$.
次に $n\in\mathbb{N}$ とし $k(n):=\max\{k\in\mathbb{N}\mid k^2\le n\}$ とおく.すると $k(n)^2\le n<(k(n)+1)^2$ より $\sqrt{n}-1<k(n)\le\sqrt{n}$ であるので,特に $\lim_{n\to\infty}k(n)=\infty$, $\lim_{n\to\infty}\frac{k(n)^2}{n}=1$. さらに $n\in\mathbb{N}$ に対し,
$Y_n-\frac{k(n)^2}{n}Y_{k(n)^2}=\frac{1}{n}\sum_{l=k(n)^2+1}^n(X_l-m)$ なので
$$\begin{aligned}\mathbb{E}\left[\left(Y_n-\frac{k(n)^2}{n}Y_{k(n)^2}\right)^2\right]&=\frac{1}{n^2}var\left(\sum_{l=k(n)^2+1}^n X_l\right)\\&\overset{\text{Prop4.14-(2)}}{=}\frac{1}{n^2}\sum_{l=k(n)^2+1}^n var(X_l)\\&\le\frac{n-k(n)^2}{n^2}v\le\frac{2k(n)+1}{n^2}v\\&\le\frac{2\sqrt{n}+1}{n^2}v\le 3vn^{-3/2}.\end{aligned}$$
よって
$$\mathbb{E}\left[\sum_{n=1}^\infty\left(Y_n-\frac{k(n)^2}{n}Y_{k(n)^2}\right)^2\right]\overset{\text{Prop1.20}}{=}\sum_{n=1}^\infty\mathbb{E}\left[\left(Y_n-\frac{k(n)^2}{n}Y_{k(n)^2}\right)^2\right]\le\sum_{n=1}^\infty\frac{3v}{n^{3/2}}<\infty$$
なので再びProp1.27-(2)により $\sum_{n=1}^\infty\left(Y_n-\frac{k(n)^2}{n}Y_{k(n)^2}\right)^2<\infty$ a.s., 従って特に $\left(Y_n-\frac{k(n)^2}{n}Y_{k(n)^2}\right)^2\xrightarrow{a.s.}0$, すなわち
$Y_n-\frac{k(n)^2}{n}Y_{k(n)^2}\xrightarrow{a.s.}0$. ゆえに前段落の結果と合わせて
$$Y_n=\left(Y_n-\frac{k(n)^2}{n}Y_{k(n)^2}\right)+\frac{k(n)^2}{n}Y_{k(n)^2}\xrightarrow{a.s.}0+1\cdot 0=0.$$
よって $\frac{1}{n}\sum_{k=1}^n X_k=Y_n+m\xrightarrow{a.s.}0+m=m$. ■

例6.5 (1) $p\in[0,1]$ とし, $\{X_n\}_{n=1}^\infty$ は i.i.d. real r.v.s で $X_1\sim Be(p)$ とする(「表の確率 p のコインを無限回投げる」).
$\mathbb{E}[X_1]=\mathbb{E}[X_1^2]=p$ なので Thm6.4により
(n 回のうち表の回数の割合) $=\frac{1}{n}\sum_{k=1}^n X_k\xrightarrow{a.s.}p$.
(2) $\{X_n\}_{n=1}^\infty$ は i.i.d. real r.v.s で $\forall k\in\{1,2,3,4,5,6\}$, $\mathbb{P}[X_1=k]=\frac{1}{6}$ とする.このとき $\forall k\in\{1,2,3,4,5,6\}$ に対し, Prop4.13-(2)より $\{\mathbb{1}_{\{k\}}(X_n)\}_{n=1}^\infty$ は独立であり, $\forall n\in\mathbb{N}$, $\mathbb{1}_{\{k\}}(X_n)\sim Be(\frac{1}{6})$ であるので(1)により
(サイコロを n 回投げたときの, k の目の回数の割合)
$$=\frac{1}{n}\#\{j\in\{1,\cdots,n\}\mid X_j=k\}=\frac{1}{n}\sum_{j=1}^n\mathbb{1}_{\{k\}}(X_j)\xrightarrow{a.s.}\frac{1}{6}.$$

i.i.d.に対しては, Thm6.4よりさらに強く,次が成り立つ.

