第3章　確率論の基礎

3.1　はじめに

前の章では，貸した資金が回収不能に陥る可能性を全く考慮せずに,「金利」を扱った．そこでは，資金が時間を通じてどのように運用されるかが最大の関心事となった．(金利の期間構造！)

しかし，他人に金を貸すという行為を考えれば，常に貸し倒れの危険を避けることはできない．実際，日本の金融機関の多くは，不良債権という名の貸し倒れの危険が異常に大きい金融資産をそのポートフォリオに抱え込む状態に陥っている．早い話，借用証書がタダの紙切れになろうとしているわけである．

危険 (risk)，リスクは，経済的損失の可能性を広範に指す．ファイナンスにおいては，将来資産価格を確実に知ることができないことから生ずる危険が，主たる研究対象となる．そこで，将来資産価格を確率変数として扱うことで数学的に危険を扱うことが，必要となる．ここでは，ファイナンスにおける危険を扱うための，確率論の基礎を学ぶ.

3.2　(有限) 離散型確率変数

ファイナンス数学Iで登場する確率変数の多くは離散型確率変数である．具体的には，とり得る値とその値が実現する確率を対としたものを (有限個) リストアップした

ものが，(有限) 離散型確率変数である．

例 2. サイコロの目 X は典型的な離散型確率変数である．X のとり得る値の可能性は *1,2,3,4,5,6* の *6* 個であり，それぞれの目が実現する確率は，歪みがないサイコロであれば

$$Pr(X = x) = \frac{1}{6}, \quad (x = 1, 2, 3, 4, 5, 6)$$

つまり

とり得る値	確率
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$

である．

一般に，高々可算な集合 $\mathcal{X}$ から区間 $[0, 1]$ への関数

$$p : \ \mathcal{X} \longrightarrow [0, 1]$$

のうち

$$\sum_{x \in \mathcal{X}} p(x) = 1$$

となるものを，確率関数とよぶ．集合 $\mathcal{X}$ 上で定義される確率関数を確率分布ともいう．確率分布と対で考える $\mathcal{X}$

を離散確率変数とよび，大文字 X で表す[1].

以下，とりうる値の集合が有限だと考えて，話を進めよう.（つまり集合 $\mathcal{X}$ は有限集合.）

3.3 期待値と分散

確率変数 X の期待値 $E[X]$ とは，

$$E[X] = \sum_{x\in\mathcal{X}} p(x) \times x$$

で定義される数値である.

演習 6. 歪みのないサイコロの目の期待値を求めよ.

離散確率変数に対して p を固定しつつ

$$g:\ \mathcal{X} \longrightarrow \mathbf{R}$$

を考えることによって，新たな確率変数 Y が $Y = g(X)$ のように作られたと考えることができる．すると，X から派生して作られた確率変数 Y に対しても

$$E[Y] = \sum_{x\in\mathcal{X}} p(x) \times g(x)$$

と定義することで期待値が定義できる.

[1] 正確には事象の集合に対して各事象に対応する値を対応させる写像を確率変数とよび，事象に集合から区間 $[0,1]$ の総和関係を満たす関数を確率分布とよぶ．しかし，離散確率分布の場合，事象の集合ととり得る値の集合を同一視しても，問題がないために，上のように確率変数を定義した.
この場合，確率変数は

$$X : \mathcal{X} \longrightarrow \mathbf{R}$$

を恒等写像に対応させているとかんがえてもよい.

例 3.

$$g(x) = 2x$$

$$g(x) = x + 3$$

$$g(x) = x^2$$

$$g(x) = \log x$$

以下 $E[X^2]$ のような記法が頻繁に登場する．

さて確率変数 X の期待値を $\mu = E[X]$ とおいて

$$E[(X-\mu)^2] = \sum_{x \in \mathcal{X}} p(x) \times (x-\mu)^2$$

で定義される数値を $Var(X)$ と記し，分散とよぶ．さらに分散の平方根 $\sigma = \sqrt{Var(X)}$ を標準偏差 (standard deviation) とよぶ．

