中級経済数学 2 講義ノート

多鹿 智哉*

0. 講義について

0.1. 概要

位相，積分，確率論について学び，ゲーム理論への応用を考える．厳密性は重視せず，豊富なトピックを簡単な例とともに学ぶ．

0.2. 講義計画

1. オリエンテーション，
2. 位相と連続性
3. 開集合と閉集合
4. ワイヤシュトラスの極値定理
5. 積分の復習
6. 総和記号と級数の収束
7. 重積分：フビニの定理
8. 重積分：変数変換
9. 知識の理論
10. 確率論の基礎
11. 条件付き確率と条件付き期待値
12. 大数の（弱）法則
13. 応用：電子メールゲーム
14. 応用：同意定理
15. 応用：オークション

0.3. 文献案内

- Liang (2023) *Information and Learning in Economic Theory*, arXiv:2212.07521v2
- Maschlar, Solan, and Zamir (2013) *Game Theory*, Cambridge University Press.
- 多鹿 (2023)「読んで理解する経済数学」新世社—ウェブ付録
- 舟木 (2004)「確率論」朝倉書店

0.4. その他情報

評価基準	**講義スタイル**	**時間・場所**
課題：7 点 × 12 回 平常点：20 点	対面	後期水曜 4 限本館 40 番教室

*日本大学経済学部，tajika.tomoya@nihon-u.ac.jp

— Notes —

Date:

— Notes — Date:

課題は受講者自身の回答について口頭で説明できることを前提とし，適宜理解の確認を行う．なお，本講義ノートの空白部はメモ用スペースである．

1. 位相と連続性

1.1. 数列と収束

自然数の集合 $\mathbb{N}$ を考える．一般の集合，A について，関数 $a : \mathbb{N} \to A$ を数列と呼ぶ．このとき，a_n を数列 a の第 n 項と呼ぶ．数列 a を $(a_1, a_2, \ldots)$ と書くこともある．

集合 A に距離関数 $d : A \times A \to \mathbb{R}$ を定める．**距離関数**は以下の性質を持つ．

1. 任意の $x, y \in A$ について, $d(x, y) \geq 0$ かつ $d(x, y) = 0$ であることと $x = y$ は同値である．
2. 任意の $x, y \in A$ について，$d(x, y) = d(y, x)$
3. 任意の $x, y, z \in A$ について，$d(x, z) \leq d(x, y) + d(y, z)$

数列 a が A の点 $x \in A$ に**収束**するとは，任意の $\varepsilon > 0$ に対して，ある自然数 N が存在して，任意の $n \geq N$ について，$a_n \in A$ かつ $d(a_n, x) < \varepsilon$ が成り立つことである．

このとき，x を数列 a の**極限**と呼び，次のように書く．

$$\lim_{n \to \infty} a_n = x$$

$\xi : \mathbb{N} \to \mathbb{N}$ を $\xi(n) < \xi(n+1)$ が全ての n について成り立つとする．このとき，数列 $b_n = a_{\xi(n)}$ を数列 a の**部分列**と呼ぶ．

$\mathbb{R}$ を実数全ての集合とする. 集合 $A \subset \mathbb{R}$ について, $\sup(A) = \min\{x : x > y, \forall y \in A\}$ と定義する．これを A の**上限**と呼ぶ．同様に，$\inf(A) = \max\{x : x < y, \forall y \in A\}$ を A の**下限**と呼ぶ．

$\mathbb{R}$ には次の性質が成り立つ．

デデキントの切断公理

$\mathbb{R}$ の部分集合のペア (A,B) が $\mathbb{R}$ の切断を定めるとは，次の条件を満たすことである．

1. $A \cup B = \mathbb{R}$
2. $A \cap B = \emptyset$
3. 任意の $a \in A$，$b \in B$ について，$a < b$ が成り立つ．

このとき，切断 (A,B) について，A の最大元か B の最小元のどちらかが実数の集合 $\mathbb{R}$ に存在する．

デデキントの切断公理を用いれば例えば $\sqrt{2}$ が実数であることを示すことができる．

$\mathbb{R}$ 上の数列について，次の性質が成り立つ．これらの性質は同値である．

1. （**ボルツァーノ・ワイヤシュトラスの定理**）ある $M \in \mathbb{R}$ について，$a_n < M$ かつ $a_n \leq a_{n+1}$ が全ての $n \in \mathbb{N}$ に対して成り立つとき，$x = \lim_{n\to\infty} a_n \in \mathbb{R}$ が存在する．
2. ある $M \in \mathbb{R}_{++}$ について，$|a_n| < M$ が全ての $n \in \mathbb{N}$ に対して成り立つとき，収束する部分列が存在する．
3. 任意の $\varepsilon > 0$ に対して，ある自然数 N が存在して，任意の $n, m \geq N$ について，$a_n \in A$ かつ $d(a_n, a_m) < \varepsilon$ が成り立つとき，a を**コーシー列**と呼ぶ．このコーシー列は収束する．

数列を用いて関数の連続性を表現することができる．以下の定義は同値である．

1. 関数 $f : A \to \mathbb{R}$ が点 $x \in A$ で**連続**であるとは，任意の数列 $a : \mathbb{N} \to A$ が x に収束するとき，数列 $b_n = f(a_n)$ が $f(x)$ に収束することである．

— Notes —

Date:

— Notes — Date:

2. (ε-δ 論法) 関数 $f : A \to \mathbb{R}$ が点 $x \in A$ で**連続**であるとは，任意の $\varepsilon > 0$ に対して，ある $\delta > 0$ が存在して，任意の $y \in A$ について，$d(x, y) < \delta$ ならば $|f(x) - f(y)| < \varepsilon$ が成り立つことである．

課題 1.

1. 次の $n \to \infty$ としたときの極限を求めよ．
   a. $a_n = \frac{1}{n}$
   b. $b_n = (-1)^n$
   c. $c_n = 1 + 2^1 + 2^2 + \cdots + 2^n$
2. 数列 $f(n) = \left(1 + \frac{1}{n}\right)^n$ の極限を求めたい
   a. $f'(n)$ を計算しなさい．
   b. $f(n)$ が単調増加であることを示しなさい．
   c. $(1 + a)^n$ を展開しなさい．
   d. $a = \frac{1}{n}$ を代入することで，$2 < f(n) < 3$ を示しなさい．
   e. 以上を用いて，$f(n)$ が収束することを示しなさい．
3. ε-δ 論法を用いて，次の関数が $x = 2$ で連続であることを示しなさい．
   a. $f(x) = 2x$
   b. $g(x) = x^2$

1.2. 開集合と閉集合

任意の $x \in A$ に対して，$B_\varepsilon(x) = \{y \in \mathbb{R} : d(x, y) < \varepsilon\}$ を x の ε-開球，あるいは**開近傍**と呼ぶ．

集合 $A \subset \mathbb{R}^n$ と任意の $x \in A$ について，ある $\delta > 0$ が存在して，区間 $B_\delta(x) \subset A$ が成り立つとき，A を**開集合**と呼ぶ．開集合の補集合を**閉集合**と呼ぶ．開集合すべてを集めた集合族（集合の集合）を**位相**と呼び，$\mathcal{O}$ などと書く．

開集合については次の性質が成り立つ[2].

1. 全体集合と空集合は開集合である.
2. 有限個の開集合の共通部分は開集合である.
3. 任意の開集合の合併は閉集合である.

ただし，有限個の開集合の共通部分は開集合であるが，無限個の開集合の共通部分は開集合とは限らない．例えば，区間 $\left(-\frac{1}{n}, 2+\frac{1}{n}\right)$ は開集合であるが，その無限共通部分 $\bigcap_{n=1}^{\infty}\left(-\frac{1}{n}, 2+\frac{1}{n}\right) = [0, 2]$ は閉集合である.

関数の連続性は開集合を用いても定義される．関数 $f : A \to \mathbb{R}^m$ が**連続**であるとは，任意の $\mathbb{R}^m$ の開集合 V について，$f^{-1}(V) = \{x \in A \mid f(x) \in V\}$ が A の開集合であることである.

課題 2.

1. 無限個の開集合の共通部分が開集合でも閉集合でもないような例を提示し，そのことを説明しなさい.
2. $f(x) = x^2$ が $\mathbb{R}$ 上で連続であることを示しなさい.
3. 開集合を用いた連続性の定義と，ε-δ 論法を用いた連続性の定義が同値であることを示しなさい.

1.3. コンパクト性とワイヤシュトラスの極値定理

集合 A について，集合族 $\mathscr{C}$ が**開被覆**であるとは，任意の $U \in \mathscr{C}$ について，C が開集合であり，かつ $A \subset \bigcup_{C\in\mathscr{C}} C$ が成り立つことである．$\mathscr{C}' \subset \mathscr{C}$ が部分被覆であるとは，$A \subset \bigcup_{C\in\mathscr{C}'} C$ が成り立つことである.

集合 $A \subset \mathbb{R}^n$ が**コンパクト**であるとは，任意の A の開被覆（A を覆う開集合の集合）について，有限個の部分被覆が存在することである.

例えば通常の距離の元で定義される $\mathbb{R}$ において，閉区間 $[0, 1]$ はコンパクトである．これを確認するために，任意の $[0, 1]$ の開被覆 $\mathscr{C}$ を考える．この

[2] 逆に，次の性質が成り立つ集合族で開集合を定めることもできる．これは開集合系と呼ばれる

— Notes —

Date:

区間のうち，有限で覆えない部分の集合を S とする．その直径を r_S と書く．つまり，$r_S = \inf\{r : S \subset B_r(x) \text{ for some } x \in S\}$ である．ここで，S を一つ一つの直径が $r_1 = \frac{r_S}{2}$ となるように分割する．これらの分割のうち，再び $\mathscr{C}$ の中の有限個で覆えない部分を S_1 とする．今度は $r_2 = \frac{r_1}{2}$ となるように設定し，再び分割を行う．これを繰り返すことで，S の中に，直径が任意に小さくなるような区間列が存在することになる．そうすると，その交わりは一点 x_* に収束する．これは $\mathscr{C}$ のいずれかの開集合 C_* に含まれる．したがって，ある $r > 0$ が存在して，$B_r(x_*) \subset C_*$ が成り立つ．しかし，直径が $\frac{r}{2}$ より小さいような区間列が無限に存在するため，十分大きな n について，その中の一つが $B_r(x_*)$ に含まれる．これはその区間列が $\mathscr{C}$ の有限個で覆えることを意味し，矛盾である．したがって，有限個の部分被覆が存在する．

一方で，開区間 $(0,1)$ はコンパクトではない．実際，開区間 $(0,1)$ の開被覆として，集合族 $\left\{\left(\frac{1}{n}, 1-\frac{1}{n}\right)\right\}_{n\in\mathbb{N}}$ を考える．これは有限の部分被覆を持たない．

さらに，$\mathbb{R}$ もコンパクトではない．例えば，開被覆として，集合族 $\{(-n,n)\}_{n\in\mathbb{N}}$ を考える．これは有限の部分被覆を持たない．

もし距離がユークリッド距離であれば，より簡単な確認方法がある．

ハイネ・ボレルの被覆定理

$d(x,y) = \sum_i |x_i - y_i|^2$ とする．A がコンパクトであることと，A が閉集合かつ有界であることは同値である．

注意したいのは，距離関数 d によってはこの同値性は成り立たない．例えば，$d(x,y) = \frac{|x-y|}{|x-y|+1}$ は距離だが，これが導く位相に従えば，任意の集合が有界であるため，$\mathbb{R}$ は有界閉集合である．しかしこれはコンパクトではない．

コンパクト性は関数の極値の存在を保証する．

ワイヤシュトラスの極値定理

関数 $f : A \to \mathbb{R}$ がコンパクト集合 A 上で連続であるとき，ある $x_*, x^* \in A$ が存在して，任意の $x \in A$ について，$f(x_*) \leq f(x) \leq f(x^*)$ が成り立つ．

これを証明するため，まず, A が有限集合のケースを考えよう．このとき，自明に最大値と最小値が存在する．

次に A が有限集合でないときを考える．任意の $x \in A$ について，$C_x = \{y \in A : f(x) > f(y)\}$ と定める．今, 最大値が存在しないことを仮定するとどんな y についても，$y \in C_x$ となるような $x \in A$ が見つかる．したがって，$A \subset \{C_x \mid x \in A\}$ である．また，C_x は開集合であるので，$\mathscr{C} = \{C_x \mid x \in A\}$ は A の開被覆である．したがって，有限個の部分被覆 $\mathscr{C}' \subset \mathscr{C}$ が存在して，$A \subset \bigcup_{C \in \mathscr{C}'} C$ が成り立つ．しかし，$\mathscr{C}'$ が有限集合であるため，その中での最大化解 $x_{\max}$ が存在する．言い換えると，任意の $y \in C_x \in \mathscr{C}'$ について，$f(x_{\max}) \geq f(y)$ となる．しかし $A \subset \mathscr{C}'$ であるので，$x_{\max}$ は f の最大化解である．これは矛盾である．したがって，最大値が存在する．同様にして最小値も存在することが示される．

課題 3.

