確率論入門

渡辺澄夫
東京工業大学

１　確率空間

確率空間

$(\Omega, \mathcal{B}, P)$ 確率空間

Ω　集合

$\mathcal{B}$　部分集合の族(完全加法族)

P　$\mathcal{B}$ から $[0,1]$ への関数(確率分布)

○　$A \in \mathcal{B}$ に対してP(A)をAの確率という。

確率空間の例1

$(\Omega, \mathcal{B}, P)$ 確率空間

$\Omega = \{1, 2, 3, 4, 5, 6\}$

$\mathcal{B} = 2^{\Omega}$

$P(A) =$ [Aの要素の個数] / 6

○ $P(\{1,2\}) = 1/3.$

確率空間の例２

（Ω, $\mathscr{B}$, P）確率空間

Ω ＝ **R**

$\mathscr{B}$ ＝ **R** の開集合全体を含む
最小の完全加法族

$$P(A) = 1/(2\pi)^{1/2} \int_A \exp(-w^2/2)\, dw$$

○ P([0,1]) ＝ 0.3413・・・

なぜ完全加法族を決めておくのか

数学は基礎論になるほど難しいので最初はあまり気にしないでよい　　どうしても気になるならば

次の二つの公理が矛盾するから
選択公理と矛盾しないで確率を定義できる
集合族をきめておく必要があるからである。

○　選択公理　　集合の集合が無限個あるときでもすべての空でない集合から要素を選び出せる。

○　公理　実数全体の集合の全ての部分集合の確率を定義することができる。

２　確率変数

確率変数

$(\Omega, \mathscr{B}, P)$ を確率空間とする。
$(\Omega', \mathscr{B}')$ を可測空間とする。

可測関数　　$X: \Omega \to \Omega'$

を（Ω' に値をとる）確率変数という

○　関数のことを確率変数と呼ぶ。
　　関数を出力と同一視（混同）する（X=X(w)）。
　　関数がランダムなわけではない。

確率変数の気持ち

$$w \longmapsto X = X(w)$$

$(\Omega, \mathcal{B}, P)$

数学的に定義されるが
観測できないものとする
運(w)の決め方は
定めないでおく

Xの値は
実世界で
ランダムでない
とはいえない

なぜこんな定義をするのか

もともとランダムに値をとるということを数学的に
定義することができなくて困っていた

(Ω , $\mathscr{B}$, P) がわからず X だけ観測できる人には
X がランダムである場合も含む定義になっている

そこで関数 X(w) とその出力値 X を同一視して
確率変数(random variable)と呼ぶことにした。

これで「ランダムでないとはいえないもの」が定義されたがランダムとは何かについてはわからないままである。

確率変数の確率分布

確率変数　X: Ω　→　Ω’ が与えられたとき

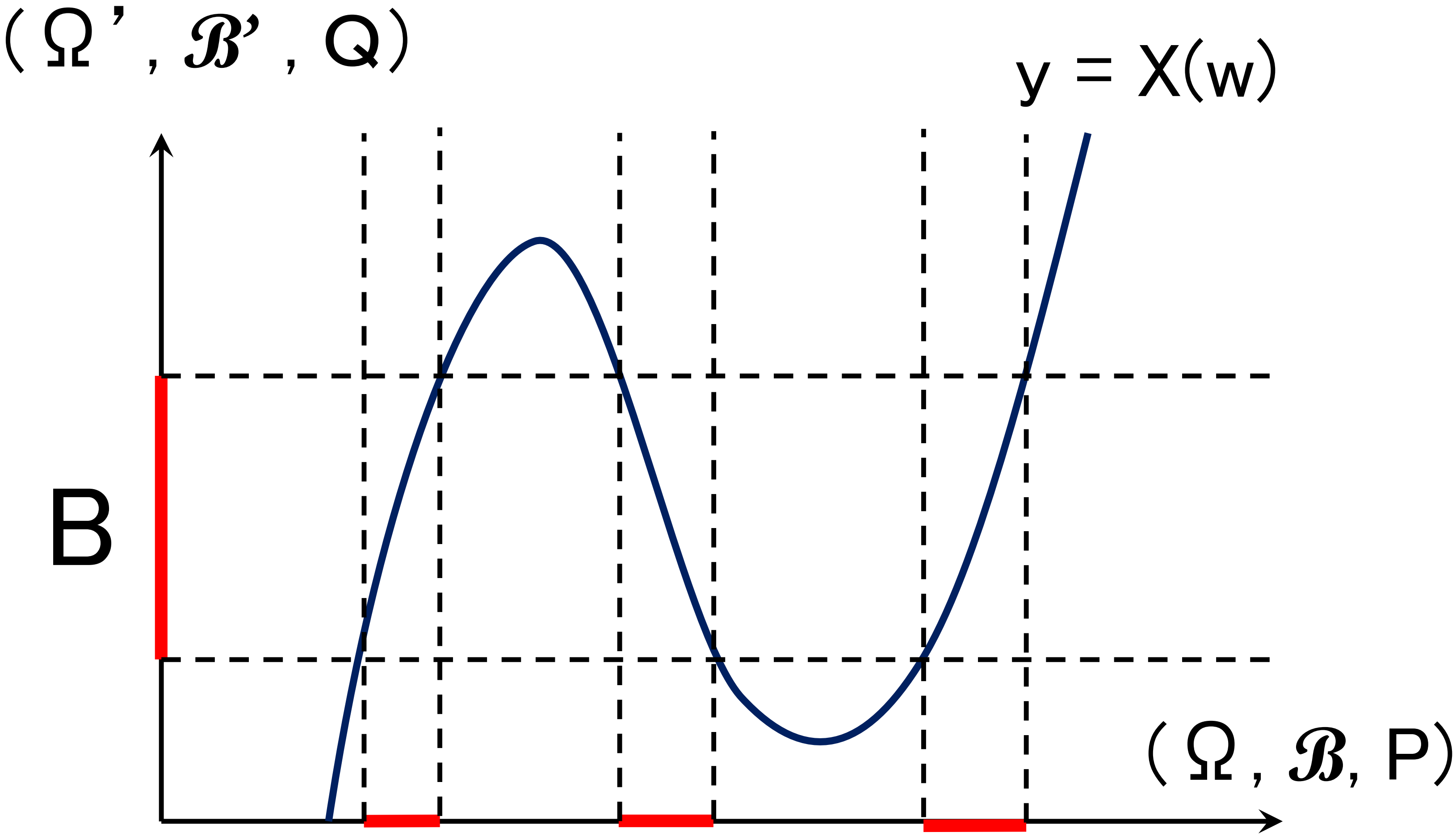
