# 2013 年度秋学期　解析応用　第９回

## 第２部・集合と確率／ 確率変数，確率分布，期待値

確率変数 (random variable) とは，「さいころを 1 回ふったときに出る目の数字」のように，「値がランダムに決まる」数です。そこで，「確率変数がある値に（ある範囲の値に）なる確率」がいくらかを考えたのが確率分布です。今回は，確率変数と確率分布，それに確率変数のとる値の平均である期待値について，前回説明した枠組みにもとづいて説明してゆきます。

### 確率変数

ランダム現象の結果の集合である標本空間 $\Omega$ と，$\Omega$ の部分集合からなる $\sigma-$ 集合体 $\mathscr{F}$ を考えるとき，これらの組 $(\Omega, \mathscr{F})$ を可測空間とよび，さらに $\mathscr{F}$ の要素である各部分集合（事象）に割り当てられた確率 $P$ を考える時，これら 3 つの組 $(\Omega, \mathscr{F}, P)$ を確率空間とよぶ，と前回説明しました。

この枠組みで，上記の確率変数を定義することを考えてみましょう。「さいころを 1 回ふったときに出る目の数字」という表現を考えると，これは「さいころのある目が出る」というランダム現象の結果に，「目の数字」という数が対応していることを表しています。そこで，確率変数は「標本空間の要素（＝ランダム現象の結果）に対して，数値を対応づける関数」と考えることができます。

しかし，標本空間の要素に確率を割り当てられるとは限りません。そこで前回は，標本空間の部分集合を事象と考え，それに確率を割り当てました。それと同様に，確率変数もどんな関数でもよいわけではなく制限があり，次のように定義されます。

> 可測空間 $(\Omega, \mathscr{F})$ において，$\Omega$ 上の実数値関数 $X$ が
>
> $$\{\omega | X(\omega) \leqq a\} \in \mathscr{F} \tag{1}$$
>
> という関係を任意の $a$ について満たすとき，$X$ を**確率変数**という。

(1) 式は，確率変数の値が「$a$ 以下」となる事象がどんな $a$ についても存在することを示しています。さらに，この定義を拡張して，次の定理がなりたちます。

> 可測空間 $(\Omega, \mathscr{F})$ において，$\Omega$ 上の実数値関数 $X$ が確率変数であるための必要十分条件は，
> $B$ を実数 $R$ 上のボレル集合体 $\mathscr{B}$ に属する任意の集合とするとき
>
> $$\{\omega | X(\omega) \in B\} \in \mathscr{F} \tag{2}$$
>
> をみたすことである。

実数 $R$ 上のボレル集合体は「すべての半開区間 $(a, b]$ を含む最小の $\sigma-$ 集合体」ですから，上の定理は，確率変数の値の「範囲の集まり」$B$ にたいして，確率変数の値がその範囲にはいるという事象（$\mathscr{F}$ の要素である，$\Omega$ の部分集合）が存在することを示しています。確率空間 $(\omega, \mathscr{F}, P)$ においては，事象には確率が割り当てられていますから，「確率変数がある範囲になる（そんな事象がおきる）確率はいくら」ということが言えることになります。

（上の定理の証明）(1) 式がなりたつとき，(2) 式が成り立つことを示します。

(2) 式をみたす集合 $B_i$ $(i = 1, 2, \ldots)$ 全体からなる集合族 $\mathscr{A}$ を考えます。すなわち, $\mathscr{A} = \{B_i|\{\omega|X(\omega) \in B_i\} \in \mathscr{F}\}$ です。