1. ハイネ・ボレル性を用いて $A = [0,1]$ がコンパクトであることを示せ．

2. リーマン積分

$a < b$ を満たす実数，a, b に対して，閉区間 $[a, b]$ を考える．$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$ となるような数列 $x = (x_0, x_1, \ldots, x_n)$ を考える．このとき，それぞれの $i = 1, \ldots, n$ について，区間 $[x_{i-1}, x_i]$ が定まり，その中の代表点 ξ_i を一つ取る．このとき，関数 $f(x)$ に対して，次の和を考える．

$$S(f, x, \xi) = \sum_{i=1}^{n} f\left(\xi_i\right)(x_i - x_{i-1})$$

これを**リーマン和**と呼ぶ．そして，そのときの最大の幅を $\Delta = \max_{i=1,\ldots,n}(x_i - x_{i-1})$ とおく．このとき，$f(x)$ が $[a, b]$ 上で有界であると

— Notes —

Date:

き，Δ が 0 に近づく極限を考える．この極限が存在するとき，その値を $f(x)$ の $[a,b]$ 上での**リーマン積分**といい，次のように表す．

$$\int_a^b f(x)\,\mathrm{d}x$$

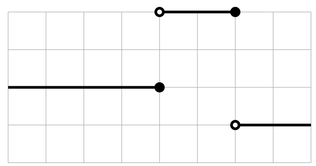
リーマン積分は，f の原始関数を F とすると，次のように表すことができる．

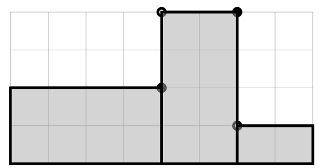
$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$$

また，$F(b)-F(a)$ を $[F(x)]_a^b$ と書くこともある．ただし，F が f の原始関数であるとは，$F'(x)=f(x)$ が成り立つことである．

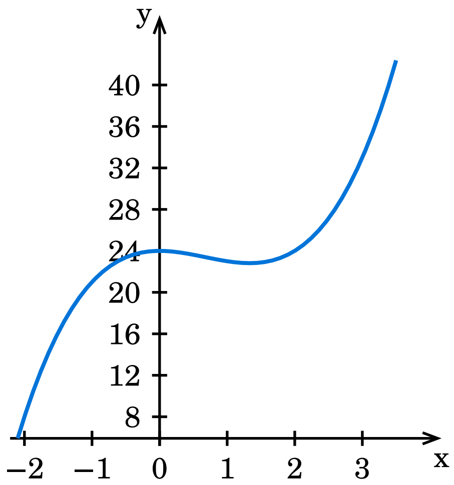
2.1. 積分と面積

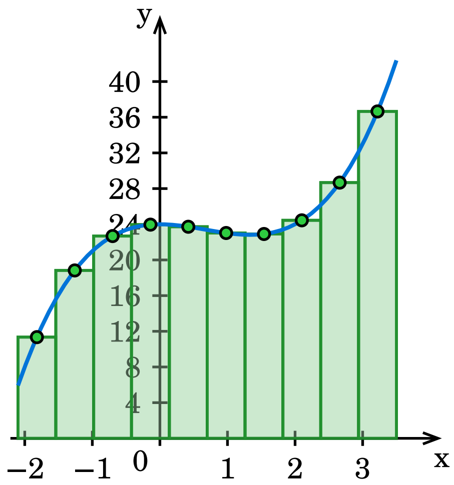
積分は関数のグラフが与える面積と一致する．例えば，次のグラフを見てみよう．



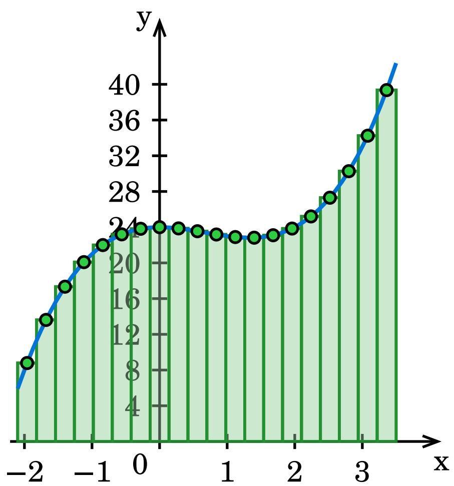


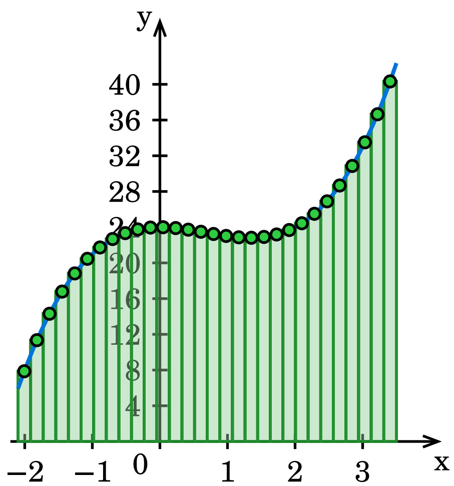
上記のグラフ（左）が与える面積は，グラフ（右）の色が塗ってある長方形の面積を全て足したものである．これは簡単に計算できることだろう．では，次のようなグラフ（左）の面積はどうやって求めたら良いか？





ひとつのアイデアはグラフ（右）のように，細かく分割をとって，そこで得られる短冊の面積を計算することである．これがリーマン和の正体である．この短冊を十分細かくしていくと，次のように実際の面積に近づく．





ではこのリーマン和の値がなぜ，原始関数として書けるのか，その理由を考えてみよう．これは，平均値の定理によるものである．つまり一定の条件のもとで，以下の公式が成り立つような ξ_i が存在する．

$$F(x_i) - F(x_{i-1}) = f\left(\xi_i\right)(x_i - x_{i-1})$$

— Notes —

Date:

— Notes —

Date:

従ってこれを全ての i について足し合わせると，次のようになる．

$$F(b)-F(a)=\sum_{i=1}^{n} f\left(\xi_i\right)(x_i-x_{i-1})$$

この右辺はリーマン和の定義そのものである．

2.2. 積分の計算公式

積分には次の計算公式が成り立つ．

1. **微分積分学の基本定理：**

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(z)\,\mathrm{d}z=f(x)$$

2.
$$\int_a^b f(x)\,\mathrm{d}x=-\int_b^a f(x)\,\mathrm{d}x$$

3.
$$\int_a^b f(x)\,\mathrm{d}x+\int_b^c f(x)\,\mathrm{d}x=\int_a^c f(x)\,\mathrm{d}x$$

4. **部分積分：**

$$\int_a^b f(x)G(x)\,\mathrm{d}x+\int_a^b F(x)g(x)\,\mathrm{d}x=[F(x)G(x)]_a^b$$

これは，積の微分の公式, $[F(x)G(x)]'=f(x)G(x)+F(x)g(x)$ の両辺を積分した結果である．

5. **変数変換：**逆関数を持つ関数 g について，

$$\int_a^b f(g(x))\,\mathrm{d}x=\int_c^d f(u)h'(u)\,\mathrm{d}u$$

ただし，$h(u)=g^{-1}(x)$ である．これは，$g(x)=u$，言い換えると $x=h(u)$ の両辺を微分して，$\mathrm{d}x=h'(u)\,\mathrm{d}u$ として，左辺に代入していると考えるとわかりやすい．ただし，積分範囲も変換する必要がある．元々の x の変域が $[a,b]$ であれば，u の積分範囲は g の値域, $g([a,b])=\{y\in\mathbb{R}\mid y=g(x),x\in[a,b]\}$ である．ここではその値域が $[c,d]$ であるとしている．

課題 4.

1. $\int_0^a e^{-x}\,\mathrm{d}x$ を計算しなさい．

2. $\int_0^a e^{-x}x\,\mathrm{d}x$ を計算しなさい.
3. $\int_0^a xe^{-x^2}\,\mathrm{d}x$ を計算しなさい.
4. 1, 2 で，$a \to \infty$ としたときの極限を求めなさい.
5. $\int_n^{n+1} \frac{1}{x}\,\mathrm{d}x \leq \frac{1}{n}$ であることを示したうえでそのことを利用し, 以下を示しなさい.

$$\int_1^\infty \frac{1}{x}\,\mathrm{d}x \leq \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

2.3. 多重シグマ記号

2.3.1. 多重シグマ記号と順序交換

多変数関数の積分の話の前に，シグマ記号を重複させて使う場合の意味について説明する．例えば次の記号はどのような意味を持つか考えてみる．

$$\sum_{i=1}^{n}\sum_{j=1}^{m} ij$$

この表な表記があるとき，基本は内側から計算する．つまり, j についての和を先に計算し，その後に i についての和を計算する．j についての計算をするときは, i は定数として扱うのでシグマ記号の外に出せることを利用すると次のような計算が可能である．

$$\begin{aligned}\sum_{i=1}^{n}\sum_{j=1}^{m} ij &= \sum_{i=1}^{n} i \sum_{j=1}^{m} j \\ &= \sum_{i=1}^{n} i(1+\cdots+n) \\ &= (1+\cdots+m)\sum_{i=1}^{n} i \\ &= (1+\cdots+n)(1+\cdots+m)\end{aligned}$$

また，これは順序を交換しても値は変わらない．実際，このようになる．

— Notes —

Date:

— Notes — Date:

$$
\begin{aligned}
\sum_{j=1}^{m}\sum_{i=1}^{n} ij &= \sum_{j=1}^{m} j \sum_{i=1}^{n} i \\
&= \sum_{j=1}^{m} j(1+\cdots+n) \\
&= (1+\cdots+m)\sum_{j=1}^{m} j \\
&= (1+\cdots+m)(1+\cdots+n)
\end{aligned}
$$

ここでは i,j は $i\times j$ というただの掛け算であったが，i,j に関する関数を考えることもできる．例えば，$f(i,j)=i^2 j$ などでも良い．そうすると次のようなものも考えることが可能である．

$$
\sum_{i=1}^{n}\sum_{j=1}^{m} f(i,j)
$$

これもシグマ記号の順序を入れ替えても良い．この理由は以下の表を見ればわかる．つまり，上の式は以下の表の $f(i,j)$ を全て足したものを計算するということで，これはどんな順番で足しても同じ値を得ることができる．

		i		
		1	$\cdots$	n
j	1	$f(1,1)$	$\cdots$	$f(n,1)$
	$\vdots$	$\vdots$	$f(i,j)$	$\vdots$
	m	$f(1,m)$	$\cdots$	$f(n,m)$

ただし，シグマ記号の順序の入れ替えは，$f(i,j)$ が有限の値を持つ場合に限られる．また，i,j の範囲が有限であることも前提となる．これが無限であれば，足し算の順序に総和の値が依存してくることがある．次の節でそのことを確認しよう．

2.3.2. 級数の収束

簡単な例で確認しよう．次の数列を見てみる．

$$1, -1, 1, -1, 1, -1, \ldots$$

これは 1 と -1 が交互に無限に繰り返される数列である．この数列を $a = (a_1, \ldots)$ とおいたときこの数列の和で作られた数列 $b_n = \sum_{i=1}^{n} a_i$ を考える．すると b_n は次のようになる．

$$1, 0, 1, 0, 1, 0, \ldots$$

これは収束しない．したがって，無限和 $\sum_{i=1}^{\infty} a_i$ は $\lim_{i \to \infty} b_i$ として定義することができなくなる．