$B \in \mathcal{B}'$ に対して　$Q(B) = P(X^{-1}(B))$

と定めると $(\Omega', \mathcal{B}', Q)$ は確率空間になる。

このとき Q を X の確率分布という。

○　P は「Xの確率分布」ではない。

確率変数の確率分布



$X^{-1}(B)$

確率変数の確率分布の例1

$\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{B} = 2^{\Omega}$

P(A) = [Aの要素の個数] / 6

$\Omega' = \{0,1\}$, $\mathcal{B}' = 2^{\Omega'}$

確率変数　X(w) = w%2

⇒ 確率変数 X の確率分布は

$Q(\phi) = 0$, $Q(\{0\})=Q(\{1\})=1/2$, $Q(\{0,1\})=1$.

確率変数の確率分布の例2

$\Omega = \mathbf{R}$, $\mathcal{B} = \mathbf{R}$ の開集合全体を含む最小の完全加法族（ボレル集合族）

$P(A) = \int_A p(w)\, dw$

確率変数 $y = X(w)$

⇒ Xの確率分布は

$$Q(B) = P(X^{-1}(B)) = \int_{X(w)\in B} p(w)\, dw$$

$$= \int_B \left\{ \int \delta(y - X(w))\, p(w)\, dw \right\} dy$$

確率変数の確率分布の例2つづき

$(\mathbf{R}^N, \mathcal{B}, P)$ から実数への確率変数 $X=X(w)$
P が密度関数 $p(w)$ をもとときき X の密度関数 $q(x)$ は

$$q(x) = \int \delta(y-X(w))\, p(w)\, dw.$$

例. $w=(w_1,w_2,\cdots,w_N)$ が正規分布 $N(0,1^N)$ に従うとき $X(w)=w_1^2+w_2^2+\cdots+w_N^2$ は 自由度 N のカイ二乗分布に従う

$$q(x) = x^{N/2-1} \exp(-x/2) \;/\; \{ 2^{N/2}\, \Gamma(N/2) \}.$$

「確率変数⇒確率分布」は１対１ではない

$\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{B} = 2^{\Omega}$

P(A) = [Aの要素の個数] / 6

$\Omega' = \{0,1\}$, $\mathcal{B}' = 2^{\Omega'}$

確率変数　X(w) = w%2,　Y(w) = (int)((w−1)/3)

確率変数としては $X \neq Y$ しかし

⇒　確率変数 X, Y　の確率分布は　どちらも

$Q(\phi) = 0$, $Q(\{0\})=Q(\{1\})=1/2$, $Q(\{0,1\})=1$.

