経験過程とは何ですか

学生のみなさまから経験過程について質問があったので説明します。

(1)　サイコロを振ると確率的にサイコロの目がでます。このサイコロの目のことを確率変数と呼びます（整数に値をとる確率変数と呼びます）。

(2) サイコロの目を X とするとき Y=X−3.5 の平均は E[Y]=0 です。$\{Y_i\}$ を Y と同じ確率分布に従う独立な確率変数とすると $(Y_1+Y_2+\cdots+Y_n)/n^{1/2}$ の確率分布は $n\to\infty$で平均0分散V[X]の正規分布に収束します（中心極限定理）。

(3) 関数空間に値をとる確率変数を確率過程と呼びます。つまり
確率過程とは「振ると関数が出てくるサイコロ」のことです。

(4) 次ページ以後では<u>関数空間に値をとる確率変数の中心極限定理</u>を考えてみます。

(5) これは統計学や学習理論に広い応用があります。

経験過程の定義

$\{X_1, X_2, \cdots, X_n\}$：実数に値をとる確率変数で独立に
X と同じ確率分布に従うものとする。

$0 \leqq a \leqq 1$ とする。

関数 $f(a,x)$ は $[0,1] \times R$ から R への関数で
$E_X[f(a,X)]=0$ をみたすものとする。

定義。確率過程（つまり関数に値をとる確率変数）

$$\psi_n(a) = (1/n^{1/2}) \sum_{i=1}^{n} f(a, X_i)$$

のことを**経験過程**という。

経験過程の法則収束をめざしたい

定義からすぐに次式が成り立つことがわかります。

(1) $E[\psi_n(a)] = 0$,
(2) $E[\psi_n(a)\psi_n(b)] = E[f(a,X)f(b,X)]$.

そこで，正規確率過程 $\psi(a)$ を次の三つの性質

(1) $E[\psi(a)] = 0$,
(2) $E[\psi(a)\psi(b)] = E[f(a,X)f(b,X)]$,
(3) 各点 a ごとに正規分布.

を満たすものと定義します。この条件を満たす正規確率過程はユニークに存在することを証明することができます。そこで

希望：確率過程の法則収束「$\psi_n(a) \to \psi(a)$」を証明したいです。
何を示せば良いですか。

経験過程の法則収束を証明する

目標：確率過程の法則収束「$\psi_n(a) \to \psi(a)$」を証明する。

以下の説明は、初めて読むときは理解できる必要はありません。無理して理解しようとせず、次の次のページに早く進みましょう。

法則収束「$\psi_n(a) \to \psi(a)$」を証明するためには、$\psi_n(a)$ と $\psi(a)$ を入れている関数空間を定義し、関数空間の位相から定まるボレル集合族を定義して可測空間をつくり、$\psi_n(a)$ と $\psi(a)$ が確率変数であることを示して、法則収束することを示します。

統計学や学習理論への応用では、結果である $E[F(\psi_n)] \to E[F(\psi)]$ を適用したい汎関数 F があらかじめ与えられていることが多いのでその汎関数に対して成り立つようにあらかじめ関数空間の位相を定義しておきます（例えば sup ノルムにより位相が定義されるコンパクト集合上の有界連続関数全体が作るバナッハ空間など）。

経験過程の法則収束を証明する（続き）

（続き）前ページとこのページは無理をしてまで読む必要はありません。早く次のページに移動してください。

極限である正規確率過程の存在は「無限次元の確率過程が自然な条件のもとでユニークに存在する」というコルモゴロフの拡張定理の帰結としてすぐに得られます。難しいポイントは、関数空間での位相の意味で法則収束の成立を示すところになります。関数空間が完備可分な場合には、このことは次のようにして可能になります。まず$\{\psi_n(a)\}$の集合が関数空間の中で一様にタイトであることを示します（注）。次にプロホロフの補題「完備可分な距離空間の中の一様にタイトな集合からは法則収束する部分列を取り出せる」を適用すると収束先がひとつしかないことを用いて法則収束を得ることができます。完備可分でないときは話は難しくなってきます。