注意 7. 任意の確率変数 X について，$Var(X)$ は非負の値である．(何故か？)

注意 8. X の期待値 μ と標準偏差 σ が与えられているとき

$$Z = \frac{X-\mu}{\sigma}$$

で定義される，新たな確率変数は

$$E[Z] = 0, \quad Var[Z] = 1$$

を満たす．(何故か？)

3.4　連続型確率変数

これまで確率変数のとりうる値は高々可算であると考えてきた.（実際には有限個と考えて，解説してきた.）

とりうる値の集合 $\mathcal{X}$ が，実数全体あるいは区間 $(0,\infty)$ のような非可算集合である場合，確率変数は連続型確率変数とよばれる.

離散型では，各値一つ一つに確率を付与する確率関数を考え，それを確率分布と考えたが，連続型ではそうはいかない．非可算無限個ある対象を「総和する」ことはできないからである.

そこで，X がある値 x 以下になる確率を x の関数として

$$F(x) = Pr(X \leq x) \tag{3.1}$$

と定義し分布関数 (distribution function) とよび，この関数を確定することが確率分布を決定することだと考える. $F(x)$ は単調非減少関数であり

$$\lim_{x\to-\infty} F(x) = 0, \quad \lim_{x\to\infty} F(x) = 1$$

という性質をもつ.

演習 7. 上のことを確認せよ.

さらに定義から，$a \leq b$ に対して

$$F(b) - F(a) = Pr(a < X \leq b)$$

である．また分布関数が導関数をもつ場合

$$f(x) = \frac{dF(x)}{dx} \tag{3.2}$$

を確率密度関数(density function)という．明らかに

$$F(x) = \int_{-\infty}^{x} f(u)du$$

である.（何故か？）

演習 8. $a \leq b$に対して

$$\int_{a}^{b} f(u)du = Pr(a < X \leq b)$$

を示せ.

連続型確率変数は，密度関数$f(x)$を用いて

$$E[X] = \int_{-\infty}^{\infty} uf(u)du \tag{3.3}$$

として期待値を定義することができる．また，$Y = g(X)$のようにXから派生して作られる確率変数の期待値も

$$E[X] = \int_{-\infty}^{\infty} g(u)f(u)du \tag{3.4}$$

で定義される.

分散も$\mu = E[X]$とおくとき

$$Var[X] = E[(X-\mu)^2] = \int_{-\infty}^{\infty} (u-\mu)^2 f(u)du \tag{3.5}$$

離散型同様，$\sigma = \sqrt{Var[X]}$を標準偏差とよぶ.

注意 9. 分布関数が連続微分可能でなく密度関数が存在しない場合でも，本来は連続型確率変数の期待値，分散を積分として定義することが可能である．その場合の積分は，諸君が習ったリーマン積分ではない．

注意 10. 確率密度関数が存在する分布のみを考察の対象とする利点は，確率論の大部分の議論を初等的な微分積分の範囲に収めることができるところにある．

おおまかに言って，離散型で展開した議論は，

$$\text{確率関数} \longleftrightarrow \text{密度関数}$$

$$\text{事象に関する総和} \sum_{x \in \mathcal{X}} \longleftrightarrow \text{実数直線上の積分} \int_{-\infty}^{\infty} \ dx$$

という対応によって，ほぼ平行して成り立つといってよい．*(*もちろん例外はある．*)*

3.5　正規分布

連続型分布の代表的なものとして正規分布が上げられる．正規分布はカイ自乗分布や t 分布のような，多くの分布の基礎となることからも重要性はずば抜けている．

正規分布とは密度関数が

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{3.6}$$

で表される分布である．正規分布にしたがう確率変数の期待値は $E[X]=\mu$，分散は $Var[X]=\sigma^2$ であり，

$$X \sim N(\mu, \sigma^2)$$

のように記す．さらに $X \sim N(0,1)$ であるとき，標準正規分布という．

注意 11. *1.* $X \sim N(\mu,\sigma^2) \Longrightarrow aX+b \sim N(a\mu+b, a^2\sigma^2)$