確率変数の確率分布

$(\Omega_1, \mathcal{B}_1, P_1)$

$w_1 \longmapsto X_1 = X_1(w_1)$

違う確率空間上に
定義された
確率変数が
同じ確率分布を
持つことはある

$(\Omega_2, \mathcal{B}_2, P_2)$

$w_2 \longmapsto X_2 = X_2(w_2)$

表記法

$(\Omega, \mathscr{B}, P)$ 確率空間、$(\Omega', \mathscr{B}')$ 可測空間

確率変数　　$X: \Omega \to \Omega'$

X の確率分布を Q とする。

確率論では略記法が用いられるので慣れる。

$$
\begin{aligned}
Q(B) &= Q(\{w' \in B\}) \\
&= P(\{w \in \Omega;\, X(w) \in B\}) \\
&= P(X \in B)
\end{aligned}
$$

さらに $B=\{w' \in \Omega';w'>1\}$ のとき $P(X>1)$ とも書く。

平均

(Ω, $\mathscr{B}$, P) 確率空間、(Ω’, $\mathscr{B}'$) 可測空間
確率変数　X: Ω → Ω’
X の確率分布を Q とする。

Ω＝Ω’=R で P, Q の密度関数を p(w), q(x) とすると

$$E[\,X\,] = \int X(w)\,p(w)\,dw$$
$$= \int x\,q(x)\,dx$$

３　確率変数の数列

確率変数の数列 $\{X_n\}$

応用上、確率変数の数列が与えられて，その収束を知りたくなることは多い。式で書くと簡単だが、その意味はきちんと理解する必要がある。

概収束　　$P(\lim X_n=X)=1$

p次（p>0）平均収束　$E(|X_n-X|^p)=0$

確率収束　$\forall \varepsilon, \delta>0, \exists N,$

$$n>N \Rightarrow P(|X_n-X|>\varepsilon)<\delta$$

確率変数列 $\{X_n\}$ の分布収束

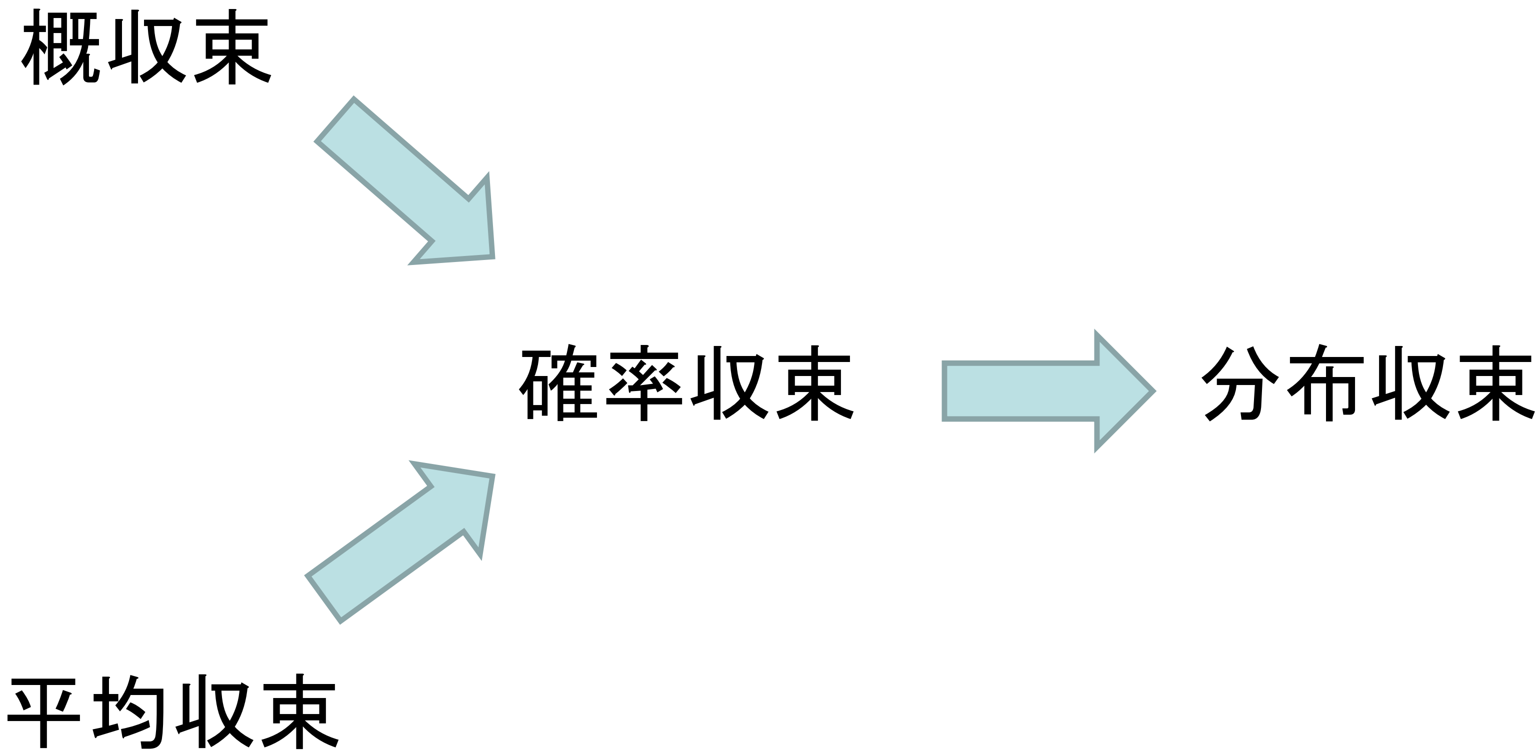
確率変数の分布収束（法則収束）は、確率変数が収束することではなく、確率変数の分布が収束することである。異なる確率空間上に定義された確率変数であっても、確率分布が収束すれば収束する。