さらに困ったことに，この数列は足し算の順番を入れ替えただけで違う値になり得る．例えば，1 だけを全て足してから，あとで -1 を足していくことにしよう．そうすると，1 が無限にあるので，その和は無限大になる．そこにいくら -1 を足していっても変わらない．まったく逆のことをすれば $-\infty$ にもなる．

したがって, 無限和を考えるときは, その収束性を確認する必要がある. そのために，次のような定義を考える．

1. 級数 $\sum_{i=1}^{\infty} a_i$ が**収束**するとは，以下の極限が存在することである．
$$\lim_{n \to \infty} \sum_{i=1}^{n} a_i$$
2. 級数 $\sum_{i=1}^{\infty} a_i$ が**絶対収束**するとは，以下の極限が存在することである．
$$\lim_{n \to \infty} \sum_{i=1}^{n} |a_i|$$

— Notes — Date:

特に絶対収束するときには，数列の足す順序を入れ替えても収束することが知られている．一方で収束する級数であっても絶対収束しない場合には，足し算の順序を入れ替えると計算結果が変わったり，収束しないことが知られている（リーマンの再配列定理）．多重シグマ記号とその順序交換についても同じようなことが言える．

積分は定義から無限和を考えなくてはならない．積分周りのややこしいところはこの辺りに起因する．

2.4. 重積分とフビニの定理

多変数関数 $f(x_1,\ldots,x_n)$ を考える．この関数の積分を考える．そのために多少の準備を行う．n 個の集合 $A_1,\ldots,A_n$ について，以下のように定まる集合を**直積集合**と呼ぶ．

$$A_1\times\cdots\times A_n=\{(a_1,\ldots,a_n)\mid a_i\in A_i\}$$

次に $i=1,\ldots,n$ について，閉区間 $[a_i,b_i]$ を考える．このとき，n 次元の直積集合 $[a,b]=[a_1,b_1]\times\cdots\times[a_n,b_n]$ を考える． この集合は**直方体**とも呼ばれる．このそれぞれの区間について

$$a_i=x_{i,0}<x_{i,1}<\cdots<x_{i,m-1}<x_{i,m}=b_i$$

となるような数列 $x_i=\left(x_{i,0},\ldots,x_{i,n}\right)$ を考える．このとき，次のような直方体を考えることができる．

$$\left[x_{1,j_1},x_{1,j_1+1}\right]\times\cdots\times\left[x_{n,j_n},x_{n,j_n+1}\right]$$

このような直方体は m^n 個存在する．このとき，この直方体の代表点を一つ取り，$\xi_{j_1,\ldots,j_n}$ と書く．ただし，$j_i=0,\ldots,m-1$ である．そうすれば次のような和を考えることができる．

$$\sum_{j_1=0}^{m-1}\cdots\sum_{j_n=0}^{m-1}f\left(\xi_{j_1,\ldots,j_n}\right)\left(x_{1,j_1+1}-x_{1,j_1}\right)\times\cdots\times\left(x_{n,j_n+1}-x_{n,j_n}\right)$$

直方体の幅の最大値を Δ と置く．具体的には次の通りである．

$$\Delta=\max_{j_1,\ldots,j_n}\max_{i=1,\ldots,n}\left(x_{i,j_i+1}-x_{i,j_i}\right)$$

このとき，$f(x_1,\ldots,x_n)$ が直方体 $[a,b]$ 上で有界であるとき，Δ が 0 に近づく極限を考える．この極限が存在するとき，その値を $f(x_1,\ldots,x_n)$ の $[a_1,b_1]\times\cdots\times[a_n,b_n]$ 上での**リーマン重積分**といい，次のように表す．

$$\int_{[a,b]} f(x_1,\ldots,x_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_n$$

リーマン重積分は，各変数についての積分の繰り返しによって計算できる．つまり

$$\int_{[a,b]} f(x_1,\ldots,x_n)\,\mathrm{d}x_1\cdots\mathrm{d}x_n = \int_{a_1}^{b_1}\cdots\left(\int_{a_n}^{b_n} f(x_1,\ldots,x_n)\,\mathrm{d}x_n\right)\cdots\mathrm{d}x_1$$

さらに，積分する順序によらず，同じ値を得ることができる（**フビニの定理**）．

具体的に計算してみよう．

$$\int_0^1\int_0^1 (x^2y^3+2x)\,\mathrm{d}x\,\mathrm{d}y$$

2.5. 変数変換

変数変換は，積分の計算を簡単にするために行う操作である．これが重積分の場合，特に重要である．これを見るため，二変数関数 $f(x,y)$ の積分，

$$\int f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

を考える．この変数に対して，$x=\varphi_x(u,v), y=\varphi_y(u,v)$ となるベクトル値関数 $\varphi=\left(\varphi_x,\varphi_y\right)$ を考える．すると，次が成り立つ，

$$\int f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int f\left(\varphi(u,v)\right)|J(u,v)|\,\mathrm{d}u\,\mathrm{d}v.$$

ただし，

$$J(u,v)=\det\left(\begin{bmatrix}\frac{\partial\varphi_x}{\partial u} & \frac{\partial\varphi_x}{\partial v}\\ \frac{\partial\varphi_y}{\partial u} & \frac{\partial\varphi_y}{\partial v}\end{bmatrix}\right)$$

であり，J は**ヤコビアン**と呼ばれる．

この式が成り立つ理由を見るため，$f(x,y)$ の積分を総和記号の極限として考えてみる．

— Notes —

Date:

$$\int f(x,y)\,\mathrm{d}x\,\mathrm{d}y \approx \sum_{i=1}^{n}\sum_{j=1}^{m} f\left(x_i, y_j\right)(x_i - x_{i-1})\left(y_j - y_{j-1}\right)$$

ここで，テイラー展開を考えれば，つぎのように書くことができる．

$$x_i - x_{i-1} = \frac{\partial \varphi_x}{\partial u}\Delta u + \frac{\partial \varphi_x}{\partial v}\Delta v$$

$$y_j - y_{j-1} = \frac{\partial \varphi_y}{\partial u}\Delta u + \frac{\partial \varphi_y}{\partial v}\Delta v$$

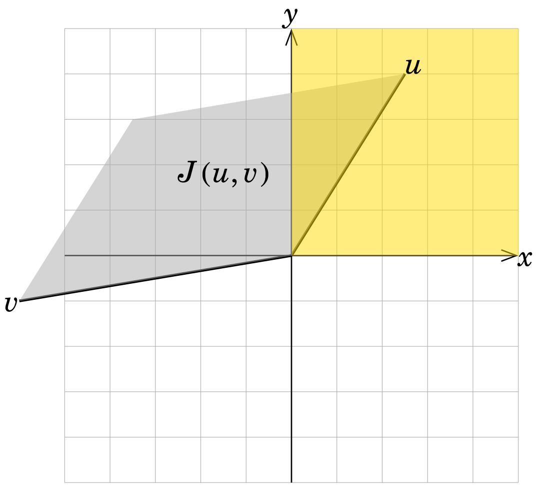
ここで，$\Delta u, \Delta v$ は $x_{i+1} - x_i \to 0, y_{j+1} - y_j \to 0$ になると，0 になるようなものである．

つぎに，$(x_i - x_{i-1}) \times \left(y_j - y_{j-1}\right)$ は i,j 番目の分割の面積である．これを，u,v の関数として表現すると，

$$\begin{bmatrix} x_i - x_{i-1} \\ y_i - y_{i-1} \end{bmatrix} = \begin{bmatrix} \frac{\partial \varphi_x}{\partial u} & \frac{\partial \varphi_x}{\partial v} \\ \frac{\partial \varphi_y}{\partial u} & \frac{\partial \varphi_y}{\partial v} \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix}$$

となる．行列式 $J(u,v)$ が，この行列が表す平行四辺形の面積を表すことを思い出す．そうすれば，i,j 番目の分割の面積は $|J(u,v)|\Delta u \Delta v$ と書ける，そうすれば，$(x_i - x_{i-1})\left(y_j - y_{j-1}\right) = |J(u,v)|\Delta u \Delta v$ というように u,v で書き直せることになる．後はこれを積分の総和に代入すれば，変数変換の公式が導かれる．

この行列が表す「平行四辺形」というのは以下の図のようなものである．



2.5.1. ガウス積分

変数変換の応用として，次の積分を考えよう．

$$\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x$$

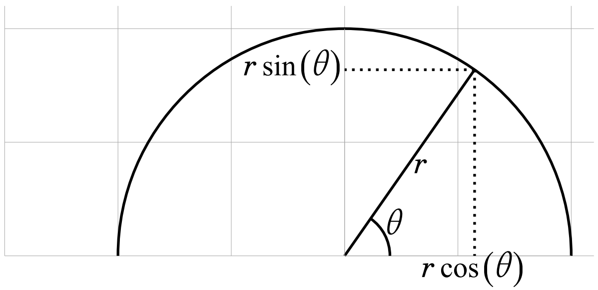
この積分を行うため，次の重積分を考える

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y$$

ここで，以下のように計算できることに注意する．

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y &= \int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x \int_{-\infty}^{\infty} e^{-y^2}\,\mathrm{d}y \\ &= \left(\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x\right)^2\end{aligned}$$

この積分を計算するために，$x = r\cos\theta, y = r\sin\theta$ という変数変換（**極座標変換**）を行う．



— Notes —

Date:

— Notes —

Date:

ただし，$(\sin\theta)' = -\cos\theta$, $(\cos\theta)' = \sin\theta$, $(\sin\theta)^2 + (\cos\theta)^2 = 1$ であることに注意せよ．そうすると，次のように計算ができる．

$$\begin{aligned}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y &= \int_0^{2\pi}\int_0^{\infty} e^{-r^2} r\,\mathrm{d}r\,\mathrm{d}\theta \\ &= \int_0^{\infty} e^{-r^2} r\,\mathrm{d}r \int_0^{2\pi}\mathrm{d}\theta \\ &= 2\pi\int_0^{\infty} e^{-r^2} r\,\mathrm{d}r\end{aligned}$$

$\left(e^{-r^2}\right)' = 2e^{-r^2}r$ であることに注意すれば，$2\pi\int_0^{\infty} e^{-r^2} r\,\mathrm{d}r = \pi$ である．よって，

$$\left(\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x\right)^2 = \pi$$

$$\Longrightarrow \int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}$$

課題 5.

