経済学研究 *57-1*
北海道大学 2007.6

統計学入門における分布の収束のある取り扱いについて

園　　　信太郎

1．はじめに

統計学の入門講義を行う際に，通常の教科書を視て，以前より気になっていることなのだが，「分布の収束」が明白に議論されていないということなのである。しかし，中心極限定理の最も基本的な場合である独立同一分布変数列に対して，この定理の結論を述べる際に，分布の収束が，標準正規分布との関りで事実上言及されているのである。そこで，二項分布の正規近似などの，正規分布による「近似」の動機づけがなされるのである。だが，

$$\Phi(z)=\int_{-\infty}^{z}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{u^2}{2}\right)du,\quad z\in\mathbb{R},$$

と置いて，分布関数（右連続としておく）の列 $\{F_n\}_{n=1}^{\infty}$ に対して，

$$F_n(z)\xrightarrow[n\to\infty]{}\Phi(z),\quad \forall z\in\mathbb{R},$$

を主張するのみならば，収束の加減が各 z に依存しているはずなので，はたして，「充分大である n」にたいして，F_n が「全体的に」正規近似されているのか，この段階では不明である。しかし，実はこの収束は z に関して一様であり，

$$\sup_{z\in\mathbb{R}}|F_n(z)-\Phi(z)|\xrightarrow[n\to\infty]{}0,$$

が成立する。筆者は，論理的な要請のみでなく，教育上の配慮からも，この分布関数列の一様収束には授業中に言及すべきであると思っている。

2．George Pólya によるある命題

分布関数の列 $\{F_n\}_{n=1}^{\infty}$ が分布 F に収束するとは，

$$F_n(z)\xrightarrow[n\to\infty]{}F(z),\quad \forall z\in\{u\in\mathbb{R}\,|\,F \text{ は } u \text{ で連続}\},$$

となることであった。F が $\mathbb{R}$ 上で連続ならば $\forall z\in\mathbb{R}$ となる。ところが Pólya（ポーヤ）は，F が $\mathbb{R}$ 上で連続ならば，この収束は z に関して一様であることを指摘した。従って，特に中心極限定理

の場合には一様である。この Pólya の命題を示しておく。

$a \leq b$ を満たす任意の実数 a，b に対して，区間 $[a,\ b]$ 上の任意の単調非減少関数 F 及び G を考えると，

$$|G(x)-F(x)| \leq F(b)-F(a)+\max\{|G(a)-F(a)|,\ |G(b)-F(b)|\},\ \forall x \in [a,\ b],$$

が従う。このことを利用する。

ε を任意の正の数とする。F が分布関数であることより，$A<B$，$F(A)<\varepsilon$，$1-F(B)<\varepsilon$，を満たす実数 A 及び B が存在する。F は $\mathbb{R}$ 上で連続であるから，有界閉区間 $[A,\ B]$ 上で一様連続である。故に，

$$\exists \delta > 0 \forall a, b \in [A,\ B],\ |a-b|<\delta \Rightarrow |F(a)-F(b)|<\varepsilon,$$

となる。この δ を取ると，$(B-A)/N<\delta$ を満たす正の整数 N が存在するので，これを取って N とする。$[A,\ B]$ 上の点列 $x_i=A+i(B-A)/N$，$i=0,\ \cdots,\ N$，を導入する。すると上掲の不等式より，

$$\begin{aligned}\forall x \in [x_i,\ x_{i+1}],\ |F_n(x)-F(x)| &\leq F(x_{i+1})-F(x_i)+\max\{|F_n(x_i)-F(x_i)|,\ |F_n(x_{i+1})-F(x_{i+1})|\} \\ &< \varepsilon+\max\{|F_n(x_i)-F(x_i)|,\ |F_n(x_{i+1})-F(x_{i+1})|\},\end{aligned}$$

が各 $i=0,\ \cdots,\ N-1$，に対して従う。また，

$$\forall x \in (-\infty,\ A),\ |F_n(x)-F(x)|<\varepsilon+|F_n(x_0)-F(x_0)|,$$
$$\forall x \in (B,\ \infty),\ |F_n(x)-F(x)|<\varepsilon+|F_n(x_N)-F(x_N)|,$$

である。故に，

$$\forall x \in \mathbb{R},\ |F_n(x)-F(x)|<\varepsilon+\max\{|F_n(x_i)-F(x_i)| i=0,\ \cdots,\ N\},$$

となり，

$$\sup_{x \in \mathbb{R}}|F_n(x)-F(x)| \leq \varepsilon+\max\{|F_n(x_i)-F(x_i)| i=0,\ \cdots,\ N\},$$

が従う。ここで「$\lim a_n=a$，$\lim b_n=b \Rightarrow \max\{a_n,\ b_n\} \to \max\{a,\ b\}$」を想起すると，

$$\varlimsup_{n \to \infty} \sup_{x \in \mathbb{R}}|F_n(x)-F(x)| \leq \varepsilon,$$

となる。ε は任意の正の数であるから，$\sup_{x \in \mathbb{R}}|F_n(x)-F(x)| \xrightarrow[n \to \infty]{} 0$ となる。

3．おわりに

以上でPólyaによる命題が証明されたわけだが，これはC.R.Rao（1973）の第2章2c節の（vi），120頁，でPólya's theoremとして言及されている。しかし証明はなく，The proof is easy．として処理されている。またPólyaのどの論文によるのかは明示されていない。しかし筆者は，少なくとも入門段階での授業では，この命題は証明されるべきであると判断する。それは，ε-δ論法，関数列の一様収束，関数の一様連続性，不等式による評価，などの基本的な事柄の復習になるからである。ヒストグラムや数値実験を用いて，分布の列が極限分布へと収束する有様を提示することは，教育上の配慮として評価されるべきである。しかし，「科学」を教授する際には，やはり論理的な「すじ」は通すべきであろう。

2007年2月27日（火）

参考文献

Rao, Calyampudi Radhakrishna, *Linear Statistical Inference and Its Applications, Second Edition*,Wiley, New York, 1973.　第一版は1965年に出ている。この傑出した書物の索引は，内容との比較において，あまりに簡潔である。またRaoは，分布関数を左連続に取っている。なお次の第二版の邦訳がある。C.R.ラオ著；奥野忠一（おくのただかず），長田 洋（おさだひろし），篠崎信雄（しのざきのぶお），広崎昭太（ひろさきしょうた），古河陽子（ふるかわようこ），矢島敬二（やじまけいじ），鷲尾泰俊（わしおやすとし）　訳，『統計的推測とその応用』，東京図書，東京，1977年11月25日。Pólyaの定理は邦訳の112頁にある。