分布収束：任意の有界連続関数 F に対して

$$\lim E[F(X_n)]=E[F(X)]$$

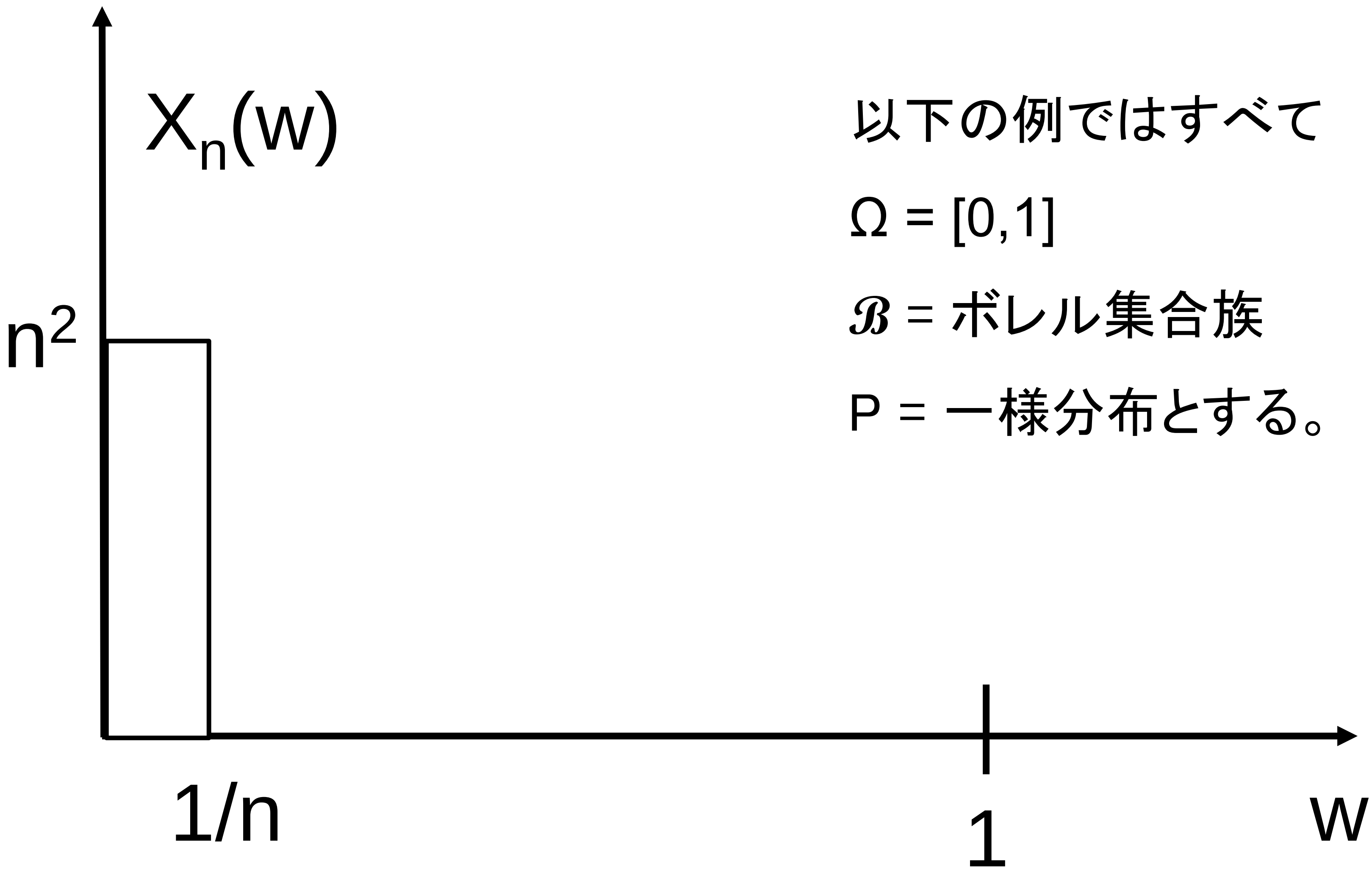
◎ 同値な命題はたくさんあり、それらのうちのいずれかを分布収束の定義とすることもある。