1. 次の式はどのような式に書き直せるか，答えなさい．

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{\ell} ijk$$

2. $I_1, ..., I_n$ を考える．ただし，$I_1 < ... < I_n$ とする．このとき，

$$G = 1 - \frac{1}{n^2\mu}\sum_{i=1}^{n}\sum_{j=1}^{n}\min\{I_i, I_j\}$$

をジニ係数と呼ぶ. ただし, $\mu = \sum_{i=1}^{n}\frac{I_i}{n}$ である. また, $f(k) = \frac{\sum_{i=1}^{k} I_i}{\sum_{i=1}^{n} I_i}$ としたとき, f が表す曲線を**ローレンツ曲線**と呼ぶ. このとき, $\frac{1-G}{2}$ はローレンツ曲線の面積と等しい．このことを $n = 3$ のケースで確認しなさい．

3. 以下の積分を計算しなさい．

$$\int_{-1}^{1}\int_{-\infty}^{\infty} e^{-(x^2+y^2+2xy)}\,\mathrm{d}x\,\mathrm{d}y$$

3. 知識の理論

何か物事を決定するとき，その物事についての知識が必要である．不確実性下の意思決定では，何が起きるか，どこまで知っていて，どこまでを知らないかを定式化しなければいけない．ゲーム理論では，人々は他人の行動を

予測する．このため，お互いに何を知っていて，何を知らないかを定式化する必要がある．それらの第一歩として，この章では知識の理論について学ぶ．つまり「知っていること」とはどういうことかを定式化する．

3.1. 知識の定義

ω を**世界の状態 (state of the world)**，あるいは単に**状態**と呼ぶ．ω は集合 Ω の要素である．Ω は**状態空間**と呼ばれる．Ω の部分集合を**事象**と呼ぶ．事象は ω の集合である．ω が事象 A に含まれるとき，$\omega \in A$ と書く．

一般に集合 A について，A の**分割**とは，A の部分集合の集合 $\{A_1, ..., A_n\}$ で，$A_i \cap A_j = \emptyset$ かつ $A = \bigcup_{i=1}^n A_i$ を満たすものである．

ある人 i の**知識構造**とは，Ω の分割のことをいう．この人の知識構造を $\mathscr{F}_i = \left\{F_i^1, ..., F_i^n\right\}$ と書く．例えば $\Omega = \{1, 2, 3, 4, 5, 6\}$ であれば $\{\{1, 2, 3\}, \{4\}, \{5, 6\}\}$ はその例である

また，$\omega \in \Omega$ について，$F_i(\omega) \in \mathscr{F}_i$ を $\omega \in F_i^k$ となるような $F_i^k \in \mathscr{F}_i$ と定める．より具体的には，状態 ω において，事象 A を**知っている**とは，$F_i(\omega) \subset A$ であることをいう．また，知識関数 K_i を次のように定義する．

$$K_i(A) = \{\omega \in \Omega \mid F_i(\omega) \subset A\}$$

「A が起きたと知っている」ということは「区別がつくこと」である．例えば $\{\{1, 2, 3\}, \{4\}, \{5, 6\}\}$ という知識構造のもとでは，$1, 2, 3$ の区別がつかない．しかし，4 ではなく $1, 2, 3$ のうち，いずれかが起きたことは区別ができるのである．

もう少し具体的な例でいってみよう．W が今日の天気，Z が株式市場の様子であるとする．それぞれ，$W \in \{0, 1\}, Z \in \{2, 3\}$ としてみよう．$0, 1$ にはテキトーに晴れ，雨，$2, 3$ には好況，不況などを当てはめてみれば解釈しやすいかもしれない．そうするとあり得る世界の状態は $\Omega = \{(0, 2), (0, 3), (1, 2), (1, 3)\}$ である．これに対して知識構造として以下の例を考えてみよう．

$$\{\{(0, 2), (0, 3)\}, \{(1, 2), (1, 3)\}\}$$

これは今日の天気はわかるが株式市場の様子がわからない人の知識構造を示している．例えば $\omega = (0,2)$ であれば，自分が $\{(0,2),(0,3)\}$ のどこかにいることがわかる．すると $W = 0$ がおきたことが確定するので，「$W = 0$ であることがわかる」ということを表現しているのである．反対にこの人は株式市場の様子はわからない．自分が $\{(0,2),(0,3)\}$ のどこかにいることがわかっていても，$Z = 2$ も $Z = 3$ の可能性も両方あるからである．

知識関数には次の性質がある．

1. $K_i(A) \subset A$
2. $K_i(A) \subset K_i(B)$ if $A \subset B$
3. $K_i(K_i(A)) = K_i(A)$

$K_i(A)$ は個人 i が A が起きたことを知っているような状態の集合である．

3.2. 相互知識と共有知識

複数の個人がいる状況を考える．個人の集合を $\mathcal{I}$ とする．事象 A が**相互知識**であるとは，A が起きたことを全員が知っている状態のことである．つまり，$\omega \in A$ ならば，$\omega \in K_i(A)$ for all $i \in \mathcal{I}$ である．集合の表記を用いれば，$\bigcup_{i\in\mathcal{I}} K_i(A)$ である．

A が相互知識であるということは, A が起きたことを全員が知っている．しかし，そのことを各個人がお互いに知っているとは限らない．例えば次の分割を考えよう．

$$\mathscr{F}_1 = \{\{1,2,3\},\{4,5,6\}\}$$

$$\mathscr{F}_2 = \{\{1,2\},\{3,4\},\{5,6\}\}$$

このとき，$\omega = 1$ において，$A = \{1,2,3\}$ は相互知識である．しかし個人 1 は個人 2 が知っていることを知らない．なぜならもし，$\omega = 3$ であれば，個人 2 は A が起きていると思わない．しかし，個人 1 は $\omega = 1$ と $\omega = 3$ の区別がつかないので「個人 2 が A が起きたことを知っている」ことを知らないのである．これを数学的に表現すれば，$1 \notin K_1(K_2(A))$ である．もっと言えば，$K_1(K_2(A)) = \emptyset$ である．

このように,「他人が知っていること」に関する知識を **2 階の知識**と呼ぶ.事象そのものに関する知識はこれに対して 1 階の知識とよばれる.さらに「他人の 2 階の知識」に関する知識を 3 階の知識などと高階層の知識を定義することもできる.A が**共有知識**であるような状態の集合は次のように定義される集合の極限として定義される.

$$\mathscr{A}^n = \bigcap_{i \in \mathscr{I}} K_i\left(\mathscr{A}^{n-1}\right)$$

ここで,$\mathscr{A}^0 = A$ である.このように,共有知識は,「「全員が A を知っていること」を全員が知っていること」を全員が知っていることを…を知っているということを表す.

例えば以下の例を考えてみよう.

$$\mathscr{F}_1 = \{\{1,2,3\},\{4,5,6\}\}$$

$$\mathscr{F}_2 = \{\{1,2\},\{3\},\{4\},\{5,6\}\}$$

今,$A = \{1,2,3\}$ であるとする.そうすると,$K_1(A) = \{1,2,3\}$ であり,$K_2(A) = \{1,2,3\}$ である.そうすると,$K_1(A) = A = K_2(A)$ であるので,$\mathscr{A}^1 = A$ である.同様にして,$\mathscr{A}^2 = \mathscr{A}^3 = \cdots = A$ となるので,$A = \{1,2,3\}$ は共有知識になる.

一方で,$A = \{1,2\}$ はどうだろうか? このときは $K_1(A) = \emptyset$ であり,$K_2(A) = \{1,2\}$ である.つまり $\mathscr{A}^1 = \emptyset$ となる.$K_i(\emptyset) = \emptyset$ であるから,$A = \{1,2\}$ は共有知識とはなり得ないのである.

ゲーム理論では何が共有知識になっているのかが重要である.これは個人の行動は他者の行動の予測に依存しており,正確な予測を行うには「他者が何を知っているのか」をちゃんと認識する必要があるからである.そして,他者の行動も他者から見た自分自身についての予測に影響している.このような高階の知識がお互いの行動を決定する上で重要になってくるのである.この実例は,のちの応用の章で検討することにしよう.

課題 6.

1. つぎの知識関数の特徴を証明しなさい

— Notes —

Date:

a. $K_i(A) \subset A$

b. $K_i(A) \subset K_i(B)$ if $A \subset B$

c. $K_i(K_i(A)) = K_i(A)$

2. $\Omega = \{1,2,3,4,5,6,7\}$, $\mathscr{I} = \{a,b\}$ とする．そして，

$$\mathscr{F}_a = \{\{1,2\},\{3\},\{4,5,7\},\{6\}\}$$

$$\mathscr{F}_b = \{\{1,2,4,5\},\{3,6\},\{7\}\}$$

とする．共有知識となりうる事象をすべて求めなさい．

3. (Muddy children's puzzle) 二人の個人が帽子をかぶっている．各個人は自分の帽子の色を見ることができないが，他人の帽子の色は見える．帽子の色は赤か黒かのどちらかである．また，少なくとも一人の帽子が赤色であることがわかっている．

a. 帽子の色の組み合わせを状態としたとき，状態の集合を書き下しなさい

b. 各個人の知識構造を書き下しなさい

c. 共有知識となりうる事象をすべて求めなさい．

d.「誰も自分の帽子の色がわからない」ということがわかった．このとき，各個人の帽子の色を求めなさい．

4. ある分割 $\mathscr{F}$ が個人 1，2 の知識構造の**共通粗大化**であるとは，$F_1 \cap F_2 \neq \emptyset$ となるようなどんな $F_1 \in \mathscr{F}_1, F_2 \in \mathscr{F}_2$ についても $F_1, F_2 \subset F \in \mathscr{F}$ となる F が見つかるようなことをいう．例えば $\mathscr{F}_1 = \{\{1,2\},\{3\},\{4\}\}$, $\mathscr{F}_2 = \{\{1\},\{2,3\},\{4\}\}$ であれば，$\mathscr{F} = \{\{1,2,3\},\{4\}\}$ や $\mathscr{F}' = \{\{1,2,3,4\}\}$ は個人 1，2 の共通粗大化になっている．しかし $\mathscr{F}'' = \{\{1,2\},\{3,4\}\}$ はそうではない．これは $\{2,3\} \in \mathscr{F}_2$ であるが，$\{2,3\}$ を含むような集合が $\mathscr{F}''$ にはないからである．このもとで，「共有知識となりうる事象」の集合が個人 1，2 の知識構造の共通粗大化になっていることを 1 の例を用いて確かめなさい．

4. 確率論の基礎

知識の理論では「知っている」とはどういうことかを定義した．しかし，「知っている」まではいかなくとも，「起こる可能性が高い」ということを考えることができる．例えば，以下の知識構造では，$\omega = 1$ であるとき，実際に起きているのが 1 か 2 か 3 かの区別がつかない．

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

しかし，1 が起きたとき，1 か 2 か 3 のうち，どれがどれほど起こりやすいのか，という予想を持つことは自然であり，そしてそれが意思決定にも重要である．確率論は，このような不確実性の下での意思決定を考えるための数学的な枠組みを提供する．

4.1. 確率の定義

Ω を状態の集合とする．Ω の部分集合の集合 $\mathscr{F}$ が次の性質を満たすとき，**σ-代数**と呼ぶ．

1. $\Omega \in \mathscr{F}$
2. $A \in \mathscr{F}$ ならば，$\Omega \setminus A \in \mathscr{F}$
3. $A_1, A_2, \ldots \in \mathscr{F}$ ならば，$\bigcup_{i=1}^{\infty} A_i \in \mathscr{F}$

確率は σ-代数上の関数として定義される．すなわち，確率関数 P は，$\mathscr{F}$ 上の関数であり，次の性質を満たす．

1. $P(A) \geq 0$ for all $A \in \mathscr{F}$
2. $P(\Omega) = 1$
3. (**加算加法性**) $A_1, A_2, \ldots \in \mathscr{F}$ が $A_i \cap A_j = \emptyset$ をみたすとき，
   $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

$A \subset \Omega$ が起きる確率，$P(A)$ は一定の条件のもと，以下のような積分としても書くことができる．

$$P(A) = \int_A p(\omega)\, \mathrm{d}\omega$$

p はラドン=ニコディム微分と呼ばれるものの一種である．

— Notes —

Date:

4.2. 確率変数

4.2.1. 確率変数と確率分布

集合 Z を適当に取る．$X:\Omega\to Z$ を Z 上の**確率変数**という．$Z\subset\mathbb{R}$ であるとき，確率変数について，次の用語を定める．

1. $G(b)=P(X(\omega)\le b)$ を確率変数 X の**分布関数**という．
2. $G(b)=\int_{-\infty}^{b}g(x)\,\mathrm{d}x$ と書けるとき，g を X の**確率密度関数**という．
3. $E(X)=\int_{-\infty}^{\infty}xg(x)\,\mathrm{d}x$ を X の**期待値**という．
4. $\int_{-\infty}^{\infty}(x-z)^k g(x)\,\mathrm{d}x$ を X の z **周りの** k **次のモーメント**という．期待値周りの 2 次のモーメントは**分散**と呼ばれる．

複数の確率変数に対しても，確率分布を定めることができる．複数の確率変数，X,Y について，それぞれの終域が $Z\subset\mathbb{R}$ であれば，2 変数の分布関数として，

$$P(X(\omega)\le a\wedge Y(\omega)\le b)=\int_{-\infty}^{a}\int_{-\infty}^{b}g(x,y)\,\mathrm{d}x\,\mathrm{d}y$$

と書けるのであれば，g は X,Y の**同時分布**（同時密度関数）と呼ばれる．

また，以下のように定義される関数を**周辺分布**と呼ぶ．

$$P(X(\omega)\le a)=\int_{-\infty}^{a}\left[\int_{-\infty}^{\infty}g(x,y)\,\mathrm{d}y\right]\mathrm{d}x$$

$$P(Y(\omega)\le b)=\int_{-\infty}^{b}\left[\int_{-\infty}^{\infty}g(x,y)\,\mathrm{d}x\right]\mathrm{d}y$$

確率密度関数，g が $g(x,y)=g_X(x)g_Y(y)$ というように二つの関数，g_X,g_Y の積で表現できるとき，確率変数 X,Y は**独立**であるという．独立でなければ**相関**があるという．つまり，例えば Y の実現する値次第で X の実現する確率が変わってくるということである．

同時分布は，ω を使って以下のように書くこともできる．

$$P(X(\omega)\le a\wedge Y(\omega)\le b)=\int_{\{\omega'\in\Omega\mid X(\omega')\le a\wedge Y(\omega')\le b\}}p(\omega)\,\mathrm{d}\omega$$

同様に期待値は次のようにも書くことができる．

$$E[X] = \int_{\Omega} X(\omega)p(\omega)\,\mathrm{d}\omega$$

課題 7.