まず，集合族 $\mathscr{A}$ が実数 $R$ 上の $\sigma-$ 集合体であることを示します。

1. $B_i$ を実数全体 $R$ と考えると，(1) 式が任意の $a$ についてなりたっているので，$\{\omega|X(\omega) \in R\} \in \mathscr{F}$ は成立し，$R \in \mathscr{A}$ です。
2. $A \in \mathscr{A}$ とします。$A^c$ を考えると，$\{\omega|X(\omega) \in A^c\} = \{\omega'|X(\omega') \in A\}^c$ です。$\{\omega'|X(\omega') \in A\} \in \mathscr{F}$ で, かつ $\mathscr{F}$ は $\sigma-$ 集合体ですから, $\{\omega|X(\omega) \in A^c\} = \{\omega'|X(\omega') \in A\}^c \in \mathscr{F}$ です。よって, $A^c \in \mathscr{A}$ です。
3. $B_i$ について,

$$\{\omega|X(\omega) \in \bigcup_{i=1}^{\infty} B_i\} = \bigcup_{i=1}^{\infty} \{\omega|X(\omega) \in B_i\} \tag{3}$$

がなりたち, さらに $\mathscr{F}$ は $\sigma-$ 集合体なので右辺は $\mathscr{F}$ に属します。よって, $B_i \in \mathscr{A}$ ならば $\bigcup_{i=1}^{\infty} B_i \in \mathscr{A}$ です。

以上から，集合族 $\mathscr{A}$ が実数 $R$ 上の $\sigma-$ 集合体であることが示されました。

一方，任意の半開区間 $(a, b]$ は，$(a, b] = (-\infty, b] - (-\infty, a]$ とすることができるので，(1) 式がなりたつならば，すべての半開区間は $\mathscr{A}$ に属します。実数 $R$ 上のボレル集合体 $\mathscr{B}$ は「すべての半開区間 $(a, b]$ を含む最小の $\sigma-$ 集合体」ですから，$\mathscr{A} \supseteq \mathscr{B}$ です。よって，$B_i \in \mathscr{B}$ なら $B_i \in \mathscr{A}$ で，$\mathscr{A}$ の定義から (2) 式が成り立っています。

逆に，(2) 式がなりたつとき，(1) 式は明らかになりたちます。よって定理は証明されました。■

**可測関数**

前節で確率変数について説明しましたが，確率変数をなんらかの関数でさらに変換した時，それをまた確率変数として扱えるのは，その「なんらかの関数」がどういう性質をもつときでしょうか？

そういう関数は，**可測関数**とよばれています。定義は次のとおりです。

> 実数の集合 $R$ とその上のボレル集合体 $\mathscr{B}$ からなる可測空間 $(R, \mathscr{B})$ において，実数から実数への関数 $f$ が
>
> $$\{x|f(x) \leqq a\} \in \mathscr{B} \tag{4}$$
>
> という関係を任意の $a$ についてみたすとき，$f$ を $(R, \mathscr{B})$ 上の可測関数という。

これは確率変数の定義とよく似ています。実際，可測空間 $(R, \mathscr{B})$ を可測空間 $(\Omega, \mathscr{F})$ に置き換えれば，確率変数の定義そのものです。

$f$ が連続関数ならば，$f$ は可測関数です。なぜならば，$f$ が連続なら，集合 $\{x|f(x) \leqq a\}$ は高々可算個の半開区間の結びで表現できるので，この集合は $\mathscr{B}$ に属するからです。これは，高々可算個の不連続点があっても同様です。

$f$ を $(R, \mathscr{B})$ 上の可測関数とするとき，$X$ が可測空間 $(\Omega, \mathscr{F})$ での確率変数ならば，$f(X)$ も確率変数です。なぜならば，$f$ は可測関数なので，任意の $a$ について $\{x | f(x) \leqq a\}$ は $\mathscr{B}$ に属します。よって，$\{\omega | f(X(\omega)) \leqq a\}$ は $\{\omega | X(\omega) \in \mathscr{B}\}$ と同じということになりますが，$X$ は確率変数ですから，前節の定理よりこれは $\mathscr{F}$ に属します。よって，$\{\omega | f(X(\omega)) \leqq a\}$ は $\mathscr{F}$ に属し，同じ定理により $f(X(\omega))$ も確率変数ということになります。