収束の関係



証明は本編で

概収束するが平均収束しない



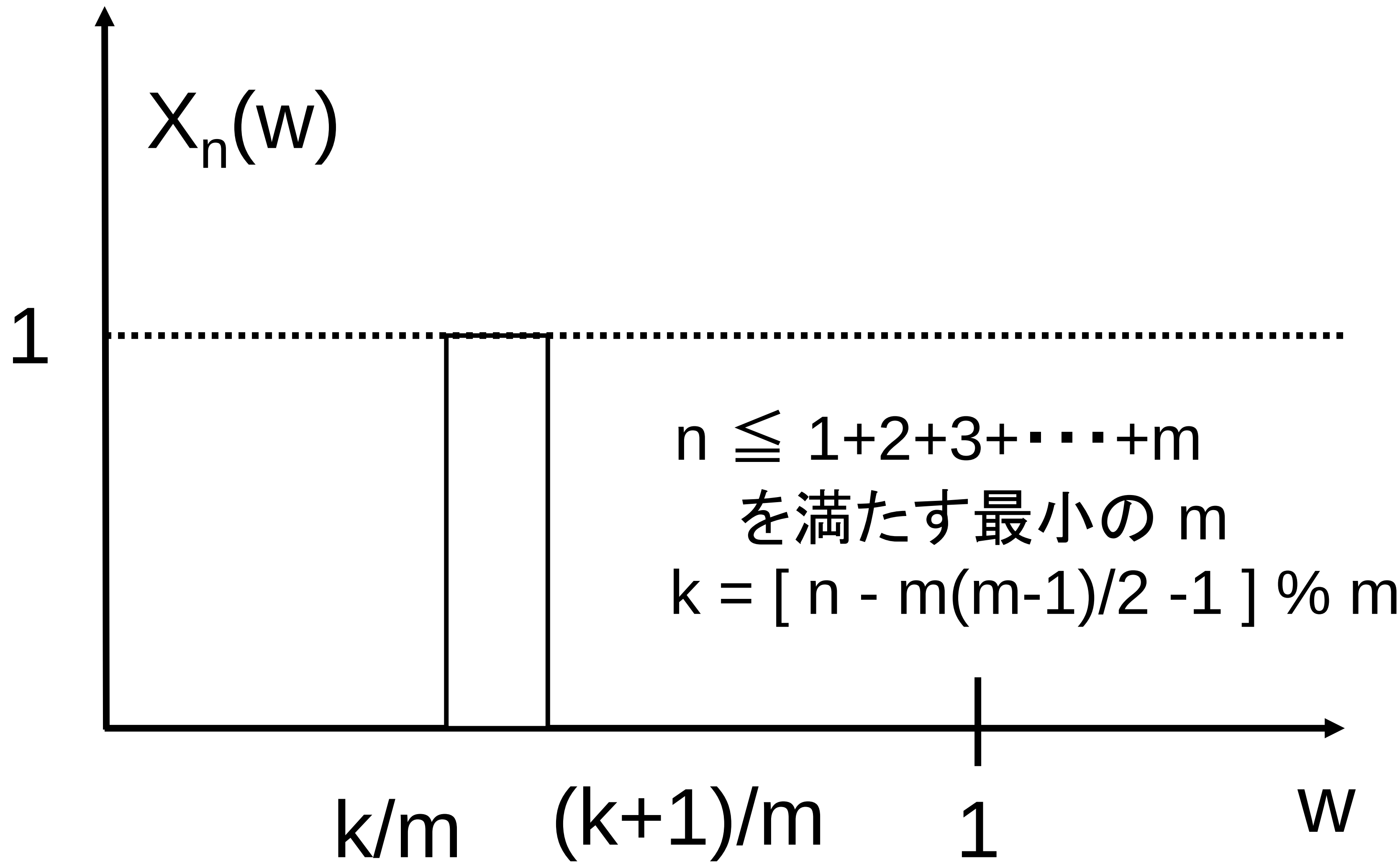
以下の例ではすべて

$\Omega = [0,1]$

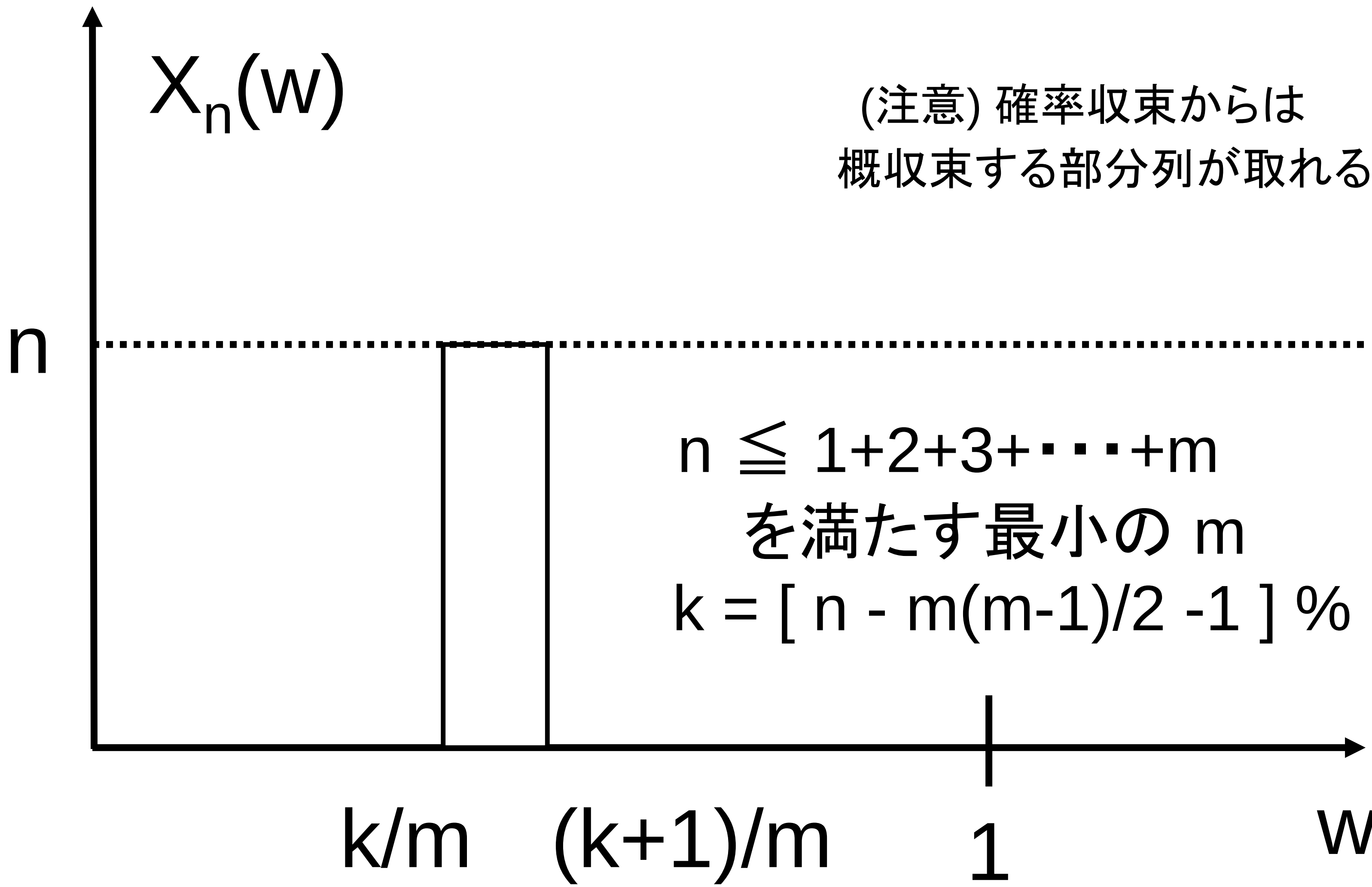