Thm6.6 (i.i.d. r.v.s $\subset\mathcal{L}^1(\mathbb{P})$ に対する大数の強法則)
$\{X_n\}_{n=1}^\infty$ を i.i.d. real r.v.s とする.このとき次が成り立つ:

● $\mathbb{E}[|X_1|]<\infty \Longrightarrow \frac{1}{n}\sum_{k=1}^n X_k \xrightarrow{a.s.} \mathbb{E}[X_1]$.

● $\mathbb{E}[|X_1|]=\infty \Longrightarrow$ a.s.に $\left\{\frac{1}{n}\sum_{k=1}^n X_k\right\}_{n=1}^\infty$ は$\mathbb{R}$において収束しない.

(☺ ノート Thm3.61)

さらにi.i.d.の場合にはThm6.4は次のように精密化できる:

Thm6.7 (重複対数の法則 (LIL, law of the iterated logarithm))
$\{X_n\}_{n=1}^\infty\subset\mathcal{L}^2(\mathbb{P})$ はi.i.d.とし, $m:=\mathbb{E}[X_1]$, $v:=var(X_1)$ とおく. このとき

$$\limsup_{n\to\infty}\frac{S_n-nm}{\sqrt{2n\log\log n}}=\sqrt{v},\quad \liminf_{n\to\infty}\frac{S_n-nm}{\sqrt{2n\log\log n}}=-\sqrt{v},\ \text{a.s.}$$

$\left(S_n:=\sum_{k=1}^n X_k\right)$

(証明は非常に難しい(ので省略せざるを得ない).)

§7 分布の収束 ($d\in\mathbb{N}$とする.)

Def7.1 $\mathcal{P}(\mathbb{R}^d):=\{\mu\mid\mu$は$\mathbb{R}^d$上の分布$\}$,
$C_b(\mathbb{R}^d):=\{f\mid f:\mathbb{R}^d\to\mathbb{R}, f$は有界連続$\}$ と定める.

以降で主に扱うのは, $\mathcal{P}(\mathbb{R}^d)$における次の収束概念である.

Def7.2 (分布の収束) $\mu\in\mathcal{P}(\mathbb{R}^d)$, $\{\mu_n\}_{n=1}^\infty\subset\mathcal{P}(\mathbb{R}^d)$ とする.
$\{\mu_n\}_{n=1}^\infty$ が μ に(弱)収束する ($\mu_n\xrightarrow{\mathcal{L}}\mu$)
$\overset{def}{\Longleftrightarrow}$ $\forall f\in C_b(\mathbb{R}^d)$, $\lim_{n\to\infty}\int_{\mathbb{R}^d} f\,d\mu_n=\int_{\mathbb{R}^d} f\,d\mu$.

● d-dim. r.v.s Xと$\{X_n\}_{n=1}^\infty$に対し, (cf. Def5.1-(3))

$$X_n\xrightarrow{\mathcal{L}}X \iff \mathcal{L}(X_n)\xrightarrow{\mathcal{L}}\underbrace{\mathcal{L}(X)}_{X\text{の分布}}$$

Def7.2の状況で, 別の収束の仕方を考えることができない訳ではない. 例えば, 次の収束はより自然に思えるかもしれない:

$$\forall A\in\mathcal{B}(\mathbb{R}^d),\ \lim_{n\to\infty}\mu_n(A)=\mu(A).\ \cdots\cdots(7.1)$$

後のThm7.11でみるように, (7.1)は $\mu_n\xrightarrow{\mathcal{L}}\mu$ より強い条件になっている. 実際には, (7.1)は強すぎて現象の記述に使えないことが多いため, あまり有用な収束概念ではない. このことは次の例7.3からも見てとれる.

例7.3 $x\in\mathbb{R}^d$とし, $A\in\mathcal{B}(\mathbb{R}^d)$に対し $\delta_x(A):=\mathbb{1}_A(x)$ と定めると, 容易に分かるように $\delta_x\in\mathcal{P}(\mathbb{R}^d)$ である. さて, $x\in\mathbb{R}^d$, $\{x_n\}_{n=1}^\infty\subset\mathbb{R}^d$ とする. このとき次が成り立つ.

(1) $\delta_{x_n}\xrightarrow{\mathcal{L}}\delta_x \iff \lim_{n\to\infty}x_n=x$.

(2) $\forall A\in\mathcal{B}(\mathbb{R}^d)$, $\lim_{n\to\infty}\delta_{x_n}(A)=\delta_x(A)$
$\iff \exists N\in\mathbb{N}, \forall n\geq N, x_n=x$.

☺ (1) ($\Longleftarrow$) $f\in C_b(\mathbb{R}^d)$ とすると fの連続性により
$\int_{\mathbb{R}^d}f(y)\delta_{x_n}(dy)=f(x_n)\xrightarrow{n\to\infty}f(x)=\int_{\mathbb{R}^d}f(y)\delta_x(dy)$.

($\Longrightarrow$) $f(y):=\min\{1,|y-x|\}$ とおくと $f\in C_b(\mathbb{R}^d)$ なので
$\min\{1,|x_n-x|\}=f(x_n)=\int_{\mathbb{R}^d}f(y)\delta_{x_n}(dy)$
$\xrightarrow{n\to\infty}\int_{\mathbb{R}^d}f(y)\delta_x(dy)=f(x)=0$.

(2) ($\Longleftarrow$) 明らか.

($\Longrightarrow$) $A=\{x\}$ とすれば $\lim_{n\to\infty}\delta_{x_n}(\{x\})=\delta_x(\{x\})=1$ なので, $\exists N\in\mathbb{N}, \forall n\geq N, |\delta_{x_n}(\{x\})-1|<1$. すると $\forall n\geq N$, $\delta_{x_n}(\{x\})>0$ すなわち $x_n=x$. //

さて, 以後の目標は次の定理の証明を与えることである:

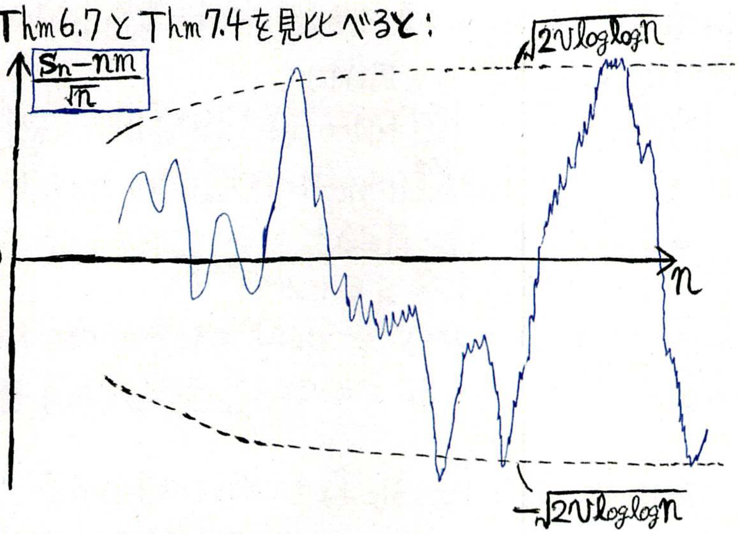
Thm7.4 (中心極限定理 (CLT, central limit theorem))
$\{X_n\}_{n=1}^\infty\subset\mathcal{L}^2(\mathbb{P})$ はi.i.d.とし, $m:=\mathbb{E}[X_1]$, $v:=var(X_1)$ とおく.

(1) $\mathcal{L}\left(\frac{S_n-nm}{\sqrt{n}}\right)\xrightarrow{\mathcal{L}}N(0,v)$. $\left(S_n:=\sum_{k=1}^n X_k\right)$

(2) $v>0$ とする. このとき $\forall x\in\mathbb{R}$,

$$\lim_{n\to\infty}\mathbb{P}\left[\frac{S_n-nm}{\sqrt{n}}\leq x\right]=\lim_{n\to\infty}\mathbb{P}\left[\frac{S_n-nm}{\sqrt{n}}<x\right]=\int_{-\infty}^x\frac{1}{\sqrt{2\pi v}}e^{-\frac{y^2}{2v}}dy.$$

Thm6.7とThm7.4を見比べると:



$v=var(X_1)>0$ ならば:

● a.s.でみると, $\frac{S_n-nm}{\sqrt{n}}$ は $\pm\sqrt{2v\log\log n}$ の間をいつまでも振動し続ける.

12月10日ここまで

● しかし振幅 $\sqrt{2v\log\log n}$ の増大は極めて遅く, 分布 $\mathcal{L}\left(\frac{S_n-nm}{\sqrt{n}}\right)$ は収束してくれる.

● 極限分布 $N(0,v)$ は $v=var(X_1)$ 以外の情報に依存しない.