### 確率分布

ここまでの説明で，確率変数 $X$ の値の範囲に対して対応する事象 $\omega$ が定まることがわかりました。そこで，事象に確率を割り当てるかわりに，直接「確率変数の値の範囲に対して確率を割り当てる」ことにし，確率変数 $X$ の範囲 $X \leqq x$ に対して確率 $P(X \leqq x)$ を関数 $F(x)$ で割り当てます。

そうすると，結局これは，前回の最後に説明した，実数の区間に対して確率 $P((-\infty, x])$ を分布関数 $F(x)$ によって割り当てることと同じになります。そこで，以後この分布関数 $F(x)$ を「確率変数 $X$ の分布関数」と考えることにし，これを**確率変数 $X$ が分布関数 $F(x)$ で表される確率分布にしたがう**といいます。

---

### 期待値

期待値は，確率変数の平均を表す量です。ここでいう平均とは，「その確率変数のとりうるすべての可能性にわたっての平均」という意味です。例えば，「さいころを 1 回ふって出る目の数の期待値」は，さいころをふるという試行を十分多くの回数行った時の，出る目の数の（無限回の試行にわたった）平均です。

確率変数 $X$ の分布関数 $F(x)$ に対して密度関数 $f(x) = F'(x)$ が定義されているとき，$X$ の期待値 $E(X)$ を

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx \tag{5}$$

と定義する，ということはどこかで習ったことがあると思います。しかし，これは密度関数が定義されているときにしか使えません。実際，離散型確率空間における確率変数（離散型確率変数）については，分布関数は階段状で微分できませんから，上の定義は使えず，確率関数 $P(X = x)$ を使って

$$E(X) = \sum_{x} x P(X = x) \tag{6}$$

という定義を習ったと思います。

このような煩雑さを避けて，期待値を一般的に定義するには，**ルベーグ積分**を用います。(5) 式のような通常の積分（リーマン積分）では，定義域（横軸）を均等に細分して，その各部分に対応する（部分の幅 × 関数の値）を合計します。このため，定義域の性質に制限があります。区分的連続な関数はリーマン積分可能ですが，それ以外の関数にはリーマン積分ができないものがあります。

これに対してルベーグ積分では，図 1 のように，関数の値を表す縦軸のほうを細分し，関数の値がある範囲に入るような定義域の区間（均等とは限らない）を取り出して，（関数の値 × 区間の幅）を合計します。この考えによる期待値の定義を，離散型確率分布に対する期待値（(6) 式）に似た考えから出発

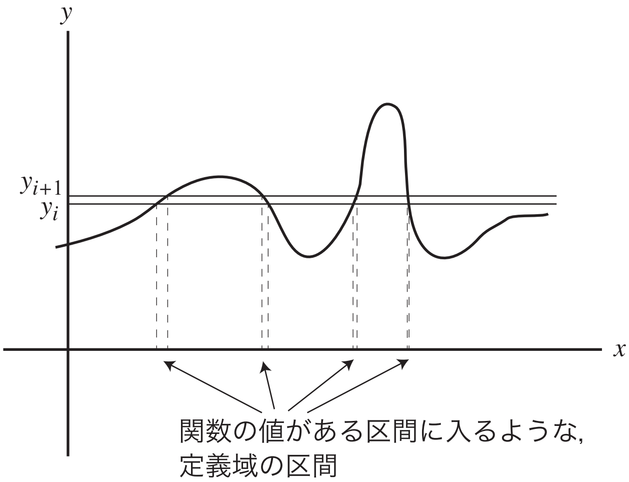


図 1: ルベーグ積分の考え方

して導いてみましょう[1]。

(6) 式は，関数 $\varphi(x)=x$ と確率 $P(X=x)$ の積の総和になっています。これをもっと一般的に考えて，実数の集合 $R$ を，排反な集合 $B_1, B_2, \ldots$ に分割します。これらはボレル集合体の要素と考えられるので，$R$ 上の確率変数 $X$ について「$X$ が $B_i$ に属するという事象」すなわち $\{\omega|X(\omega)\in B_i\}$ には，確率が割り当てられています。これを $P(\{\omega|X(\omega)\in B_i\})$ とします。

一方，関数 $\varphi(x)$ は，集合 $B_i$ に対しては値が一定値 $a_i$ であるとします。このような関数は，集合 $B$ の**定義関数** $I_B$ を

$$I_B(x)=\begin{cases}1 & x\in B\\ 0 & x\notin B\end{cases} \tag{7}$$

として

$$\varphi(x)=\sum_i a_i I_{B_i}(x) \tag{8}$$

で表されます。このような関数 $\varphi(x)$ を**単関数**といいます。いわば，「階段状の関数」[2]です。

このように考えると，「関数と確率の積」の合計は $\sum_i \varphi(x(\omega))P(\{\omega|X(\omega)\in B_i\})$ すなわち

$$\sum_i a_i P(\{\omega|X(\omega)\in B_i\}) \tag{9}$$

と表すことができます。$B_i$ の区切りがどれだけ細かくなり，単関数 $\varphi(x)$ の階段の幅がどれだけ狭くなっても，それぞれに対して確率は割り当てられています。そして，確率測度には完全加法性がありますから，ある事象を可算無限個に分割しても，分割された各事象に対する確率の合計は，それらの和事象である元の事象の確率になっています。

以上から，$B_i$ の区切りが十分に細かくなったときの，(9) 式の極限を

$$\int_R \varphi(x(\omega))dP(\omega) \tag{10}$$

[1]本当はもう少し細かい議論がありますが，概略を示します。詳しくは前回あげた参考文献を見てください。
[2]ただし，段の幅は一定とは限りません。

と書き，関数 $\varphi$ の $P$ に関するルベーグ積分といいます。単関数は可測関数ですから，$\varphi$ が可測関数のときにこのように積分を定義できます。

さて，分割 $B_i$ を $B_i = (x_{i-1}, x_i]$ という区間とするとき，分布関数 $F$ を用いると

$$\begin{aligned}\sum_i a_i P(\{\omega|X(\omega) \in B_i\}) &= \sum_i a_i P(\{\omega|x_{i-1} < X(\omega) \leqq x_i\}) \\ &= \sum_i a_i [F(x_i) - F(x_{i-1})]\end{aligned} \tag{11}$$

とあらわせるので，(10) 式の積分は

$$\int_R \varphi(x) dF(x) \tag{12}$$

で，すなわち関数 $\varphi$ の $F$ に関するルベーグ積分ということもできます。

分布関数 $F(x)$ で表される確率分布にしたがう確率変数 $X$ の期待値 $E(X)$ は，この積分で $\varphi(x) = x$ とおいた場合，すなわち

$$E(X) = \int_R x dF(x) \tag{13}$$

となります。直感的には，密度関数 $f(x)$ が存在する時は $dF(x) = f(x)dx$ と考えれば，これが (5) 式の期待値の定義と同じになることがわかります[3]。また，(10) 式を $E(\varphi(X))$ で表し，$\varphi(X)$ の期待値といいます。

---

**問題**

$\Omega$ 上の実数値関数 $X$ が可測空間 $(\Omega, \mathscr{F})$ での確率変数であることと，任意の $a$ について $\{\omega|X(\omega) < a\} \in \mathscr{F}$ が同値であることを示してください。

(ヒント：

$$\{\omega|X(\omega) < a\} = \bigcup_{n=1}^{\infty} \left\{\omega|X(\omega) \leqq a - \frac{1}{n}\right\} \tag{A1}$$

であることを用います)

[3] あくまで，直感的には，です。