2. $X \sim N(\mu,\sigma^2) \Longrightarrow \frac{X-\mu}{\sigma} \sim N(0,1)$

注意 12. ファイナンスにおいては，金融商品の連続複利収益率が正規分布にしたがうと想定されることが多い．

3.6 同時分布: 離散型

二つの金融商品の価格あるいは収益率を確率変数と考え，両者の確率的な特性を調べるには同時分布を考える必要が生ずる．

離散型では二つの確率変数 X と Y のとりうる値の集合 $\mathcal{X}$ と $\mathcal{Y}$ の上で定義され，非負の値をとり，

$$\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} p(x,y) = 1$$

となる関数 $p(x,y)$ を確率関数と考える．

また

$$p_X(x) = \sum_{y\in\mathcal{Y}} p(x,y), \quad p_Y(y) = \sum_{x\in\mathcal{X}} p(x,y) \tag{3.7}$$

をそれぞれ周辺分布とよぶ．周辺分布は，それぞれの確率変数がもう一方の確率変数の値に関係なくとる確率を考えたものである．

さらに

$$p(x,y) = p_X(x)p_Y(y) \tag{3.8}$$

であるとき，二つの確率変数 X と Y は独立であるといわれる．

演習 9. 白と黒の *2*個のサイコロを用意する．白のサイコロの目を確率変数 X で，黒のサイコロの目を確率変数 Y で表すことにするとき，同時分布関数を求めよ．また，各周辺分布を求め，独立か否かを判定せよ．

注意 13. 独立であるとき，二つの周辺分布の確率関数から同時分布が決まる．しかし，一般に周辺分布の情報のみから同時分布関数を構成することはできない．

二つの確率変数に依存する確率変数 $g(X,Y)$ の値の期待値は

$$E[g(X,Y)] = \sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} g(x,y)p(x,y) \tag{3.9}$$

で定義される．

演習 10. 演習 *9*において，確率変数 X，Y のそれぞれの期待値・分散を求めよ．また，$X+Y$ の期待値・分散を求めよ．

さて同時分布が与えられているとき，$g(x,y)=x, g(x,y)=y$ とおくことで期待値 $\mu_X=E[X]$ と $\mu_Y=E[Y]$ が得られる．さらに $g(x,y)=(x-\mu_x)^2$，$g(x,y)=(y-\mu_y)^2$ とおくことで，X の分散 σ_X^2 と Y の分散 σ_Y^2 が定義される．

ここで t を実数として $g(X,Y)=tX+Y$ とおき，$g(X,Y)$ の分散を計算する．

$$
\begin{aligned}
Var[tX+Y] &= E[(tX+Y-(t\mu_X+\mu_y))^2] \\
&= t^2E[(X-\mu_X)^2]+2E[(X-\mu_X)(Y-\mu_Y)]t \\
&\quad +E[(Y-\mu_Y)^2] \\
&= t^2\sigma_X^2+2E[(X-\mu_X)(Y-\mu_Y)]t+\sigma_Y^2
\end{aligned}
$$

となる．

分散は，負になりえないから，上で計算した左辺は t の値にかかわらず非負でなくてはならない．そのためには，上の式の右辺を t に関する 2 次式とみなしたとき，判別式が正の値をとらないことが必要である．つまり

$$E[(X-\mu_X)(Y-\mu_Y)]^2-\sigma_X^2\sigma_Y^2\leq 0$$

が必要である．ここで

$$\sigma_{XY}=E[(X-\mu_X)(Y-\mu_Y)]$$

とおき，共分散とよぶことにする．また

$$\sigma_{XY}^2\leq\sigma_X^2\sigma_Y^2$$

とかかれるから，$\sigma_X^2 > 0, \sigma_Y^2 > 0$ のとき

$$\rho_{XY}^2 = \frac{\sigma_{XY}^2}{\sigma_X^2 \sigma_Y^2} \leq 1$$

である．さらに σ_{XY} が負の値もとりうることを考慮して

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

と定義し，相関係数とよぶ.