$\mathscr{B}$ = ボレル集合族

P = 一様分布とする。

平均収束するが概収束しない



確率収束するが概収束も平均収束もしない

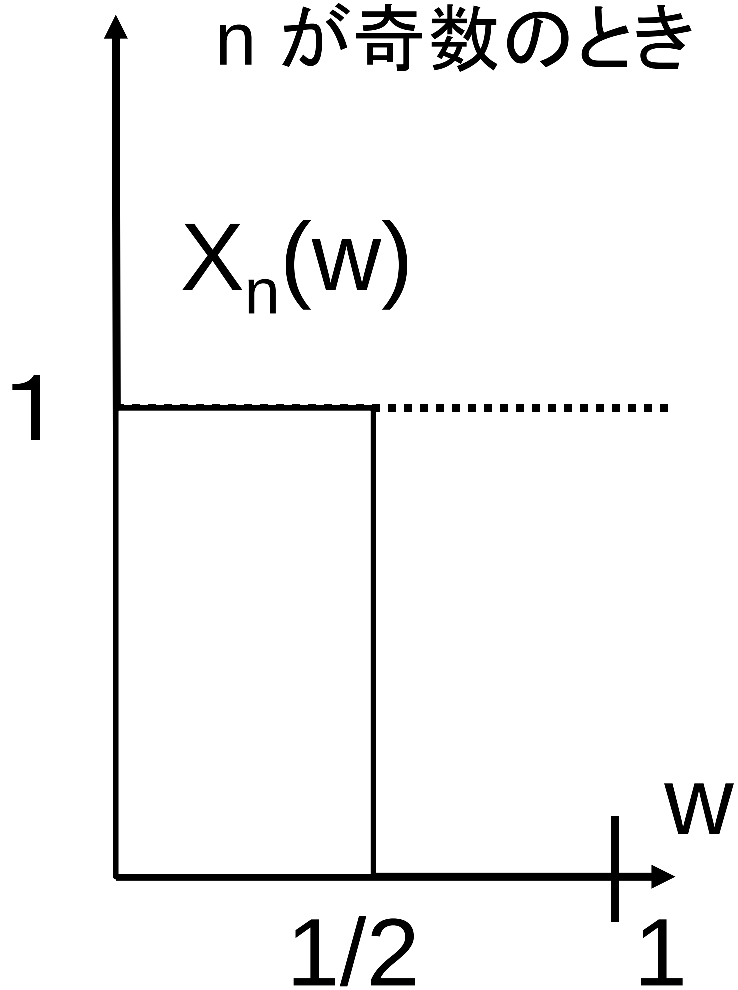


(注意) 確率収束からは
概収束する部分列が取れる

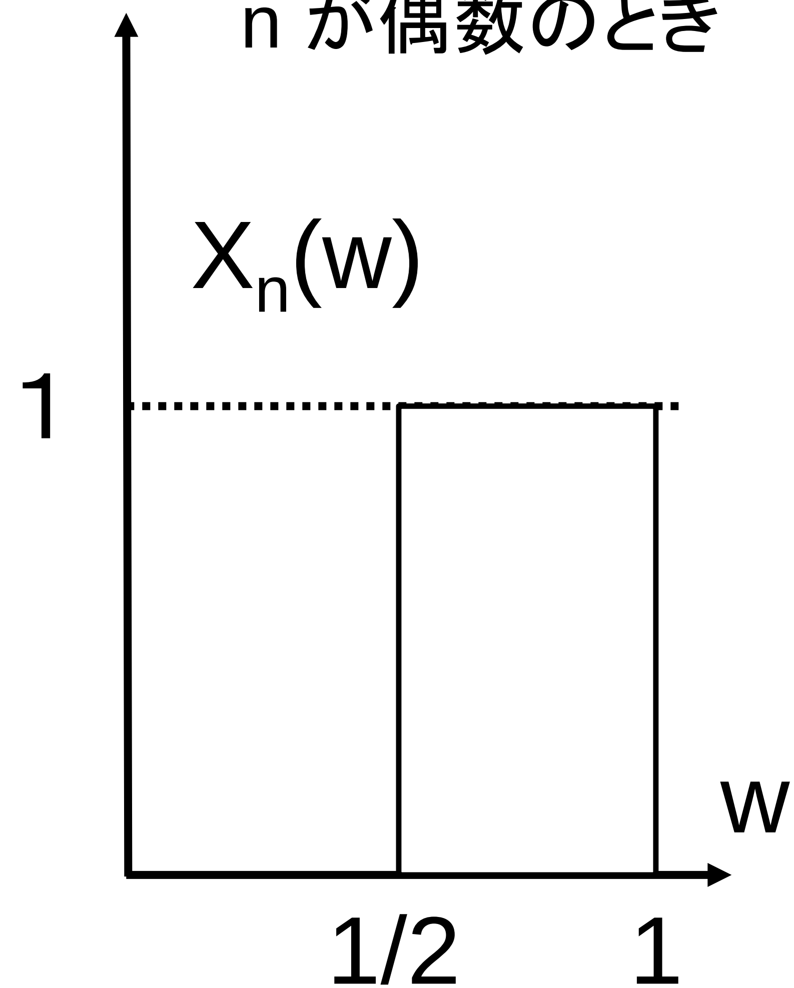
$n \leqq 1+2+3+\cdots+m$
を満たす最小の m
$k = [\ n - m(m-1)/2 - 1\ ]\ \%\ m$

分布収束するが確率収束しない

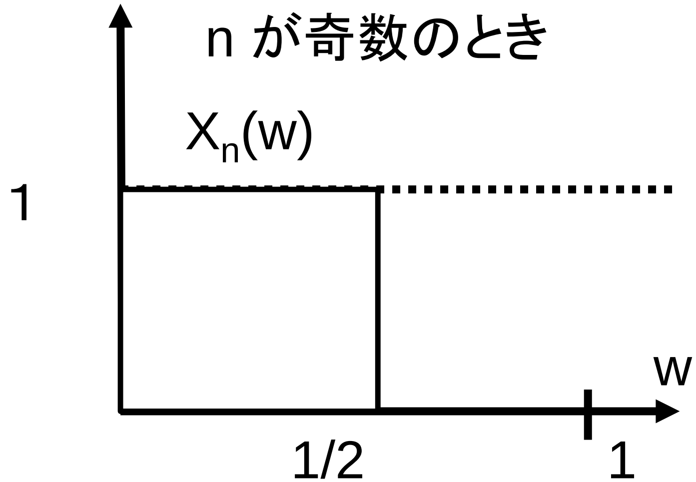
n が奇数のとき

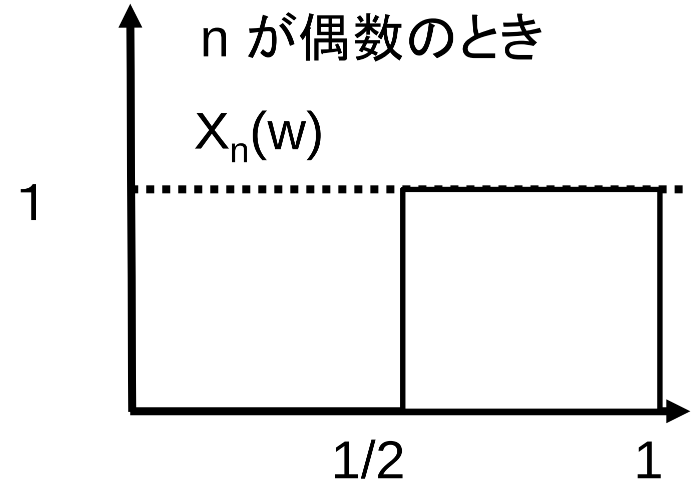


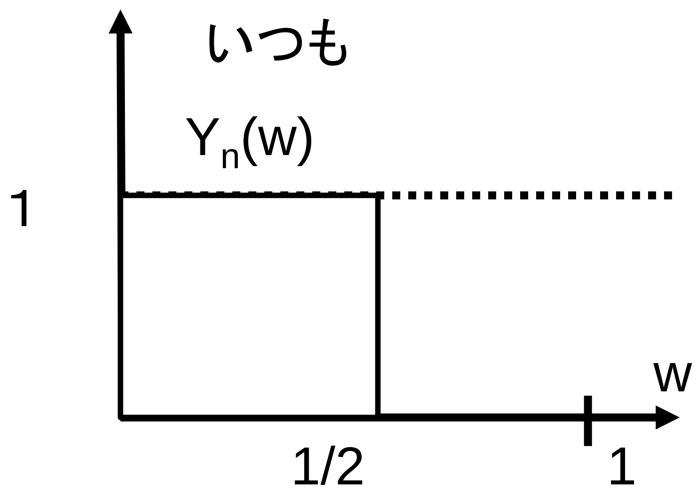
n が偶数のとき



「X_n分布収束＋Y_n概・平均収束」でも分布収束しない例







注意

○「定数に分布収束」は確率収束

○「定数に確率収束＋分布収束」は分布収束する

○ 完備可分な距離空間の一様にタイトな確率変数の列からは分布収束する部分列が取れる