1. $A = \{1, 2, ..., n\}$ であるとき，$P(A) = \sum_{\omega \in A} P(\omega)$ と書けることを示しなさい．わからなければ $A = \{1, 2\}$ のケースで考えなさい．
2. 三つの確率変数，X, Y, Z について，そのうちのどの二つの組み合わせが独立でも，三つ同時の分布を考えると相関があることがある．このことを以下の例を考えることで示しなさい．
   - 二つのコインの出目の組み合わせが Ω である．X は 1 枚目のコインが表なら 1, 裏なら 0 をとる．Y は 2 枚目のコインが表なら 1, 裏なら 0 をとる．Z は表が出たコインの枚数がちょうど一枚であるとき，1 をとり，それ以外なら 0 をとる．
3. 以下のように定義する．

$$\mathrm{Var}[X] = \int_{\Omega} (X(\omega) - E[X])^2 p(\omega)\,\mathrm{d}\omega$$

$$\mathrm{Cov}[X, Y] = \int_{\Omega} (X(\omega) - E[X])(Y(\omega) - E[Y])p(\omega)\,\mathrm{d}\omega$$

   このとき，以下の事実を示せ．
   - $\mathrm{Var}[X] = \int_{-\infty}^{\infty} (X(\omega))^2 p(\omega)\,\mathrm{d}\omega - (E[X])^2$
   - $E[aX + bY] = aE[X] + bE[Y]$
   - $\mathrm{Var}[aX + bY] = a^2\,\mathrm{Var}[X] + 2ab\,\mathrm{Cov}[X, Y] + b^2\,\mathrm{Var}[Y]$
4. X と Y が独立ならば，$\mathrm{Cov}[X, Y] = 0$ であることを示せ．
   (ヒント) 独立なら $g(x, y) = g_X(x)g_Y(y)$ とかけ，さらに共分散が以下のように書けることを利用する．

$$\mathrm{Cov}[X, Y] = \int\int (x - E[X])(y - E[Y])g(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

4.2.2. 条件付き確率

$A \subset \Omega$ が観察された場合,「どの状態が起こりやすいか」についての予想がアップデートされるのが自然である. このような確率を**条件付き確率**といい, $P(\omega \mid A)$ は，A が起きたとわかったときの ω が出る確率であることを表す．

例えば，次の知識構造を考えよう．

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

このとき，$\omega = 1$（2 でも 3 でも同じ）が実際の実現値であるならば，$\{1,2,3\}$ が起きたことは知っているということになる．この場合，$A = \{1,2,3\}$ が起きたとわかるのである．そうすると，$1,2,3$ に対する予想や，$4,5,6$ に対する予想は $\omega = 1$ が起きる前と後では当然変わってくる．$A \subset \Omega$ が確定している以上，$\Omega \setminus A$ に属する状態が起きる確率は 0 である．他方，確率すべてを足すと 1 であることを維持しなければならないので，$A \subset \Omega$ についてもそのままというわけにはいかない．このとき，次のようなルールで情報を更新する方法を**ベイズ更新 (Bayesian update)** と呼ばれる．

$$P(\omega \mid A) = \begin{cases} \dfrac{P(\omega)}{P(A)} & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

これは，A が起きれば，起こりうるすべての可能性は A の中にあるので，A が全体になり，そのときの確率を $P(A \mid A) = 1$ としたい事情から来ている．

確率変数，X, Y について，もし一方の確率変数の値がわかっており，X, Y の間に相関があるのであれば，例えば次のような条件付き確率を考えることができる．

$$P(X \leq a \mid Y = y) = \frac{\int_{-\infty}^{a} g(x,y)\,\mathrm{d}x}{\int_{-\infty}^{\infty} g(x',y)\,\mathrm{d}x'}$$

よって密度関数は次のように書ける．

$$g(x \mid y) = \frac{g(x,y)}{\int_{-\infty}^{\infty} g(x',y)\,\mathrm{d}x'}$$

条件つき確率を使って期待値を考えることもできる．例えば $E[X \mid y]$ は $Y(\omega) = y$ がわかったときの X の期待値であり，これは次のように計算できる．

$$E[X \mid y] = \int_{-\infty}^{\infty} x g(x \mid y)\,\mathrm{d}x$$

課題 8.

1. 次のようなものを繰り返し期待値と呼ぶ．

$$E[E[X \mid Y]] = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} x g(x \mid y)\,\mathrm{d}x \right] g_Y(y)\,\mathrm{d}y$$

このとき, $E[E[X \mid Y]] = E[X] = \int_{-\infty}^{\infty} x g_X(x)\,\mathrm{d}x$ であることを**繰り返し期待値の法則**という．この事実を証明しなさい．ただし g_X, g_Y は周辺分布の密度関数であり，次のように定義される．

$$g_Y(y) = \int_{-\infty}^{\infty} g(x,y)\,\mathrm{d}x$$

$$g_X(x) = \int_{-\infty}^{\infty} g(x,y)\,\mathrm{d}y$$

4.2.3. 和の分布と畳み込み

二つの確率変数の和，$X+Y$ の分布をどのように求めるかを考える．これは，次のように計算できる．

$$P(X+Y \le a) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{a-x} g(x,y)\,\mathrm{d}y \right] \mathrm{d}x$$

これは，$X+Y \le a$ が，$Y \le a-X$ と変形できることによる．a で微分することにより，密度関数も次のように求められる．

$$\int_{-\infty}^{\infty} g(x, a-x)\,\mathrm{d}x$$

このような積分を**畳み込み積分**と呼ぶ．

畳み込み積分を用いた応用例として，ノイズのついた情報から真の情報を推測する方法を考えよう．観測されるデータが，$Y = X + N$ と言う形で表されるとする．つまりは X, N は直接観察することはできないが，その和である Y は観察できると言う状況である．欲しい情報は X だが，そこにノイズ N がかかっているときにはこの状態に近い．このときに観察されたデータ $y = Y$ の情報から X の分布を推測したいという状況がよくある．例えば，学生の真の実力 X を測りたいときに学力テストをするわけだが，学力テストの点数には，対策がうまく行った，体調が悪かった，などの実力以外の要素が反映されてしまう．この部分がノイズというわけだ．こういったとき，X の分布をどのように計算するのか見てみよう．

ここで，N と X は独立であるとしてみよう．つまり同時分布は $g(x,\varepsilon) = g_X(x)g_N(\varepsilon)$ と書ける．ちなみにノイズは ε と書くことが多いのでそう書いている．今，$Y = y$ と言う情報が観測されたとする．このとき，x の分布を推

— Notes —

Date:

Notes

Date:

測したい．この計算には条件付き確率と和の分布の知識を使う．まず，$Y = y$ と言う情報が観測される確率を考えよう．これは，$\int_{-\infty}^{\infty} g_X(x) g_N(y-x)\,\mathrm{d}x$ と書ける．このときの X の分布は，次のように計算できる．

$$P(X \leq a \mid Y = y) = \frac{\int_{-\infty}^{a} g_X(x) g_N(y-x)\,\mathrm{d}x}{\int_{-\infty}^{\infty} g_X(x) g_N(y-x)\,\mathrm{d}x}$$

分布を特定して，具体的に計算してみよう．g_X が平均 μ, 分散 σ^2 の正規分布の密度関数であるとは，$g_X(x) = \frac{1}{2\sqrt{\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ と書けることをいう．このときは確率変数 X は平均 μ, 分散 σ^2 の正規分布に従うという．X が平均 μ, 分散 σ^2 の正規分布に従い，N が平均 0, 分散 γ^2 の正規分布に従うとき，$P(X \leq a \mid Y = y)$ として求められる分布が平均 $\frac{\gamma^2\mu + \sigma^2 y}{\gamma^2 + \sigma^2}$, 分散 $\frac{\gamma^2\sigma^2}{\gamma^2+\sigma^2}$ の正規分布であることを示す．

このためには $g_X(x) g_N(y-x)$ の計算をすれば良い．正規分布の $\frac{1}{2\sqrt{\pi}\sigma}$ は積分して 1 になるように調整するためのものであるからあまり重要ではない[3]．全体を $\left(\int_{-\infty}^{\infty} g_X(x) g_N(y-x)\,\mathrm{d}x\right)$ で割っているので，$P(X \leq \infty \mid Y = y)$ は必ず 1 になるのである．したがって，$g_X(x) g_N(y-x)$ の計算結果が $e^{-\frac{(x-\text{平均})^2}{2\text{分散}}}$ という形をしていることを確認すれば十分である．実際に計算してみよう．

$$\begin{aligned} & e^{-\frac{(x-\mu)^2}{2\sigma^2}} e^{-\frac{(y-x)^2}{2\gamma^2}} \\ &= e^{-\frac{(x-\mu)^2}{2\sigma^2} - \frac{(y-x)^2}{2\gamma^2}} \\ &= e^{-\frac{1}{2\sigma^2}(x^2 - 2\mu x + \mu^2) - \frac{1}{2\gamma^2}(y^2 - 2xy + x^2)} \end{aligned}$$

以降は，指数の部分の計算がメインになるので，そこだけ取り出して計算していこう．さらに，x に関する項だけまとめることを考えよう．これは x に関する積分なので，そこ以外は外に括り出せるからである．

[3] $\int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \sqrt{2\pi}\sigma$ になることが実際知られている．これはガウス積分と変数変換を使えば求められる．

$$
\begin{aligned}
&-\frac{1}{2\sigma^2}\left(x^2-2\mu x+\mu^2\right)-\frac{1}{2\gamma^2}\left(y^2-2xy+x^2\right)\\
&=-\frac{\sigma^2+\gamma^2}{2\gamma^2\sigma^2}\left(x^2-2x\frac{\gamma^2\mu+\sigma^2 y}{\sigma^2+\gamma^2}\right)+\text{定数あ}\\
&=-\frac{\sigma^2+\gamma^2}{2\gamma^2\sigma^2}\left(x-\frac{\gamma^2\mu+\sigma^2 y}{\sigma^2+\gamma^2}\right)^2+\text{定数い}\\
&=-\frac{\left(x-\frac{\gamma^2\mu+\sigma^2 y}{\sigma^2+\gamma^2}\right)^2}{2\frac{\gamma^2\sigma^2}{\sigma^2+\gamma^2}}+\text{定数い}
\end{aligned}
$$

2 番目から 3 番目の変形には平方完成を使っている．そうすると $Z=\frac{\gamma^2\mu+\sigma^2 y}{\sigma^2+\gamma^2}$, $V=\frac{\gamma^2\sigma^2}{\sigma^2+\gamma^2}$ とすれば以下のようにまとめられる．

$$
g_X(x)g_N(y-x)=\text{定数}e^{-\frac{(x-Z)^2}{2V}}e^{\text{定数い}}
$$

よって，

$$
\begin{aligned}
\frac{\int_{-\infty}^{a}g_X(x)g_N(y-x)\,\mathrm{d}x}{\int_{-\infty}^{\infty}g_X(x)g_N(y-x)\,\mathrm{d}x}
&=\frac{\int_{-\infty}^{a}e^{-\frac{(x-Z)^2}{2V}}e^{\text{定数い}}\,\mathrm{d}x}{\int_{-\infty}^{\infty}e^{-\frac{(x-Z)^2}{2V}}e^{\text{定数い}}\,\mathrm{d}x}\\
&=\frac{\int_{-\infty}^{a}e^{-\frac{(x-Z)^2}{2V}}\,\mathrm{d}x}{\int_{-\infty}^{\infty}e^{-\frac{(x-Z)^2}{2V}}\,\mathrm{d}x}\\
&=\int_{-\infty}^{a}\frac{1}{\sqrt{2\pi V}}e^{-\frac{(x-Z)^2}{2V}}\,\mathrm{d}x
\end{aligned}
$$

これで，正規分布の密度関数が現れることを確認できたのである．

課題 9.