注意 14. $\rho_{XY} = 1$ のとき X と Y の間に正の完全相関があるといい，$\rho_{XY} = -1$ のとき X と Y の間に負の完全相関があるという.

これらのとき，$\sigma_{XY} = \sigma_X \sigma_Y$ あるいは $\sigma_{XY} = -\sigma_X \sigma_Y$ となる.

また $\rho_{XY} = 0$ のとき無相関である.

注意 15. 二つの確率変数が独立であるとき

$$\begin{aligned}
\sigma_{XY} &= E[(X - \mu_X)(Y - \mu_Y)] \\
&= \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} (X - \mu_X)(Y - \mu_Y) p(x, y) \\
&= \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} (X - \mu_X)(Y - \mu_Y) p(x) p(y) \\
&= \sum_{x \in \mathcal{X}} (X - \mu_X) p(x) \sum_{y \in \mathcal{Y}} (Y - \mu_Y) p(y) \\
&= 0
\end{aligned}$$

となる.

演習 11. 二つの確率変数が，独立なら無相関であることはわかった．実は無相関でも独立とはかぎらない．具体的な例を示せ．

3.7 同時分布: 連続型

1変数の場合と同様，とりうる値が連続的であるような2つの確率変数についても同時分布を考えることができる．

$$F(x,y) = Pr(X \le x, Y \le y)$$

で2変数の同時分布関数という．

$$F(-\infty,-\infty) = 0, \quad F(\infty,\infty) = 1$$

である．

ここでは，簡単に密度関数

$$f(x,y) = \frac{\partial^2}{\partial x \partial y} F(x,y)$$

が存在する場合だけを考える．密度関数と分布関数は，次の関係でも結ばれる．

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)dudv$$

周辺分布関数，周辺密度関数はそれぞれ，

$$F_X(x) = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(u,v)dudv$$

$$F_Y(y) = \int_{-\infty}^{\infty} \int_{-\infty}^{y} f(u,v)dudv$$

$$f_X(x) = \int_{-\infty}^{\infty} f(x,v)dv$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(u,y)du$$

で与えられる．

　また離散型同様，独立性も定義できる．すべての x, y について

$$f(x,y) = f_X(x)f_Y(y) \tag{3.10}$$

であるとき，X と Y は独立であるという．

　また $g(X,Y)$ の期待値と，それを用いて計算される共分散も

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)f(x,y)dxdy$$

$$\sigma_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-\mu_X)(y-\mu_Y)f(x,y)dxdy$$

と定義される．ただし $\mu_X = E[X], \mu_Y = E[Y]$ である．他に相関係数も同様に定義される．

注意 16. 独立であれば共分散はゼロである．ただし，共分散がゼロであるからといって，独立とは限らない．

3.8　リスク

ファイナンスでは収益率や資産価格を確率変数と考える．その場合，リスクを分散で測ることが多い．

さて2つの確率変数XとYから作られる確率変数$aX+bY$の分散は，これまで示したことを用いると

$$Var[aX+bY]=a^2\sigma_X^2+b^2\sigma_Y^2+2ab\sigma_{XY}$$

である．このことは，2つの資産の共分散に関する性質によっては，2つの資産をともに所有することでリスクが小さくできる可能性を示唆している．

3.9　多次元分布

さらに複数の確率変数を考えて同時分布を考えることもできる．

演習 12. 各自，多次元の同時分布関数や同時密度関数を調べよ．

なお，ファイナンスでよく登場する正規分布に関しても多次元正規分布が考えられる．$\mathbf{X}=(X_1X_2\cdots X_n)$とするとき，$\mu_{\mathbf{X}}=E[\mathbf{X}]$とおいて$\Sigma=E[({}^t\mathbf{X}-{}^t\mu_{\mathbf{X}})(\mathbf{X}-\mu_{\mathbf{X}})]$と考えると，$\Sigma$は正の定符号行列である．

$$f(\mathbf{x})=(2\pi)^{-\frac{n}{2}}|\Sigma|^{-\frac{1}{2}}e^{-\frac{1}{2}({}^t\mathbf{X}-{}^t\mu_{\mathbf{X}})\Sigma(\mathbf{X}-\mu_{\mathbf{X}})} \tag{3.11}$$

であらわされる．

3.10　条件付確率

これまで，2つあるいはそれ以上の数の確率変数の同時分布を考える状況として，複数の金融商品の将来価格(あるいは収益率)が独立でないことを念頭においた．

しかし，複数の確率変数の間に独立性がないことを考える他の状況として，ある金融商品（たとえば株式）の異なる時点の価格同士に関心をもつ場合が挙げられる．その場合，最初に初期時点の株価と将来時点の株価の間の同時分布を考え，次に初期時点の株価が確実に知られた後の，将来時点の株価の確率分布を考える．このように初期時点での株価が実現した後の将来時点の株価の分布を，初期時点の株価の条件付確率分布という．この節では，条件付確率分布について学ぶ.

3.10.1　条件付確率：離散型

確率変数の対XとYのそれぞれのとりうる値の高々可算な集合を$\mathcal{X}$，$\mathcal{Y}$と記すことにすると，(x, y)という数値の対が実現する確率関数

$$p : \mathcal{X} \times \mathcal{Y} \longrightarrow \mathbf{R}$$

は同時分布を表現する．

定義 1. $x \in \mathcal{X}, y \in \mathcal{Y}$に対して，$p_X(x) \neq 0$のとき

$$\frac{p(x, y)}{p_X(x)}$$

を $X=x$ が与えられたときの $Y=y$ である条件付確率とよび，$p(y|x)$ あるいは $Pr(Y=y|X=x)$ と記す.

注意 17.

$$\begin{aligned}\sum_{y\in\mathcal{Y}} p(y|x) &= \sum_{y\in\mathcal{Y}} \frac{p(x,y)}{p_X(x)} \\ &= \frac{1}{p_X(x)} \sum_{y\in\mathcal{Y}} p(x,y) \\ &= \frac{p_X(x)}{p_X(x)} = 1\end{aligned}$$

であるから，条件付確率も確率分布である.

注意 18. 確率変数 X，Y が独立ならば，*(3.8)*が成立するから，任意の $x\in\mathcal{X}, y\in\mathcal{Y}$ に対して

$$p(y|x) = \frac{p_X(x)p_Y(y)}{p_X(x)} = p_Y(y)$$

である.

演習 13. 上の注意と逆に，任意の $x\in\mathcal{X}, y\in\mathcal{Y}$ に対して

$$p(y|x) = \frac{p_X(x)p_Y(y)}{p_X(x)} = p_Y(y)$$

であるとき，*2*つの確率変数 X，Y は独立である．このことを示せ．*(*超カンタン！*)*

定義 2. $X=x$ の条件付確率関数 $p(y|x)$ を与えたとき，確率変数 Y と実数値関数 $g(\cdot)$ から作られる確率変数 $g(Y)$ か

ら計算される

$$\sum_{y\in\mathcal{Y}} g(y)p(y|x) \tag{3.12}$$

を $X = x$ のもとでの $g(Y)$ の条件付期待値とよび，

$$E[g(Y)|X = x]$$

と記す.

ここで $E[g(Y)|X = x]$ は，$x \in \mathcal{X}$ に対してある実数を表すことに注意しよう．これにより写像

$$\mathcal{X} \ni x \longmapsto E[g(Y)|X = x] \in \mathbf{R} \tag{3.13}$$

を考えることができる．これは (3.13) が確率変数であることを意味している．このように実現した $x \in X$ を特定せず，確率変数のように考えた条件付期待値を

$$E[g(Y)|X]$$

と記す.

3.10.2　条件付確率：連続型

連続型を確率密度関数をもつ確率変数に限定すると，条件付確率の議論は，すぐ上の離散型の議論とほぼ平行して行なうことができる.

今同時確率密度関数 $f(x, y)$ を考える.

定義 3. $x \in \mathcal{X}, y \in \mathcal{Y}$に対して，$f_X(x) \neq 0$のとき

$$\frac{f(x,y)}{f_X(x)}$$

を$X = x$が与えられたときの$Y = y$である条件付確率密度関数とよび，$f(y|x)$と記す.

注意 19.

$$\begin{aligned}\int_{-\infty}^{\infty} f(y|x)dy &= \int_{-\infty}^{\infty} \frac{f(x,y)}{f_X(x)}dy \\ &= \frac{1}{f_X(x)}\int_{-\infty}^{\infty} f(x,y)dy \\ &= \frac{f_X(x)}{f_X(x)} = 1\end{aligned}$$

であるから，条件付確率も確率密度関数である.

注意 20. 連続型確率変数X，Yが独立ならば，*(3.10)*が成立するから，任意の$x \in \mathcal{X}, y \in \mathcal{Y}$に対して

$$f(y|x) = \frac{f_X(x)f_Y(y)}{f_X(x)} = f_Y(y)$$

である.

演習 14. 上の注意と逆に，任意の$x \in \mathcal{X}, y \in \mathcal{Y}$に対して

$$f(y|x) = \frac{f_X(x)f_Y(y)}{f_X(x)} = f_Y(y)$$

であるとき，*2*つの連続型確率変数X，Yは独立である. このことを示せ.

定義 4. $X = x$ の条件付確率密度関数 $f(y|x)$ を与えたとき，確率変数 Y と実数値関数 $g(\cdot)$ から作られる確率変数 $g(Y)$ から計算される

$$\int_{-\infty}^{\infty} g(y)f(y|x)dy \tag{3.14}$$

を $X = x$ のもとでの $g(Y)$ の条件付期待値とよび，離散型の場合同様

$$E[g(Y)|X = x]$$

と記す.

ここで $E[g(Y)|X = x]$ は，離散型の場合同様 $x \in \mathcal{X}$ に対してある実数を表すことに注意しよう．これにより写像

$$\mathcal{X} \ni x \longmapsto E[g(Y)|X = x] \in \mathbf{R} \tag{3.15}$$

を考えることができる．これは (3.15) が確率変数であることを意味している．このように実現した $x \in X$ を特定せず，確率変数のように考えた条件付期待値を

$$E[g(Y)|X]$$

と記す．これも離散型の場合同様である.

3.10.3　条件付確率の重要性

不確実な世界に直面するわれわれが，確率論によるアプローチをとるとき，最終的な目標は関心の対象となる確率変数群の同時分布の構造を知ることである.

しかし，神様はおいそれと不確実な世界の構造のヒントを，一挙に見せてくれることはない．大抵の場合，神様は一部の確率変数の実現値を小出しにわれわれに見せることで，上記の同時分布の構造のヒントとなさる．

確率変数としての条件付期待値の分布を明らかにすることは，関心の対象となる確率変数群の同時分布の構造を知ることの第1歩なのである．それゆえ，条件付確率や条件付期待値が重要性をもつのである．

確率論：より厳密な展開のために

以上で示した確率に関する議論は，密度関数を中心して解析学の範囲内で展開する連続型確率論と，形式的に似せるために離散型の確率論の展開を通常のテキストと異なる形で示してある．以下では，厳密に展開される確率論の雛型がどういうものかを示し，この授業で示した解説とどのように関係するかを述べる．

確率を考える数学的な枠組みは次の3つのものの組である．

1. 標本空間 Ω
2. $\sigma-$加法族 ($\sigma-$field) $\mathcal{S}$
3. 確率測度 P

3つをたばねて $(\Omega, \mathcal{S}, P)$ のように書いて，**確率空間**とよぶ．

標本空間 Ω は，観測や賭けの実行などの**試行**によって起

こりうる個々の結果の全体を指している．Ωはいかなる抽象的なものの集まりでもかまわない．

例 4. 「サイコロ*1*個を１回振る」という試行の結果の全体としての標本空間は，出た目の数に注目すると

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

である．

次に，$\sigma-$加法族$\mathcal{S}$は，標本空間の要素を集めた集合である事象(event)を集めたものである．例4において，事象とは例えば「出た目の数が偶数」という言明に対応する標本空間の要素を集めた集合

$$\{2, 4, 6\}$$

である．よって$\sigma-$加法族$\mathcal{S}$は，Ωのすべての部分集合を集めた冪集合(power set)2^{Ω}の部分集合になっている．$\mathcal{S}$の要素を可測事象とよび，すぐあとに解説する確率が計算できるような事象という意味でつけられている．

$$\mathcal{S} \subset 2^{\Omega}$$

である．さらに以下の要件を課す

1. $\Omega \in \mathcal{S}$
2. $A \in \mathcal{S} \Longrightarrow A^c \in \mathcal{S}$
3. $A_n \in \mathcal{S}\ (n \in \mathbf{N}) \Longrightarrow \bigcup_{n\in\mathbf{N}} A_n \in \mathcal{S}$

これらは，確率というものが整合的に議論するために必要な最低限の条件である．

注意 21. 要件 *1* と *2* から，$\phi \in \mathcal{S}$ がわかる．また要件 *3* は，有限個の可測事象の和集合も可測事象であることを含意することに注意せよ．*(なぜか？)*

最後に確率測度 P は，数学的には

$$P:\ \mathcal{S} \longrightarrow [0,1]$$

で，つぎの性質を満たすものである．

1. $(\forall A \in \mathcal{S})\ P(A) \geq 0$
2. $P(\Omega) = 1$
3. $\mathcal{S}$ から可算無限個の要素を $\{A_1, A_2, A_3 \dots, \}$ をとって，任意の 2 つの事象の対について $A_i \cap A_j = \phi$ であるとき $P(\bigcup_{n\in\mathbf{N}} A_n) = \bigcup_{n\in\mathbf{N}} P(A_n)$

3 番目に登場した，2 つの事象の対について $A_i \cap A_j = \phi$ であることを，事象が排反しているという．

注意 22. 確率測度 P については，以下の性質が導かれる．

1. $P(\phi) = 0$
2. $P(A) + P(A^c) = 1$
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
4. $A \subset B \Longrightarrow P(A) \leq P(B)$

例 5. Ωが有限集合であるとき，Ωの各要素*1*点からなる集合に等確率$\frac{1}{\#\Omega}$を与え，$\sigma-$加法族$\mathcal{S}$として2^{Ω}をとった確率空間を一様確率空間とよぶ．なお，$\#\Omega$はΩの要素の個数を表す．

この確率空間は，場合の数によって確率が規定されるモデルの確率空間に対応している．

すでにふれたように，標本空間は実数の集合，複素数の集合に限定される必要はない．(標本空間にあるカテゴリの関数の集合をとることはよくある．) しかし，実数の集合$\mathbf{R}$の部分集合を標本空間に，またその部分集合の開部分集合の全てを含む最小の$\sigma-$加法族を考えることも多い．その場合のボレル$\sigma-$加法族といい$\mathcal{B}$とかく．

実用的には，授業で示したように，生起する事象として数に関することが多い．まず標本空間の各点$\omega \in \Omega$に対して実数を対応させる写像

$$X:\ \Omega \longrightarrow \mathbf{R}$$

を考えることで，事前にどの数が生起するかわからないことを表現する．Xの値域は実数全体$\mathbf{R}$の部分集合である．さて

$$(\forall B \in \mathcal{B})\ X^{-1}(B) \in \mathcal{S}$$

であるとき，Xは可測であるといい，Xを確率変数とよぶ．

確率空間 $(\Omega, \mathcal{S}, P)$ と，確率変数が与えられるとき，

$$P_X(B) = P(X^{-1}(B)) \quad (B \in \mathcal{B})$$

で定義される写像

$$P_X : \mathcal{B} \longrightarrow [0, 1]$$

は確率測度となる．これを確率変数 X の分布とよぶ．

授業で確率変数と常に対で，確率関数あるいは確率分布 (あるいは確率密度関数) を扱ったが，実は背後に根源的な確率評価の構造，つまり確率空間率空間 $(\Omega, \mathcal{S}, P)$ があったと考えるのが正確なのであった．