（注）確率変数 X を固定したとき、関数の集合 $\{f(x,a);a\in A\}$ が一様にタイトであるためには何らかの条件を満たす必要があります。

例　計算機で簡単に作れる

$X_1,X_2,\cdots,X_n$ 独立に[0,1]上の一様分布に従う。

$Y_1,Y_2,\cdots,Y_n$ 独立に平均0分散1の正規分布に従う。

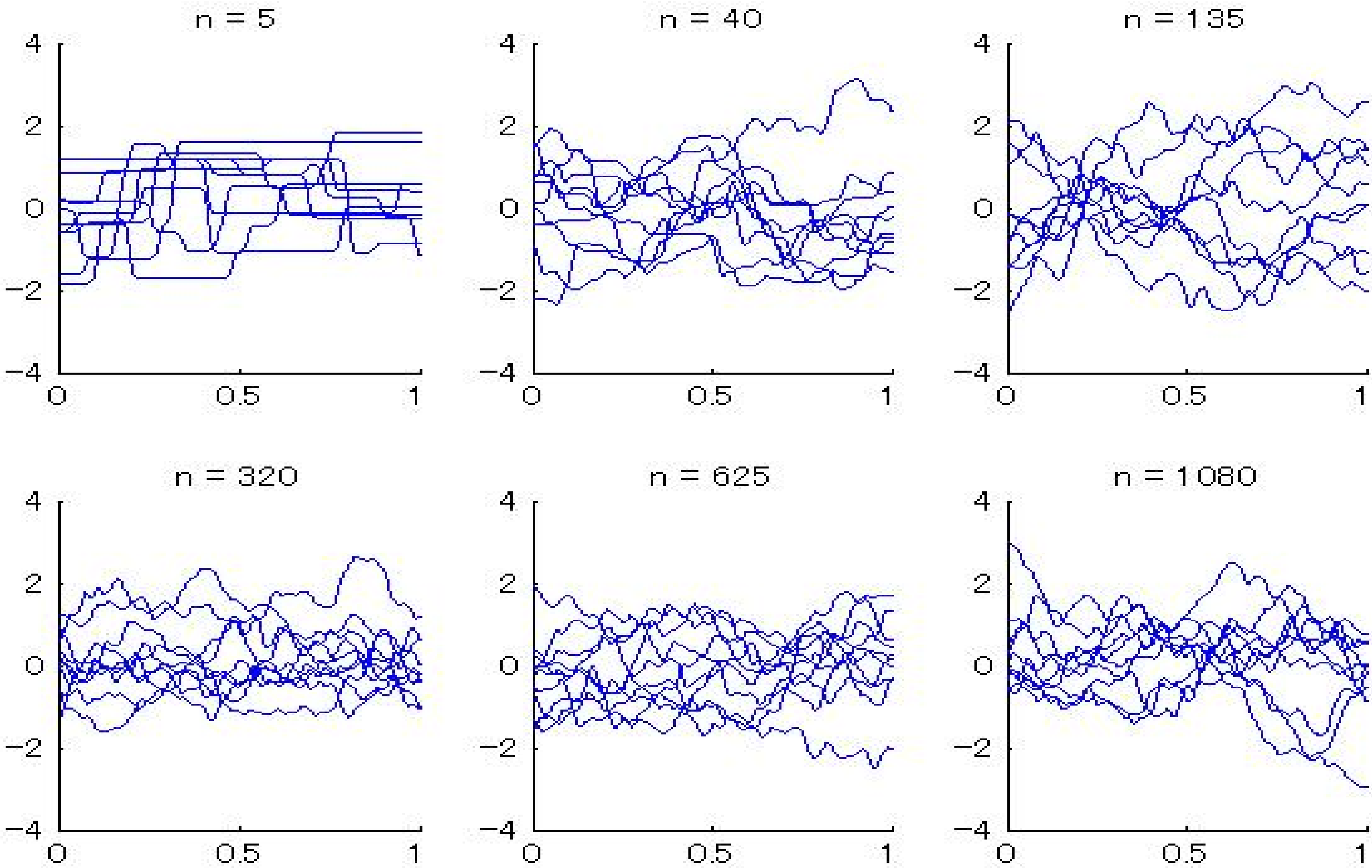
$0\leqq a\leqq 1$ の具体的な関数を

$$\psi_n(a) = (1/n^{1/2}) \sum_{i=1}^{n} Y_i \tanh(100(X_i - a))$$

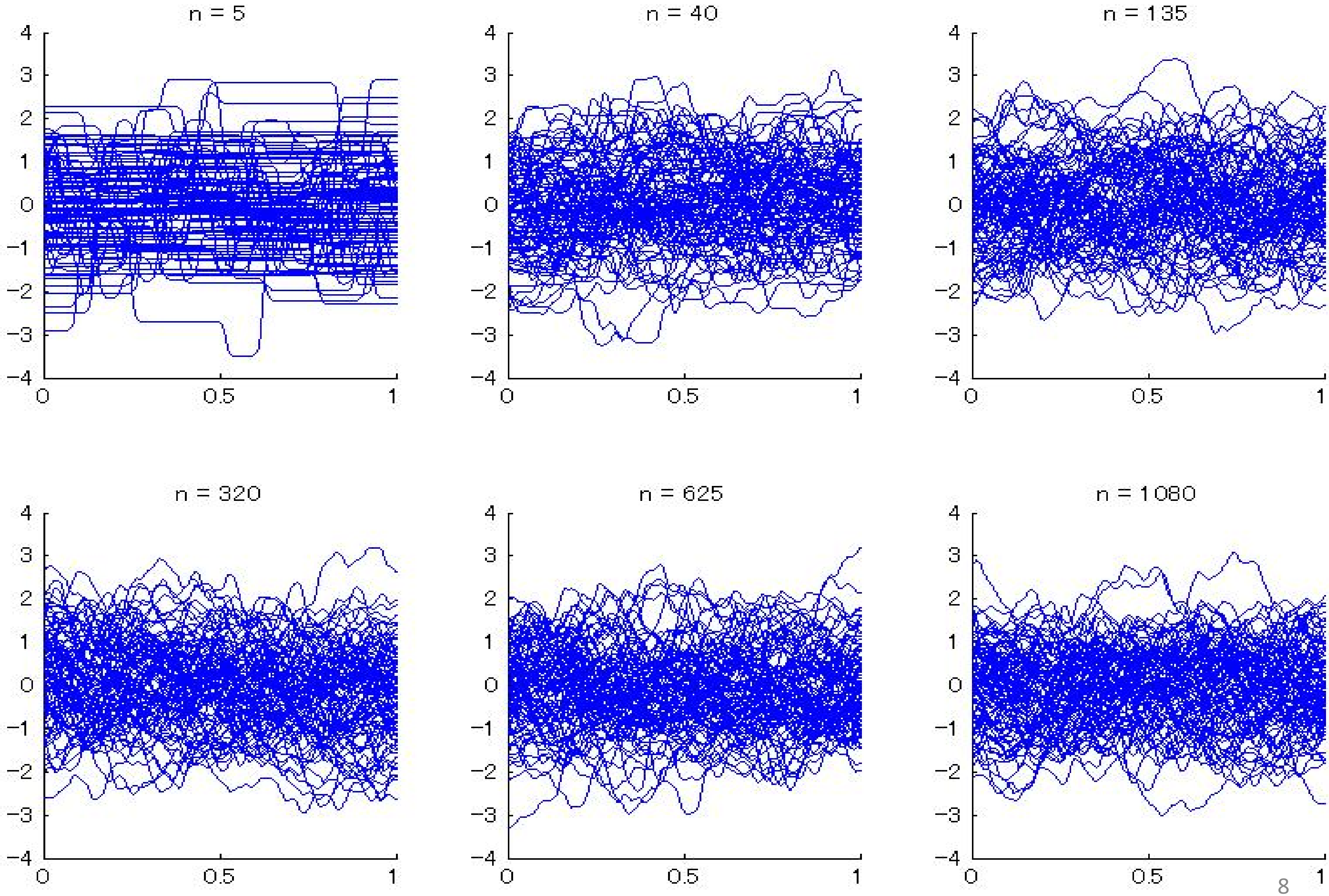
と定義する。
関数 $\psi_n(a)$ は a の関数ですが $X_1,X_2,\cdots,X_n$ と $Y_1,Y_2,\cdots,Y_n$ がでるたびに確率的に変動する確率過程です。

(注) 上記の関数 $\psi_n(a)$ は計算機で簡単に作れます。サイコロが具体的なものだと感じられるのと同じ程度に具体的なものだという感覚が得られるまでなんどでもこの確率過程を振ってみましょう。

例　関数を10個ずつ出してみた



例　関数を100個ずつ出してみた



関数空間上の中心極限定理とは

実世界で具体的な問題を扱うにはさらに次の課題が必要になります。

(1) もともとの関数空間が完備可分でない場合でも同じことができますか。
(2) 経験過程が関数空間上で可測になることは心配しなくてもよいですか。
(3) 正規確率過程 $\psi(a)$ が（確率１で）何回まで連続微分可能であるかを もとの関数 $f(a,x)$ から知るためにはどうしたらよいですか。
(4) 有界連続でない汎関数 F に対しても $E[F(\psi_n)] \to E[F(\psi)]$ を証明 したい場合には、さらにどのような条件が必要になりますか。
(5) これらのことは統計学や学習理論で何の役に立つのですか。

☆これらは非常に重要な課題ですが数学的に易しい問題だとは限りません。数学科で確率論を専門とする友人に相談してみると良いでしょう。上記の問題に関心がある人のために次の本があります。しかしながら準備なしで登頂できる峰ではないと思います。あらかじめそのことをご理解のうえ、図書館に行き本を手にとってみましょう。

Aad W. Van der Vaart and Jon A. Wellner, Weak Convergence and Empirical Processes. Springer, 1996.