次のうち，一つの問いに答えなさい．

1. X, Y が $[a, b]$ 上の確率変数であるとする．ただし，$b>a>0$ とする．このとき，$\frac{X}{Y}$ の分布関数を求めなさい．

2. X, Y が $[a, b]$ 上の確率変数であり，X, Y は独立であるとする．ただし，$b > a$ とする．このとき，$\max\{X, Y\}$ の分布関数を求めなさい．
   - ヒント: $\max\{X, Y\} = z$ としてみよう．$X(\omega) > Y(\omega)$ なら $X(\omega) = z$ である．$X > Y$ となるケースだけを考えれば，$X(\omega) = z$ のとき，Y は $[a, z]$ 上にあるはずである．つまりその確率は $G_Y(z)$ である．あとは，$X \leq z$ になる確率を考えれば答えが出る．$X(\omega) < Y(\omega)$ となるケースも同様に考えて，その二つを足せば良い．

4.3. 大数の弱法則

確率変数が平均からどれだけ離れるか，というのはバラツキを考える上で大事である．ほとんど平均から離れることがないのであれば，ほとんど確実に起こると言えるし，そうでなければかなり不確実性が高いと言える．

これがどれだけ離れるかというのは分散を使って評価できる．実際，次の不等式が成り立つ．

チェビシェフの不等式

X の平均を μ，分散を σ^2 とする．このとき，任意の実数 $k > 0$ について，次の不等式が成り立つ．

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

(証明) χ という関数を次のように定める．

$$\chi(\omega) = \begin{cases} 1 \text{ if } |X - \mu| \geq k\sigma \\ 0 \text{ otherwise} \end{cases}$$

すると，$P(|X - \mu| \geq k\sigma) = E[\chi]$ と書くことができる．

このとき，$|X - \mu| \geq k\sigma$ というのは $\frac{(|X-\mu|)^2}{\sigma^2}\frac{1}{k^2} \geq 1$ であることに注意する．また，$\frac{(|X-\mu|)^2}{\sigma^2}\frac{1}{k^2} \geq 0$ は必ず成り立つ．ということは，$\chi(\omega) = 1$ のとき，$\frac{|X(\omega)-\mu|^2}{\sigma^2}\frac{1}{k^2} \geq 1$ となるわけなので，$\frac{|X(\omega)-\mu|^2}{\sigma^2}\frac{1}{k^2} \geq \chi(\omega)$ であることが言える．ここから，次のように計算できる．

$$\begin{aligned} P\left(|X-\mu|^2 \geq k\sigma\right) &= E[\chi] \\ &\leq E\left[\frac{|X-\mu|^2}{\sigma^2}\frac{1}{k^2}\right] \\ &= \frac{1}{k^2}\frac{E\left[|X-\mu|^2\right]}{\sigma^2} = \frac{1}{k^2} \end{aligned}$$

ここで，$E\left[|X-\mu|^2\right] = \sigma^2$ は分散の定義である．

さて，この実現値と期待値の乖離であるが，数多くのサンプルをとるとその平均値と期待値の乖離は次第に縮小していく．このことをチェビシェフの不等式を用いて示そう．$X_1, X_2, \ldots$ を複数の確率変数の列とし，それぞれの i について X_i を平均 μ, 分散 σ^2 に従うとする．また，X_i と X_j は $i \neq j$ なら独立であるとする．その上で，確率変数の平均 $\overline{X}_n = \frac{\sum_{i=1}^n X_i}{n}$ を考える．この分布を考えよう．まず平均は $E\left[\overline{X}_n\right] = \mu$ である．また分散は独立性より $i \neq j$ ならば $\mathrm{Cov}\left[X_i, X_j\right] = 0$ であるので，

$$\mathrm{Var}\left(\overline{X}_n\right) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

となる．そうするとチェビシェフの不等式から，以下の不等式が言える．

$$P\left(\left|\overline{X}_n - \mu\right| \geq k\frac{\sigma}{\sqrt{n}}\right) \leq \frac{1}{k^2}$$

ここで，k は正の実数ならなんでもいいので，$n^{\frac{1}{4}}$ としよう．そうすると，

$$P\left(\left|\overline{X}_n - \mu\right| \geq \frac{\sigma}{n^{\frac{1}{4}}}\right) \leq \frac{1}{n^{\frac{1}{2}}}$$

となる．ここで，$n \to \infty$ という極限をとってやれば，$P\left(\left|\overline{X}_n - \mu\right| \geq 0\right) \leq 0$ ということになる．これは期待値からの乖離が 0 より大きい確率が 0 になるということである．つまり極限では期待値と確率変数の平均は一致することになる．この性質を**大数の弱法則**という．

— Notes — Date:

大数の弱法則の応用として，**コンドルセの陪審定理**を紹介する．次のような状況を考えよう．$x \in \{0,1\}$ とする．「陪審」定理の名に従って，これは被告が真犯人であるかどうかを示すものとしよう．$x = 0$ ならこの被告は無実であり，各陪審員は確率 $q \in \left(\frac{1}{2}, 1\right)$ で無罪判決を出す．逆に $x = 1$ ならこの被告は真犯人であり，各陪審員は確率 $q \in \left(\frac{1}{2}, 1\right)$ で有罪判決を出す．陪審の結審は陪審員たちの多数決で行われる．このとき，多数決が正しい審判を下す確率はどれほどだろうか？

もし，$x = 0$ なら無罪の票は q だけの割合で入る．確率変数 χ_i を陪審員 i が無罪の票を投じれば 1, 有罪票を投じれば 0 という確率変数とするとその期待値は q である．陪審員の数が無限大に多く，なおかつ彼らの判断が独立であれば，大数の弱法則から χ_i の平均, $\frac{\sum \chi_i}{n}$ は q に収束し，その確率は 1 である．χ_i の平均, $\frac{\sum \chi_i}{n}$ というのは実際に無罪票に投じた陪審員の割合だ．$q > \frac{1}{2}$ であるので，無罪票に投じた陪審員の割合は過半数を確率 1 で超えることになる．つまり，無実なら，無罪の確率が 1 という結論になる．

被告が真犯人である時にも同様の結論が導かれる．このように，陪審員が無限にいて，それぞれの判断が独立であれば，判決は確率 1 で「正しい」というのがコンドルセの陪審定理が主張することである．

課題 10.

1. 大数の弱法則はそれぞれの確率変数の分散や期待値が異なっても成り立つ．$X_1, X_2, \ldots$ を複数の確率変数の列とし，それぞれの i について X_i を平均 μ_i, 分散 σ_i^2 に従うとする．ただし，$\frac{\sum_i \mu_i}{n} \to \mu$ であり，$\lim_{n\to\infty} \frac{\max_{k\le n} \sigma_k}{n^{\frac{1}{2}}} = 0$ とする．このとき，確率変数の平均 $\overline{X}_n = \frac{\sum_{i=1}^n X_i}{n}$ がどのような値に収束するか，答えなさい．また，その乖離が 0 に収束することを証明しなさい．

5. ゲーム理論への応用

5.1. ゲーム理論について

自分が何か行動するとき，自分自身の選択肢や好みは当然重要であるが，他者が関わってくるときは他者がどう行動するかを頭に入れなくてはならない．例えば，経済学部でよく例として挙げられる同質財の寡占市場では，自分がどれだけ生産すべきかどうかは，相手がどれだけ生産してくるかに依存してくる．相手がたくさん生産すればそれだけ財価格が低下し，自分は多く生産すべきではない．反対に相手があまり生産しないのであれば，価格はそれほど低下しないので，逆に自分はたくさん生産すべきである．

このように，自分がどれだけ生産すべきか，というのは相手がどれだけ生産してくるかという「予想」に依存する．この予想と実際の行動が一致するとき，ゲーム理論では一般的にその状態を「均衡」と呼ぶ．その詳細によって均衡にはいろいろな定義があるが，その精神は大体同じである．つまり，均衡を考えるためには，人が他者の行動をどう予想するか，ということにかかっている．

最も単純なパターンとして次の例を考えよう．

		相手	
		L	R
自分	U	$1, 2$	$2, 3$
	D	$3, 1$	$1, 5$

この状況を「ゲーム」と呼ぶ．「自分」の選択肢は U か D かであり，相手の選択肢は L か R かである．例えば，自分が U, 相手が L を選んだときには左上のマスが実現し，「自分」には 1 という利得を，相手には 2 という利得を与える．ゲームではこの利得を最大にすることが目的である．

その観点からすると，相手が L を選ぶと予想するなら自分は U を選べば利得 1，D を選べば利得 3 を得られるので，D を選ぶべきとなる．逆に R を選

— Notes —

Date:

— Notes —

Date:

んでくると予想すれば，自分は U を選ぶべきである．このように自分の選ぶべき行動が相手が何をしてくるかという予想に依存してくる．

では相手の行動をどう予想すべきか？これを考えるために相手の立場に立って考えてみよう．相手が自分が U を選んでくると予想するなら，相手は R を選ぶべきである．相手が自分が D を選ぶと予想するなら，相手は R を選ぶべきである．ということは自分が何を選んでも相手は R を選ぶべきである．このような選択肢のことは **(強) 支配戦略**と呼ばれる．そうすると，相手は R を選んでくるという予想がたつ．そうすれば自分は何をすべきかというのは非常に簡単で，これは U を選ぶべきだということになる．

ではそのような強支配戦略がないようなゲームだとどうだろうか？次のようなゲームはそのような状況である．

		相手	
		L	R
自分	U	2, 2	0, 1
	D	1, 0	1, 1

このゲームを考えると, , 相手が L を選ぶと予想するなら自分は U を選べば利得 2，D を選べば利得 1 を得られるので，D を選ぶべきとなる．逆に R を選んでくると予想すれば，自分は U を選ぶべきである．この点は前のゲームと変わり映えはしない．しかし，相手の立場では違いが出てくる．自分が U を選ぶと予想するなら，相手は L を選ぶべきである．逆に D を選ぶと予想するなら，相手は R を選ぶべきである．このような状況では，自分が何を選ぶかというのは相手が何を選ぶかという予想に依存している．逆に相手が何を選ぶかというのは自分が何を選ぶかという予想に依存している．では，この場合はお互いにどういう選択肢を選ぶべきだろうか？

ポイントは「予想」と「（最適な）行動」の二つが一致するということである．ここで，自分と相手がお互いに (U, L) を取ると予想したとしてみよう．そうすると，自分は U を選ぶべきであるし，逆に相手も L を選ぶべきである．

この状況では彼らの予想とそれに対する最適な行動が一致している．この状況を**ナッシュ均衡**と呼ぶ．

ナッシュ均衡は一つとは限らない．例えば同じゲームで，(D, R) を取ると予想したとしてみよう．このとき，自分は D を選ぶべきであるし，逆に相手も R を選ぶべきである．この状況でも彼らの予想とそれに対する最適な行動が一致している．

これらのゲームでのポイントは，行動と利得に関する情報がお互いに完全にわかっているということである．言い換えれば，ゲームの情報が共有知識であるということだ．このゲームの情報の一部が共有知識になっていないとき，「予想」をどう考えるかが重要になる．これを次の節で見ていこう．

5.2. 高階の信念

5.2.1. 電子メールゲーム

次のようなゲームを見てみよう．状態 X が起きる確率は $\frac{2}{3}$ であるとする．

状態 X

A 氏 \ B 氏	L	R
U	$1, 1$	$0.1, -2$
D	$-2, 0.1$	$0, 0$

状態 Y

A 氏 \ B 氏	L	R
U	$0, 0$	$0.1, -2$
D	$-2, 0.1$	$1, 1$

状態 X であるとわかりきっていれば，ナッシュ均衡は (U, L) である．状態 Y であるとわかりきっていればナッシュ均衡は (U, L), (D, R) の二つである．では，状態が X か Y かわかりきっていないとき，どのような行動を取ればいいだろうか？今，A 氏だけが状態が X か Y かを知ったとし，これを B 氏に伝えることを考える．具体的には次のような手続きで B 氏に状態を知らせることを考えよう．

- 状態 Y のときのみ，A 氏が相手にメールを自動で送信する

- B 氏が A 氏からのメールを受信すると，自分にメールを自動で送信する
- A 氏が B 氏からのメールを受信すると，相手にメールを自動で送信する
- 以下, 交互にメールを送り合う. メールが一度でも届かないとそこでコミュニケーションは終了する.
- メールは確率 $1-\varepsilon$ で届く. ε で届かない.
  - $\varepsilon \in (0,1)$ とする.

このコミュニケーションの後，我々はどのような行動を取るべきであろうか？ポイントは，A 氏が「B 氏がどちらの状態にいるのか知っていること」を必ずしも知らないということである.

まず，A 氏が状態が X であると知ったとしよう．そうすると，A 氏は U を取るべきである．これは U が支配戦略であるからだ．一方で, Y であると知ったときには B 氏の行動に依存してくる．そうすると，B 氏がどのような行動をしてくるかという予想を考えなくてはならない．従って B 氏の立場に立って考えよう.

B 氏の立場に立って考えるとき，メールが何回送られてきたかは重要である．なぜなら B 氏の状態に関する情報は自分がどれだけメールを送ったかに依存するからである．最初にメールが全く送られてこなかったケースを考えよう．そうすると，次の二つの可能性がある.

1. A 氏は X であると知ってメールを送信しなかった
2. A 氏は Y であると知ってメールを送信したが，B 氏に届かなかった

Y であると知って，送信ミスをした確率は，$\frac{1}{3}\varepsilon$ である．一方で，X であると知ってメールを送信しなかった確率は $\frac{2}{3}$ である．そうすると，B 氏の立場に立って，メールが送られてこなかったという条件のもと，X である確率はベイズ更新の公式により $\frac{\frac{2}{3}}{\frac{2}{3}+\frac{\varepsilon}{3}} = \frac{2}{2+\varepsilon} > \frac{2}{3}$ である.

「状態が X であるならば A 氏は U を選んでいる」と B 氏は予想する．なぜならそれが支配戦略であるからだ．したがって，B 氏が L を選ぶときの利得の期待値は少なくとも次の値以上になっているはずである．

$$\frac{2}{2+\varepsilon}\times 1>\frac{2}{3}$$

これは状態が Y だったとしても，B 氏が何を選んでいようと利得が 0 以上になるからである．一方で，R を選ぶとき，利得の期待値は最大でも次の値より大きくない．

$$\frac{2}{2+\varepsilon}\times(-2)+\frac{\varepsilon}{2+\varepsilon}\times 1=\frac{-4+\varepsilon}{2+\varepsilon}<0$$

よって，メールが全く送られてこなければ，B 氏は L を選ぶべきである．

次に，A 氏は Y であると知ってメールを送ったが，返信が一度も来なかったケースを考えてみる．このとき，次の可能性が考えられる．

1. A 氏はメールを送信したが，B 氏に届かなかった（確率 ε) で発生）
2. B 氏はメールを受信し，返信したが，A 氏に届かなかった（確率 $(1-\varepsilon)\varepsilon$ で発生）．

前者のケースを考えよう．条件付き確率を考えるとこれが発生した確率は $\frac{\varepsilon}{(1-\varepsilon)\varepsilon+\varepsilon}=\frac{1}{2-\varepsilon}>\frac{1}{2}$ である．B 氏に届いてなかった場合，B 氏は L を選んでくるはずである．つまり，B 氏は $\frac{1}{2-\varepsilon}$ 以上の確率で L を選んでくると予想できる．このとき，U を選ぶときの利得の期待値は 0 以上になるはずである．

一方で，D を選ぶ場合，最大でも次の値より大きくない

$$\frac{1}{2-\varepsilon}\times(-2)+\frac{1-\varepsilon}{2-\varepsilon}\times 1=\frac{-1-\varepsilon}{2-\varepsilon}<0.$$

よって，メールの返信が一度も来なかった場合，A 氏は U を選ぶべきである．このように A 氏は Y であることを知っていても D ではなく，U を選んでしまうのである．

このプロセスを繰り返して考える．B 氏は A 氏からのメールが来て，返信をしたが，2 回目の返信が来なかったとしてみよう．そうすると，次の可能性が考えられる．

— Notes —

Date:

1. 返信したが届かなかった (確率 ε で発生，A 氏は U を取る)
2. 返信して届いたが，A 氏が送信ミスをした (確率 $(1-\varepsilon)\varepsilon$ で発生)

そうすると，さっきと同じような考え方によって，A 氏は確率 $\frac{1}{2}$ 以上で U をとってくると相手は確信することになる．そうすれば, L を取ることがベストになる．同様にして，メールが届かなかった時点で，情報を更新して何を選ぶべきかを考えると，A 氏は U を選び，B 氏は L を選ぶべきとなるのである．そうすると, A 氏も B 氏も状態が Y であることがわかっているのに，(U, L) しか均衡になれないのである．

これを知識構造の観点から考えてみよう．A 氏にメールが届いた回数を T_A とする．同様に T_B を B 氏にメールが届いた回数とする．ここで世界の状態として (状態，A 氏に届いた回数，B 氏に届いた回数) をペアとして考える．記号では (s, T_A, T_B) と書く. メールが一度も届かないとき，B 氏にとっては $(X, 0, 0)$ と $(Y, 0, 0)$ の区別がつかない．また, A 氏に返信が一度も来ないとき，$(Y, 0, 0)$ と $(Y, 0, 1)$ の区別がつかない．このとき，A 氏と B 氏の知識構造は次のようになる．

$$\mathscr{F}_{\text{A氏}} = \{\{(X,0,0)\},\{(Y,0,0),\quad(Y,0,1)\},\{(Y,1,1),\quad(Y,1,2)\},\dots\}$$

$$\mathscr{F}_{\text{B氏}} = \{\{(X,0,0),\quad(Y,0,0)\},\{(Y,0,1),\quad(Y,1,1)\},\{(Y,1,2),\dots\}$$

こうすると，「Y が発生したこと」が共有知識になっていないことがわかる．しかしながら，メールが何千何万回も飛び交った状態というのはほとんど共有知識みたいなもののはずだ．Y が発生したことはお互いにほとんど共有知識みたいなものになっているのにも関わらず，Y が共有知識のときのナッシュ均衡 (D, R) は均衡にはならないのである．この意味でゲームの構造が共有知識であることは，ナッシュ均衡には重要である．

5.2.2. 同意しないことに同意することの不可能性

共有知識というものは，相手がどれだけの確率で事象が起きるのかという予想（これを**信念 (belief)** と呼ぶ）についても定義することができる．この信念が一致するかどうかは重要である．次の例を考えてみよう．

- A と B は決闘をするかどうかを考えている.
- 決闘に敗れた方の利得は -2 である.
- 決闘に勝利した方の利得は 1 である.
- 申し込むことのコストは 0.1 である.
- お互いが決闘をすることを申し込めば, 決闘が起きるが, 片方が申し込まなければ決闘は起きない. このときの利得は両者 0 である.
- A が勝利する確率を p と表記する.

お互いに相手が決闘を申し込むという前提のもとではAが決闘に申し込むためには，少なくとも $p > \frac{2}{3}$ でなければならないし，B が決闘に申し込むためには多くとも $p < \frac{1}{3}$ でなければならない（確かめてみよ）. したがって,「A が勝利する確率」に関する予想が一致しているのであれば決闘は起きない. 逆に言えば，決闘が起きるためには，お互いに自分が勝利する可能性の方が高いと思っている必要があるのである.

ではこのようなことが起きる，つまり信念の不一致が共有知識になることはあるだろうか？ これを考えるため，次のように考えてみよう.

- Ω を状態の集合とする.
- p を事前共通信念と呼ぶ. これは Ω 上の確率関数である.
- $\mathscr{F}_i$ を $i \in \{A, B\}$ の知識構造とする.
- $\omega \in \Omega$ が実現したとき，$\omega \in F \in \mathscr{F}_i$ とすると，i は情報を次のようにアップデートする.

$$p_i(\omega) = p(\omega \mid F) = \frac{p(\omega)}{p(F)}$$

例えば，知識構造を次のように考える.

$$\mathscr{F}_A = \{\{1,2\},\{3,4\},\{5,6,7\}\}$$

$$\mathscr{F}_B = \{\{1,2,3\},\{4,5\},\{6,7\}\}$$

— Notes —

Date:

そして, $p(3) = 0.05, p(4) = 0.2, p(5) = 0.5$ としてみよう. そうすると, $\omega = 4$ であった場合，A は $\omega = 4$ である確率を

$$p_A(4) = p(4 \mid \{3,4\}) = \frac{p(4)}{p(\{3,4\})} = \frac{0.2}{0.05 + 0.2} = \frac{2}{2.5} = \frac{4}{5} > \frac{2}{3}$$

と考える．これに対して，B は $\omega = 4$ である確率を

$$p_B(4) = p(4 \mid \{4,5\}) = \frac{p(4)}{p(\{4,5\})} = \frac{0.2}{0.2 + 0.5} = \frac{2}{7} < \frac{1}{3}$$

と考える．こうすると，確かに $\omega = 4$ である確率は確かに A と B で異なる．問題は，このことが共有知識になっているかということである．

実はこのことは共有知識ではない．A の視点に立って考えよう．A は $\omega = 3$ か $\omega = 4$ かの区別がつかない．もし，$\omega = 3$ であれば，B は $\omega = 4$ である確率は 0 だと思っているはずである．したがって，A は B の信念について 0 なのか $\frac{1}{3}$ なのかの区別がつかないので, $p_B(4) = \frac{1}{3}$ であることは共有知識ではない．

このことが「決闘」の参加の意思決定にどう影響を及ぼすのかを見てみよう．仮に $\omega = 4$ であれば A が勝利，それ以外は B が勝利するとしてみよう．そうすると $\omega = 4$ がおきれば，A も B も，相手が決闘に申し込むなら自分も申し込みたいと思うはずである．

A が決闘を申し込むとすれば $\omega \in \{3,4\}$ のケース以外にあり得ないことに注意せよ．そうでなければ $\omega \neq 4$ がわかってしまうので，申し込むのは得にならない. ここで B の選択を考えてみよう. $\omega = 4$ なら, B は $\omega = 4, 5$ のどちらかであることがわかっている．場合分けをして考えてみよう．もし $\omega = 4$ なら B は申し込むと，相手が申し込んでいれば確実に敗北するし，申し込んでいなければ 0 の利得を得るのでどちらにしても申し込まない方が良い. $\omega = 5$ なら，A は申し込んでこないので，確実に利得は 0 以下である．したがって, B は $\omega = 4$ のときには申し込むことはないのである. そして, そのことがわかっていれば，A も同じように申し込まない．1 階の信念だけ考えて，

決闘を申し込むことがありえたとしても 2 階の信念をきちんと考えれば両者が共に申し込むことが均衡になることがないということがこれでわかる．

それでは一般的に，$p_A(\omega) = q_A \neq q_B = p_B(\omega)$ が共有知識になりえないかということが疑問になる．次の結果はその答えを示す．

> ω が発生したとき，$p_A(\omega) = q_A$ と $p_B(\omega) = q_B$ が共有知識だとする．このとき，$q_A = q_B$ である．
>
> ---
>
> **(証明)** $p_A(\omega) = q_A$ が共有知識であるとするとき，そのような情報集合を $\omega \in F \in \mathscr{F}_A$ とする．このとき，$F \in \mathscr{F}_B$ である．もし，$F \notin \mathscr{F}_B$ であれば，$\omega \in F' \in \mathscr{F}_B$ かつ $F \neq F'$ という F' が存在することになる．そうすると，$\omega' \in F' \setminus F$ という状態が存在することになる．しかし ω' がおきれば，A にとって，$p_A(\omega) = 0$ であることがわかる．したがって，F' において，$p_A(\omega) = q_A$ であることは B は知り得ない．つまり，$p_A(\omega) = q_A$ は共有知識になっていない．したがって，
>
> $$p_A(\omega) = \frac{p(\omega)}{p(F)}$$
>
> である．全く同じ論理により，$p_B(\omega) = \frac{p(\omega)}{p(F)}$ である．よって，$p_A(\omega) = p_B(\omega)$ である．

これはオーマン (R. Aumann) の**同意定理**と呼ばれるものの簡易バージョンである．一般的にはどんな集合 E についても「ω が発生したとき，$p_A(E) = q_A$ と $p_B(E) = q_B$ が共有知識だとする．このとき，$q_A = q_B$ である．」ということが言える．

課題 11.

1. 電子メールゲームにおいて，共有知識を求めなさい．

2.

$$\mathscr{F}_A = \{\{1,2\},\{3\},\{4\}\}$$
$$\mathscr{F}_B = \{\{1\},\{2,3\},\{4\}\}$$

とする．ただし，$p(1) = 0.1, p(2) = 0.3, p(3) = 0.4, p(4) = 0.2$ とする．$\tilde{p}_A(\omega)$ を ω が発生したとき，A が ω に対して割り当てる確率とする．このとき，$\{\omega \in \{1,2,3\} \mid \tilde{p}_A(\omega) \neq \tilde{p}_B(\omega)\}$ が共有知識になり得るかどうかを考えなさい．

3. 2 の結果はオーマンの同意定理に反するか？反しないか？理由を述べよ．

5.3. オークション理論

ゲーム理論を現実に応用する例として活発なものが，オークションへの応用である．いろいろな定式があるが，以下が最も単純なものである．

- 入札者が n 人いる．
- 各入札者は一つの財を手に入れたい. 財は一つしかなく, 誰かが手に入れれば他の誰も手に入れることができない.
- i という入札者にとっての価値（評価値という）は v_i であり，これは $[0,1]$ 上の分布関数 F に従う確率変数である
- $i \neq j$ であれば v_i と v_j は独立である
- お互いに自分以外の評価値は知らない.

オークションで，これらの入札者に売るとき，色々な方法が知られているが代表的なものは次のとおりである．

- **封印入札オークション**
  - 入札額を書き，それを封筒に入れ主催者に渡す
  - 落札者は入札額の最も高い者である．
- **公開入札オークション**
  - 一般的なオークションのイメージに近い

さらに封印入札や公開入札の中にも，どのように支払いを決めるかによってルールが分かれる．この講義ノートでは封印入札に絞って考えてみよう

5.3.1. 第１価格オークションと第２価格オークション

第１価格オークションとは，封印入札オークションで落札額が落札者の入札額になっているオークションのことである．例えば，入札者が 5 人いて，入札額が，$b_1 = 0.5, b_2 = 0.2, b_3 = 0.7, b_4 = 0.3, b_5 = 0.3$ となっていれば，落札者は 3 で，落札額は 0.7 である．

これに対して，**第２価格オークション**とは，落札額が，2 番目に高い入札額になっているオークションである．さっきの例で言えば，落札者は 3 で，落札額は 0.5 である．

それではそれぞれのオークションについて，どのように各入札者が入札すれば良いのかを考える．このために，**入札戦略**というものを導入する．これは，評価値を入力とし，入札額を出力とする関数である．これを $b_i(v_i)$ などと書くことにしよう．話を単純にするために，全員の入札戦略が同じであるとしよう．つまり，$b_i = b$ としてみる．

さらに単純化して，$n = 2$ としてみよう．まず第 1 価格オークションから考える．このとき，入札者 1 の入札額が s であれば利得の期待値は次のような計算ができる．

$$P(s > b(v_2))(v_1 - s)$$

つまり，入札者 1 の入札額が 1 位になるとき，落札でき，そのときに利得 v_1 と支払い s が発生しているのである．これをもう少し書き下してみよう．

$$P(s > b(v_2))(v_1 - s) = F\big(b^{-1}(s)\big)(v_1 - s)$$

ここで，b^{-1} は b の逆関数である．これにより，利得が s の関数になったため，これを微分することにより，最適な s を見つけることができる．そうすると，最適な s の条件は以下の通りである[4]．

[4]逆関数の微分は，$b\big(b^{-1}(s)\big) = s$ という式の両辺を微分することによって得られる．

$$f\left(b^{-1}(s)\right)(v_1 - s)\frac{1}{b'\left(b^{-1}(s)\right)} - F\left(b^{-1}(s)\right) = 0$$

ただし，$F' = f$ である．ここで，s というのは v_1 のときの最適な s であるので，$b(v_1) = s$ をみたす．よって上の式は以下のように書き換えられる．

$$f(v_1)(v_1 - b(v_1))\frac{1}{b'(v_1)} - F(v_1) = 0$$

$$\Rightarrow \frac{f(v_1)}{F(v_1)}(v_1 - b(v_1)) = b'(v_1)$$

この方程式は，微分する前の関数 b と微分した後の関数 b' の両方が含まれている．このような方程式を**微分方程式**と呼ぶ．さてこのままでは計算が難しいが，微分方程式の一部はよく知られた解き方がある．ここでは wolfram alpha などの数式処理ソフトに頼ってみると具体的な形を計算できる．ここではそれはややこしいので省略しよう．

さて，通常は評価値が高ければ入札額も高い．数式でかければ $b'(v_1) > 0$ である．このことを飲み込むと $\frac{f(v_1)}{F(v_1)}(v_1 - b(v_1)) = b'(v_1)$ という形をしていることから $v_1 > b(v_1)$ であることがわかる．つまり入札額は必ず評価値を超えない．この理由は簡単である．まず評価値を超えると落札したときの利得がマイナスになるのであり得ない．一歩位で自分の評価値より低く入札しても勝つ可能性があるので，低く入札する結果になる．

これに対して，第 2 価格オークションでの入札戦略を考えよう．相手の入札額が b_2 であるとし，これの分布が G としよう．これは別にどんな形をしていても構わない．そうすると，入札額を s とするときの期待値は次のとおりである．

$$\int_0^s (v_2 - b_2)g(b_2)\,\mathrm{d}b_2$$

この関数を微分すれば，$(v_2 - s)g(s)$ ということがわかる．つまり，$s = v_2$ のときに期待値が最大になる．つまり入札額を評価値通りにするのが最適になるのである．これは第 1 価格オークションとの大きな違いである．第 1 価格オークションでは，自分の入札額がそのまま落札額になるので，どうして

も評価値より低くしてしまう．一方で，第 2 価格オークションでは，落札額は他人の入札額になるので，自分ではコントロールできないのである．そうすると，第 1 価格オークションと比べ，入札額を大きくする意味が出てくるのである．

5.3.2. 収入同値定理

さて，第 1 価格オークションと第 2 価格オークションでは最適な入札戦略は異なっていた．ではどういったオークションを導入すればいいのだろうか？気にするべきことの一つは主催者の収入である．では第 1 第 2 どちらが高いのか比べてみよう．

まず第 2 価格オークションの入札額の期待値は次のとおりである．

$$\int_0^1 \left(\int_0^{v_1} v_2 f(v_2)\,\mathrm{d}v_2\right) f(v_1)\,\mathrm{d}v_1$$

これは第 2 価格オークションではそのまま評価値をそのまま入札することが最適になっていたので, 2 番目の評価値の期待値を計算している．まず, v_1 が一番評価値が高いケースを考え，その分布を考える．そのとき，v_2 は v_1 以下でなければいけないので，その確率は $\int_0^{v_1} f(v_2)\,\mathrm{d}v_2$ である．あとは v_2 を中に入れ, 期待値を計算する. v_2 の方が評価値が高いケースも同じなので, 平均すればこの値が収入の期待値である．

部分積分を使って見やすい形に変形してやると，以下の通りである．

$$\int_0^1 vf(v)(1-F(v))\,\mathrm{d}v$$

これに対して，第 1 価格オークションでは，入札戦略 b の期待値を考える．これは次のように計算できる．

$$\int_0^1 b(v_1)F(v_1)f(v_1)\,\mathrm{d}v_1$$

これでは全然形が違うので計算できないと思われるかもしれないが，b は具体的な形がわかっている．$\frac{f(v_1)}{F(v_1)}(v_1-b(v_1)) = b'(v_1)$ を満たすのである．

これを使うため, 部分積分を用いて形を変形する．$\left(F^2\right)' = 2Ff$ であることに注意しよう．すると，次のような計算ができる．

— Notes —

Date:

$$\int_0^1 b(v_1)F(v_1)f(v_1)\,\mathrm{d}v_1 = \frac{1}{2}\Big[\left[b(1)\left(F(1)\right)^2 - b(0)\left(F(0)\right)^2\right] - \int_0^1 b'(v_1)\left(F(v_1)\right)^2 \mathrm{d}v_1\Big]$$

$F(1)=1, F(0)=0$ であることに注意する．すると，

$$\begin{aligned}\int_0^1 b(v_1)F(v_1)f(v_1)\,\mathrm{d}v_1 &= \frac{1}{2}\Big[b(1) - \int_0^1 b'(v_1)\left(F(v_1)\right)^2 \mathrm{d}v_1\Big] \\ &= \frac{1}{2}\Big[b(1) - \int_0^1 (v_1 - b(v_1))F(v_1)f(v_1)\,\mathrm{d}v_1\Big] \\ &= b(1) - \int_0^1 v_1F(v_1)f(v_1)\,\mathrm{d}v_1\end{aligned}$$

ここで，$b(0)=0$ とする．これは評価値 0 の人は入札額が 0 ということである．

さらに $b(1)$ を計算するために，$\frac{f(v_1)}{F(v_1)}(v_1 - b(v_1)) = b'(v_1)$ を変形した以下の式を用いる．

$$f(v_1)(v_1 - b(v_1)) = b'(v_1)F(v_1)$$

両辺を積分すれば

$$\int f(v_1)(v_1 - b(v_1))\,\mathrm{d}v_1 = \int b'(v_1)F(v_1)\,\mathrm{d}v_1$$

部分積分を用いると，

$$\int b'(v_1)F(v_1)\,\mathrm{d}v_1 = [b(v_1)F(v_1)]_0^1 - \int b(v_1)f(v_1)\,\mathrm{d}v_1$$

$F(1)=1, F(0)=0$ であることに注意する．また，$b(0)=0$ であるとする．このとき，以下の計算が可能である．

$$\int f(v_1)(v_1 - b(v_1))\,\mathrm{d}v_1 = b(1) - \int b(v_1)f(v_1)\,\mathrm{d}v_1$$

$$\Rightarrow b(1) = \int f(v_1)(v_1)\,\mathrm{d}v_1$$

これを収入の式に代入すれば

$$\int_0^1 b(v_1)F(v_1)f(v_1)\,\mathrm{d}v_1 = \int_0^1 v_1(1 - F(v_1))f(v_1)\,\mathrm{d}v_1$$

となって，これは第 2 価格オークションの収入と完全に一致する．このように，収入が一致することは**収入同値定理**として知られている．実はさらに一

般的に「一番評価値が高い人に落札してもらう」「評価値 0 の人には支払いが発生しない」という 2 点が一致すればどんなオークション方式でも収入が同じになることが知られている．

つまりは収入の面ではどんな方式にしようと変わらないのである．これは少し設定を崩すと成り立たなくなったりするが，比較のための基準として重要な結果である．

課題 12.

1. $F(v) = v$ とするとき，$n = 2$ のときの第 1 価格オークションの入札戦略を計算しなさい．ただし，数式処理ソフトを用いて良い．
2. 部分積分によって以下を示せ．

$$\int_0^1 \left(\int_0^{v_1} v_2 f(v_2)\, \mathrm{d}v_2 \right) f(v_1)\, \mathrm{d}v_1 = \int_0^1 v f(v)(1 - F(v))\, \mathrm{d}v$$

— Notes —

Date: