数理・データサイエンスのための基礎数学[*1]

石田 淳

aishida@kwansei.ac.jp

目次

第 1 章

数理・データサイエンスのための基礎数学：イントロダクション

1.1　シラバス

1.1.1　授業目的

本講義は，数理モデリング，データサイエンスのために必要となる基礎的な数学的知識を紹介することを目的とする．

前半は，各種応用数学の基本言語となる論理・集合・確率の基本を導入する．

後半は，数理モデリング・データサイエンスでの応用を前提にさらに実践につながるいくつかのトピックを紹介する．

1.1.2　到達目標

学部レベルのデータ分析系科目や数理社会学を受講する際の数学的前提知識を得ること，同時に数学やデータ分析への苦手意識をなくすこと．

1.1.3　授業方法

対面授業形式．数理モデリング，データサイエンスのために必要となる基礎的な数学的知識をいくつかの項目に分けて紹介していく．詳細な展開や証明のフォローよりも，各項目についての直感的理解と，データ分析系科目や数理社会学とのつながりを理解することに重点を置く．練習問題などで実際に手を動かして理解する機会を設ける．また，Python や R を用いた計算方法も紹介する．

1.1.4 成績評価

定期試験 80%，平常リポート（授業内で提出のリポート，LUNA 提出のテスト・リポート）20%

ただし，定期試験が行われなかった場合，平常リポートと LUNA によるテストで成績を付ける．

1.1.5 授業計画

	内容	授業日	備考
(1)	1 章 イントロダクション	4/14	
(2)	2 章 論理 (1)	4/21	
(3)	3 章 論理 (2)	4/28	
(4)	4 章 集合 (1)	5/12	
(5)	5 章 集合 (2)	5/19	
(6)	6 章 確率 (1)	5/26	
(7)	7 章 確率 (2)	6/2	
(8)	8 章 総和	6/9	
(9)	9 章 微分	6/16	
(10)	10 章 積分	6/23	
(11)	11 章 指数と対数	6/30	
(12)	12 章 ベクトル	7/7	
(13)	13 章 行列	7/14	
(14)	14 章 さらに学ぶために	7/21	

1.2 授業方法 (対面の場合)

- 今のところ対面講義方式を予定している．本講義ノートと参考資料をもとに板書と説明によって授業を進める．
- 各自感染症対策をしっかりと行うこと．とくにマスク着用は必須でありマスクは鼻口を覆う形で適切に着用すること．感染症対策について，講師の指示に従

わない人は教室の安全確保のため受講を拒否することがある．

- 原則として，授業開始後 30 分以降の入室は不可とする．
- 各章の内容は可能な限り事前に OneDrive にアップロードしておく．
- 各回ごとに 1-2 回で数理・データサイエンスにかかわる基本的な数学のトピックを紹介していく．
- 想定する事前知識は，高校数学の数 I レベルであるが，できるだけ基本的なレベルから講義をすすめるので，最悪中学校レベルでもよい．
- データ分析系科目 (入門，基礎，応用)，数理社会学とのつながりを意識したトピックを選択している．これらの科目と並行して取るとよいだろう．
- 実際の計算過程を体験してもらうために R と Python での計算例を参考資料において紹介する．関心のある人は自分で試してみよう．
- 平常リポートとして「課題」もしくは「小テスト」を LUNA 上で課すことがある．課題内容については，授業中のみ指示し，授業外では一切繰り返さない．
- 課題ファイルはコピペチェッカーにかける．異なる受講者から同一ファイルが提出された場合は，どのような事情であろうと提出した受講者全員を不可とする．
- 他の受講者の受講の妨げとなるような行為は厳禁である．逸脱行為にはゼロトレランスで対応する．
- 質疑応答はメール (aishida@kwansei.ac.jp) もしくはオフィスアワー (木曜 4 限，第一教授館本館 103) で行う．

1.3　参考文献

数学（含む統計）の参考書はそれこそ星の数ほどあって，質もレベルもバラバラである．一番いいのは，図書館や大きな本屋に行って，自分の勉強したい分野について片っ端から確認して，合ってそうなものを選ぶことである．

ここでは，授業全体にかかわる参考文献を挙げる．

■本ノートの参考文献　本講義ノートの下敷きとして，論理と集合については，

矢野健太郎，1968，『教養の数学（改訂版）』裳華房．

そして，確率，微分，積分，ベクトル，行列については，

　　矢野健太郎・田代嘉宏，1993,『社会科学者のための基礎数学（改訂版）』裳華房.

を用いた．矢野・田代（1993）は，社会科学者に求められる基礎数学の知識をまとめたものとしてちょうどよい．ただし，いくつか誤植が残っているが，誤植を見つけることも勉強である．

■数学全般　中学数学レベルから，ちゃんと数学をやり直したいという場合は，なんといっても

　　松坂和夫，2019,『新装版　数学読本 1-6』岩波書店.

である．独学用にめちゃめちゃ丁寧に一歩一歩解説されている．逆にこれであかんかったらなにをやってもあかん（かも）．あと，読み物的だが

　　結城浩『数学ガールの秘密ノート』シリーズ（SB クリエイティブ）

も一つのトピックを丁寧に説明していてよい．

■論理・集合

　　嘉田勝，2008,『論理と集合から始める数学の基礎』日本評論社.
　　中内伸光，2009,『ろんりと集合』日本評論社.
　　石田淳，2017,『集合論による社会的カテゴリー論の展開——ブール代数と質的比較分析の応用』勁草書房.

■線形代数

　　平岡和幸・堀玄，2004,『プログラミングのための線形代数』オーム社.
　　岡太彬訓，2008,『データ分析のための線形代数』共立出版.
　　ステファン・ボイド，リーヴェン・ヴァンデンベルグ，2021,『スタンフォード　ベクトル・行列からはじめる最適化数学』講談社.

■統計関連

永野裕之 (著)・岡田謙介 (監修)，2016,『この 1 冊で腑に落ちる　統計学のための数学教室』ダイヤモンド社
永田靖・棟近雅彦，2001,『多変量解析法入門』サイエンス社
倉田博史・星野崇宏，2009,『入門統計解析』新世社

1.4　プログラム

数学的概念を自分のものとするためには，実際にいろいろと動かしてみるのがよい．もちろん，プリミティブには「手を動かす」ことになるのだが，プログラミング言語を使って動かしてみることもできる．本講義では，参考として Python と R によるプログラムを随時紹介する．

■Python　Python は，近年データサイエンスの領域で主流になりつつあるスクリプト型プログラミング言語である．汎用言語なので，各種のパッケージを導入することでいろいろなことができる．

数式演算やデータ分析目的の Python 利用は，現時点では Anaconda (`https://www.anaconda.com/`) を導入して Jupyter を使うのがもっとも手っ取り早い．Anaconda に汎用テキストエディタである Visual Studio Code (`https://code.visualstudio.com/`) を組み合わせてもいいだろう．とりあえず，web ベースで Jupyter ライクに Python を使ってみる場合には，Google Colaboratory (`https://colab.research.google.com/`) というサービスがある．

■R　R は，データ分析のためのオープンソース・フリーソフトウェアのプログラミング言語環境である（`https://www.r-project.org/`）．R 単体での利用も可能であるが，近年は R のための統合開発環境（IDE）である RStudio からの利用が一般的になっている（`https://rstudio.com/`）．

とりあえず, R や RStudio を触ってみたい場合には, RStudio と同じ機能を web 上で利用できるサービスである RStudio Cloud がある（`https://rstudio.cloud/`）．

がっつり使いたい場合は, 自分のマシンに R と RStudio をインストールする．基本

的に, OS (Windows, macOS, Linux) の違いに応じてそれぞれのインストラクションに沿ってインストールしていけば問題ないだろう．Windows については, ユーザー名を全角文字で設定している場合, R や RStudio がうまくインストールできないという問題がある．この問題への対処を含めて，現時点で日本語で一番まとまっている情報は, 矢内勇生氏のウェブサイト（`https://yukiyanai.github.io/jp/resources/`）である．

■コード置き場　以下の URL に Jupyter notebook ファイル (.ipynb)，rmarkdown ファイル（.rmd）とその HTML へのリンクを置いている．

`https://github.com/aishidajt9/BasicMath`

1.5　授業を始めるにあたって

以下，授業を始めるにあたっていくつか伝えたいことを列挙する．

- 数学は言語
  日本の教育システムの悪しき影響により，数学は理系のもの，という意識がいまだに強いかもしれないが，数学は学問分野を問わず共通の科学的言語である．ただし，どれくらい使うか，どのように使うかは分野によって異なる．
- 数学は道具
  数学科でやるような数学のための数学は例外として，われわれは数学を何かのための道具として使っている．道具の使い方を学習しよう．
- ギャップのない理解を目指す
  理解にギャップ（ブラックボックス）があると気持ち悪いし，わかったという感覚になりにくい．時間がかかってもいいのでできるだけ，ロジックを一歩一歩フォローしよう．
- 飲み込みは悪くてよい
  一部の天才を除き，飲み込みがよいというのは逆に何も理解せずに噛み砕かずに飲み込んでいることを示唆している．それだと消化不良を起こすであろう．それよりも，ゆっくりと咀嚼するほうがよい．

課題 1.1. 以下のことを教えてください.

(1) なぜこの講義を受けようと思ったか. 受講動機

(2) 現在の自分の数学レベルはどれくらいか. 自己評価

(3) この講義で何を得たいか. 目標

(4) そのほか, 要望など何でも.

第 2 章

論理 (1)

2.1 論理を学ぶ目的

「数理・データサイエンスのための基礎数学」のはじめに**論理** (logic) の基礎の基礎を学ぶ[*1].

論理とは，単純にいえば，正しい論証の仕方を形式的に示したものである．この論理は，日常生活における論理的思考法はもちろんのこと，数学においてもっとも重要となる議論の進め方である「証明」の形式的基礎を与えるものであり，大変重要である．また，プログラミング言語を学ぶ上でも論理演算の基礎は必要になってくる．さらに，いわゆる「論理的思考」を身につけるには，形式論理を学ぶのがもっとも近道である．

もっとも基本となる論理には大きく分けて

- **命題論理** (propositional logic)：それ単体で真偽の定まる命題を扱う論理
- **述語論理** (predicate logic)：不特定の対象 x を含む命題関数を扱う論理

の 2 種類がある．本章と次章では，命題論理を中心に学び，後の集合のところで述語論理にも若干触れる．

2.2 命題

命題 (proposition) とは「意味のある文章で，それが真（意味的に正しい）であるか偽（間違い）であるかがはっきりと判定できるもの」をいう．重要な点は真偽が定

[*1] 石田 (2017) の 1 章と大部分重複している．

まっているということであって，真偽が定まっていれば，どんな文章も命題である．例えば

「こんにちは」
「石田淳」
「$1 + 2 =$ カバ」

は命題ではない．しかし，

「昨日は雨が降った」
「石田淳は男である」
「$1 + 2 = 3$」
「$1 + 2 = 50$」

は命題である．1 番目の命題は場所を特定すれば真偽を判定できる．2, 3 番目の命題は真の命題であり[*2]，4 番目の命題は偽の命題である．

ここでいう真，偽は命題が取り得る値であり**真理値** (truth value) とも呼ばれる．以下では簡単のため「真」を T (true)，「偽」を F (false) とも表す[*3]．

2.3　基本的な論理演算

数学言語においても日常言語においても，「推論を進めていく」あるいは「論証を進めていく」という論理的操作は，いくつかの基礎的な真の命題から出発して，それらを組み合わせていくことによって，新しい真の命題を求めていくことに他ならない．ゆえに，正しい推論のためには，命題の合成の仕方と，合成した後の命題の真偽がどのようになるのかについて知っておく必要がある．

以下では，1 つ 1 つの命題を抽象化して記号で $p, q, r, \cdots$ などと記す．1 つ 1 つの命題を合成して新しい命題を作り出す操作を**論理演算** (logical operation) と呼び，そのときに用いられる演算記号を**論理演算子**と呼ぶ．合成される前の 1 つ 1 つの命題を**単一命題**，合成された命題を**合成命題**と呼ぶ．

[*2] はたしてどのような根拠でこれらの命題が真であると言えるのだろうか？
[*3] 真を 1，偽を 0 と表す場合もある．ブール値ともいわれる．

本節では，もっとも基本的な論理演算である論理和，論理積，否定を導入する．

2.3.1　論理和 (選言)

命題 p と q を「または」で結んでできる合成命題を，p と q の**論理和** (**選言**)(logical OR, disjunction) とよび，論理演算子 $\vee$ を用いて，

$$p \vee q \tag{2.1}$$

と表す．論理和 $p \vee q$ は，p と q のいずれか一方，もしくは両方が真のとき真であることを示している．言い換えると，論理和 $p \vee q$ が偽であるのは，p と q の両方が偽のときだけである．

このように p と q の真理値の組み合わせによって，合成命題 $p \vee q$ の真理値を求めることができる．これを視覚的にわかりやすくするために表の形にしたものが**真理表** (truth table) である．論理和の真理表は表 2.1 のようになる．

表 2.1　論理和 $\vee$ の真理表

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

2.3.2　論理積 (連言)

命題 p と q を「かつ」で結んでできる合成命題を，p と q の**論理積** (**連言**)(logical AND, conjunction) とよび，論理演算子 $\wedge$ を用いて，

$$p \wedge q \tag{2.2}$$

と表す．論理積 $p \wedge q$ は，p と q の両方が真のときだけ真であり，それ以外の真偽の組み合わせは偽であることを示している．論理積の真理表は表 12.1 のようになる．

表 2.2　論理積 $\wedge$ の真理表

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

2.3.3　否定

命題 p を「でない」によって打ち消す命題を p の**否定** (negation) とよび，論理演算子 $\neg$ を用いて，

$$\neg p \tag{2.3}$$

と表す．否定の真理表は表 2.3 のようになる．なお，単一命題 p とその否定 $\neg p$ を合わせて**リテラル** (literal) という．

表 2.3　否定 $\neg$ の真理表

p	$\neg p$
T	F
F	T

2.3.4　さらに複雑な合成命題

ここまで導入した論理和，論理積，否定は命題論理演算のもっとも基本的なものであって，これらを組み合わせていくことによって，さらに複雑な合成命題を作ることができる．逆に言えば，後で導入する条件文や双条件文などの論理演算を含む合成命題は，すべてこれら 3 つの基本的な論理演算に還元できることが知られている．

では，例えば，

$$\neg p \vee \neg q \tag{2.4}$$

という合成命題の真理表はどのようになっているのだろうか．ここでは，矢野 (1968) に解説されている方法に従って，表 2.4 のように，1 から順に真理表を構成しよう．2 で単一命題 p, q の真理値を与え，3 でそれぞれの否定，4 で否定命題の論理和の真理値を求める．この手順によって合成命題の真理値を得る．

表 2.4　合成命題の真理表作成例

1

p	q	$\neg$	p	$\vee$	$\neg$	q
T	T					
T	F					
F	T					
F	F					

2

p	q	$\neg$	p	$\vee$	$\neg$	q
T	T		T			T
T	F		T			F
F	T		F			T
F	F		F			F

3

p	q	$\neg$	p	$\vee$	$\neg$	q
T	T	F	T		F	T
T	F	F	T		T	F
F	T	T	F		F	T
F	F	T	F		T	F

4

p	q	$\neg$	p	$\vee$	$\neg$	q
T	T	F	T	F	F	T
T	F	F	T	T	T	F
F	T	T	F	T	F	T
F	F	T	F	T	T	F

以下，複雑な合成命題において括弧がある場合は，その内部を先に演算せよという意味であることを約束する．

練習問題 2.3

次の命題の真理表を作れ（矢野 (1968: 9) より）．

(1) $\neg(p \wedge q)$　(2) $\neg(p \vee q)$　(3) $\neg p \wedge \neg q$

(4) $p \vee \neg p$　(5) $p \wedge \neg p$　(6) $\neg(p \vee q) \wedge p$

(7) $(p \vee \neg q) \wedge \neg p$　(8) $\neg[(p \wedge q) \vee (\neg p \wedge \neg q)]$

2.4　条件法と双条件法

本節ではさらに条件法 $\rightarrow$ と双条件法 $\leftrightarrow$ という論理演算を導入する．

2.4.1　条件法

日常言語の「ならば（if...then）」に対応する論理演算が**条件法** (conditional) であり，論理演算子として $\rightarrow$ を用いる．「p ならば q である」という合成命題は

$$p \rightarrow q \tag{2.5}$$

と表される．合成命題 $p \rightarrow q$ を条件文という．

p と q がともに真であれば $p \rightarrow q$ は明らかに真であり，p が真であるにもかかわらず q が偽であれば $p \rightarrow q$ は明らかに偽である．一方，p が偽の時は q の真偽にかかわらず，合成命題 $p \rightarrow q$ は真であると**約束する**．

p が偽の時の $p \rightarrow q$ の真理値，とくに p が偽で q が真の時の真理値は，われわれの日常言語における直感にやや反する．たとえば，「雨が降れば地面がぬれる」という命題が，雨が降っていないときにも真といえるかどうかは少々心許ない．そこで，この条件文を導入するときには，いろいろな教科書でいろいろな理由が述べられている．直感的な説明としては「条件である p がそもそも偽なのだから，p が成り立てば q が成り立つという主張をしても**間違いではない**」という言い方がある．そして，命題論理においては真偽の 2 値だけを許容しているので，「間違いではない」場合は真を当てるのが妥当であるとの説明の仕方だ．やや納得感にかけるかもしれないが，そのように定義することが 2 値命題論理を展開する上でもっともリーズナブルであるし

扱いやすいと考えてほしい*4.

結局，条件法の真理表は表 2.5 のようになる.

表 2.5　条件法 $\to$ の真理表

p	q	$p \to q$
T	T	T
T	F	F
F	T	T
F	F	T

2.4.2　双条件法

これまでに導入した論理演算 $\wedge$ と $\to$ を用いて，さらに複雑な合成命題

$$(p \to q) \wedge (q \to p) \tag{2.6}$$

を考えることができる．これは日常言語に直すと「p ならば q であり，かつ，q ならば p である」ということである．このような合成命題を作る論理演算を**双条件法** (biconditional) と呼び，双条件文を

$$p \leftrightarrow q \tag{2.7}$$

と書くことにする．この双条件法の真理表は表 2.6，まとめると表 2.7 のようになる．ただし，表下の数字は真理表を書き込んでいく順番である.

練習問題 2.4

次の命題の真理表を作れ（矢野 (1968: 13) より）.

---

*4 構成論的な立場から，条件法の定義を正当化すると次のようになる（ただし次章の知識が必要）．日常言語における条件文の特徴として最低限満たすべき 2 つの特徴を考える．それは，(1) 逆命題との非同値性，(2) 推移性である．これを命題論理の言葉に直すと，(1) $p \to q$ と $q \to p$ は論理的同値ではない，(2) 合成命題 $((p \to q) \wedge (q \to r)) \to (p \to r)$ がトートロジーである，ということになる．そして，これらの要件を満たす定義は，表 2.5 の定義だけである．詳しくは，戸田山和久，2000，『論理学をつくる』名古屋大学出版会，pp. 81-3，を参照.

表 2.6　合成命題 $(p \to q) \land (q \to p)$ の真理表

p	q	$(p$	$\to$	$q)$	$\land$	$(q$	$\to$	$p)$
T	T	T	T	T	T	T	T	T
T	F	T	F	F	F	F	T	T
F	T	F	T	T	F	T	F	F
F	F	F	T	F	T	F	T	F
		1	2	1	3	1	2	1

表 2.7　双条件法 $\leftrightarrow$ の真理表

p	q	$p \leftrightarrow q$
T	T	T
T	F	F
F	T	F
F	F	T

(1) $q \to (p \lor q)$　(2) $(p \land q) \to p$　(3) $(p \to p) \lor (p \to \neg p)$

(4) $(p \land q) \leftrightarrow (q \land p)$　(5) $\neg p \leftrightarrow (p \to \neg q)$　(6) $(p \to q) \leftrightarrow (\neg p \lor q)$

2.5　恒真命題と恒偽命題

命題を構成する単一命題の真偽の組み合わせのすべての論理的可能性に対して，**常に真**となるような合成命題を**恒真命題**もしくは**トートロジー** (tautology) という．

例えば，命題 $p \lor \neg p$ は表 2.8 のとおり恒真命題である．

この命題は，「p であるか p でないかのいずれかである」という意味なので，p が真であっても偽であっても，合成命題が真になるのは当たり前のように感じるだろう．命題 $p \lor \neg p$ が恒真命題であるという性質は**排中律**と呼ばれる．

次に，やや複雑になるが $p \to (p \lor q)$ もやはり恒真命題である（表 2.9）．

表 2.8　合成命題 $p \vee \neg p$ の真理表

p	p	$\vee$	$\neg$	p
T	T	T	F	T
F	F	T	T	F

表 2.9　合成命題 $p \to (p \vee q)$ の真理表

p	q	p	$\to$	$(p$	$\vee$	$q)$
T	T	T	T	T	T	T
T	F	T	T	T	T	F
F	T	F	T	F	T	T
F	F	F	T	F	F	F

反対に，命題を構成する単一命題の真偽の組み合わせのすべての論理的可能性に対して，**常に偽**となるような合成命題を**恒偽命題**もしくは**矛盾命題** (contradiction) という．例えば，合成命題 $p \wedge \neg p$ は恒偽命題である（表 2.10）．これを**矛盾律**という．

表 2.10　合成命題 $p \wedge \neg p$ の真理表

p	p	$\wedge$	$\neg$	p
T	T	F	F	T
F	F	F	T	F

ほぼ自明のことであるが，恒真命題の否定は恒偽命題であり，恒偽命題の否定は恒真命題である．

恒真命題は，それを構成する命題のどのような真偽の可能性に対しても真となる命題である．つまり，内容はともかくとして**形式的に正しい**命題である[*5]．この形式的

[*5] 反対に，恒偽命題は，それを構成する命題のどのような真偽の可能性に対しても偽となる命題であり，**形式的に誤った**命題である．

な正しさが妥当な推論，そして妥当な推論による証明において決定的に重要となる．

練習問題 2.5

(1) 恒真命題の具体例を作れ．

(2) 恒偽命題の具体例を作れ．

第 3 章

論理 (2)

引き続き論理を学んでいく．ここでは妥当な推論と証明の形式を学ぶ．

3.1　論理的含意と論理的同値

これまで導入してきた論理演算は，いくつかの単一命題から新たな命題を作り出すためのものであったが，以下に導入する論理的含意と論理的同値は，記述のレベルが異なり，**命題間の関係**を述べるための概念である．

3.1.1　論理的含意

$p \wedge q$ という命題と $\neg p \vee q$ という命題の関係性を見るために真理表を作成する．表 3.1 から，$p \wedge q$ が真であるときには，$\neg p \vee q$ も必ず真であることが分かる．

表 3.1　$p \wedge q$，$\neg p \vee q$ の真理表

p	q	$p \wedge q$	$\neg p \vee q$
T	T	T	T
T	F	F	F
F	T	F	T
F	F	F	T

このように，2 つの合成命題 P と Q について，P が真であるようなすべての論理的可能性に対して，Q が常に真であるとき，P は Q を**論理的に含意する** (logically imply) といい，

$$P \Longrightarrow Q \tag{3.1}$$

で表す．上の例では，

$$p \wedge q \Longrightarrow \neg p \vee q \tag{3.2}$$

つまり，命題 $p \wedge q$ は命題 $\neg p \vee q$ を論理的に含意する．

同じ矢印なので混同しやすいが，**条件法 $\rightarrow$ が合成命題を作る論理演算子であるのに対して，論理的含意 $\Longrightarrow$ は命題間の関係を述べるための記号であり，記述のレベルが異なる**．このことは強調してもしすぎることがない[*1]．しかしながら，条件法 $\rightarrow$ と論理的含意 $\Longrightarrow$ は，恒真命題という考え方を通して密接に関連する．

ここで条件文 $(p \wedge q) \rightarrow (\neg p \vee q)$ の真理表を構成してみる．表 3.2 から分かるとおり，条件文 $(p \wedge q) \rightarrow (\neg p \vee q)$ は恒真命題である．そして，条件文が恒真命題の時には，前件が真であれば結果も必ず真になる．反対に，前件が真のとき結果も必ず真となっていれば，その条件文は必ず恒真命題である．ゆえに，以下のことが成立する．

> **条件文 $P \rightarrow Q$ が恒真命題であるとき，そしてそのときに限り，P は Q を論理的に含意する．すなわち，$P \Longrightarrow Q$．**

表 3.2　$(p \wedge q) \rightarrow (\neg p \vee q)$ の真理表

p	q	$p \wedge q$	$\rightarrow$	$\neg p \vee q$
T	T	T	T	T
T	F	F	T	F
F	T	F	T	T
F	F	F	T	T

[*1] ただし，テキストによっては論理演算子 $\rightarrow$ そのものを「含意」と呼ぶ場合があるので，注意が必要である．

3.1.2　論理的同値

次に，$\neg p \wedge q$ と $\neg(p \vee \neg q)$ という命題の関係性を検討する．表 3.3 から，p と q との組み合わせに対して，$\neg p \wedge q$ と $\neg(p \vee \neg q)$ の真理値はつねに等しいことが分かる．このように 2 つの命題 P と Q があって，それらを構成する単一命題の組み合わせの論理的可能性に対して，つねに真理値が等しいとき，P と Q は**論理的に同値である** (logically equivalent) といって，

$$P \Longleftrightarrow Q \tag{3.3}$$

と表す[*2]．上の例でいうと，

$$(\neg p \wedge q) \Longleftrightarrow \neg(p \vee \neg q) \tag{3.4}$$

である．

表 3.3　$\neg p \wedge q$，$\neg(p \vee \neg q)$ の真理表

p	q	$\neg p \wedge q$	$\neg(p \vee \neg q)$
T	T	F	F
T	F	F	F
F	T	T	T
F	F	F	F

条件法と含意の関係と同様に，**双条件法 $\leftrightarrow$ が合成命題を作る論理演算子であるのに対して，論理的同値 $\Longleftrightarrow$ は命題間の関係を述べるための記号であり，記述のレベルが違う．**

さて双条件文 $\neg p \wedge q \leftrightarrow \neg(p \vee \neg q)$ の真理表を構成する．表 3.4 から分かるとおり，双条件文 $\neg p \wedge q \leftrightarrow \neg(p \vee \neg q)$ は恒真命題であり，双条件文が恒真命題であるということは，結局論理的同値であることと同じことである．まとめると，以下のことが成立する．

[*2] 同値の記号として $\equiv$ を使うこともある．

双条件文 $P \leftrightarrow Q$ が恒真命題であるとき，そしてそのときに限り，P と Q は論理的に同値である．すなわち，$P \Longleftrightarrow Q$．

表 3.4　$\neg p \wedge q \leftrightarrow \neg(p \vee \neg q)$ の真理表

p	q	$\neg p \wedge q$	$\leftrightarrow$	$\neg(p \vee \neg q)$
T	T	F	T	F
T	F	F	T	F
F	T	T	T	T
F	F	F	T	F

P と Q が論理的に同値であるということは，その命題の実質的意味内容がなんであれ，論理的には，P を Q に，もしくは Q を P に置き換えても，いっこうに差し支えないということを意味している．

また，条件法と双条件法との関係から，含意と同値の関係について次のことが成り立つ．

$P \Longrightarrow Q$ と $Q \Longrightarrow P$ が同時に成立するのは，$P \Longleftrightarrow Q$ のときそのとき限りである．

3.1.3　条件文の逆，裏，対偶

条件文 $p \to q$ に対して，$q \to p$ とした命題を，条件文 $p \to q$ の**逆**という．また，条件文 $p \to q$ に対して，$\neg p \to \neg q$ とした命題を，条件文 $p \to q$ の**裏**という．最後に，条件文 $p \to q$ に対して，$\neg q \to \neg p$ とした命題を，条件文 $p \to q$ の**対偶** (contraposition) という．条件文 $p \to q$ の対偶はまた，条件文 $p \to q$ の逆命題の裏，あるいは裏命題の逆と言い換えることもできる．条件文 $p \to q$ についての逆，裏，対偶の関係を図 3.1 に示す．

命題 p と q について，条件文 $p \to q$ とその逆，裏，対偶命題の真理表は表 3.5 のようになる．

表から，条件文 $p \to q$ とその対偶 $\neg q \to \neg p$ は論理的同値であることが分かる．こ

$$\begin{tikzcd}
p \to q \arrow[r, "逆"] \arrow[d, "裏"'] \arrow[dr, "対偶"] & q \to p \arrow[d, "裏"] \\
\neg p \to \neg q \arrow[r, "逆"'] & \neg q \to \neg p
\end{tikzcd}$$

図 3.1　条件文の逆，裏，対偶の関係図

表 3.5　条件文の逆，裏，対偶の真理表

p	q	$p \to q$	$q \to p$	$\neg p \to \neg q$	$\neg q \to \neg p$
T	T	T	T	T	T
T	F	F	T	T	F
F	T	T	F	F	T
F	F	T	T	T	T

のことは，条件文 $p \to q$ が真であることを論証したい場合，その対偶 $\neg q \to \neg p$ が真であることを論証することで置き換えてもよい，ということを示している[*3].

3.1.4　必要条件，十分条件，必要十分条件

いま，条件文 $p \to q$ が恒真命題である，つまり，p は q を論理的に含意する，と仮定する．つまり，$p \Longrightarrow q$．ここで，$p \to q \Longleftrightarrow \neg q \to \neg p$ であったので，$p \Longrightarrow q$ は $\neg q \Longrightarrow \neg p$ と言い換えてもよい．$\neg q \Longrightarrow \neg p$ は，「q であるときは p であるかもしれないしそうでないかもしれないが，q でないときには絶対に p ではない」ことを表している．言い換えると

「p であるためには，ぜひとも q であることが必要である」

[*3] 条件文 $p \to q$ の裏（$\neg p \to \neg q$）と逆（$q \to p$）も論理的同値である．この 2 つの命題はそれぞれに対して対偶であることに注意せよ．

ことを示している．したがって $p \Longrightarrow q$ のとき，q を p であるための**必要条件** (necessary condition) とよぶ．

同時に，$p \Longrightarrow q$ は，

「p であれば，それで十分 q が成立する」

ことを示している．したがって $p \Longrightarrow q$ のとき，p を q であるための**十分条件** (sufficient condition) とよぶ．

一方，p と q が論理的同値であるとき，つまり $p \Longleftrightarrow q$ のときには，$q \Longrightarrow p$ と $p \Longrightarrow q$ が同時に成り立つのであるから，**p は q であるための必要条件であり，同時に，p は q であるための十分条件**である．この場合，**p は q であるための必要十分条件** (necessary and sufficient condition) であるという．同様にして，$p \Longleftrightarrow q$ のとき，**q は p であるための必要十分条件**である．

必要条件，十分条件，必要十分条件は，言葉による説明だけでは直感的に分かりにくい．これらは，集合論を導入した後に，ベン図によって図式化したほうが理解が容易であるので，5 章において再び取り上げることにする．

練習問題 3.1

(1) $[(p \to q) \wedge (q \to r)] \to (p \to r)$ が恒真命題であること，言い換えると $[(p \to q) \wedge (q \to r)] \Longrightarrow (p \to r)$ が成立することを示せ（三段論法）．

(2) $\neg(p \vee q) \Longleftrightarrow \neg p \wedge \neg q$ が成立することを示せ（ド・モルガンの法則 (1)）．

(3) $\neg(p \wedge q) \Longleftrightarrow \neg p \vee \neg q$ が成立することを示せ（ド・モルガンの法則 (2)）．

3.2　推論と証明

3.2.1　推論の方法

いくつかの命題——これらを**前提** (premise) という——から他の命題——これを**結論** (conclusion) という——が論理的に導かれる（含意される）ことを主張することを**推論** (inference) という．

推論は，前提の論理積が結論を論理的に含意するとき，**妥当**であるという．すなわ

ち，推論の諸前提を $A_1, A_2, \ldots, A_n$ と表し，結論を C とすると，**妥当な推論**とは，

$$A_1 \wedge A_2 \wedge \cdots \wedge A_n \Longrightarrow C \tag{3.5}$$

が成り立つような推論に他ならない．前提の論理積が結論を論理的に含意するということは，**前提すべてが真であれば，結論も間違いなく真である**ということに他ならない．反対に，妥当でない推論を**誤った推論**という．

推論は，

$$\begin{array}{l} \text{前提 1} \\ \text{前提 2} \\ \hline \text{結論} \end{array}$$

などとと書かれる．

以下に，いくつかの妥当な推論，そして誤った推論を示す．まず最初に，

$$\begin{array}{l} p \to q \\ p \\ \hline q \end{array}$$

は妥当な推論である．すなわち，

$$(p \to q) \wedge p \Longrightarrow q \tag{3.6}$$

が成立する[*4]．これは条件法を用いたもっとも単純な妥当な推論であり，**肯定式** (modus ponens) と呼ばれる．次に，

$$\begin{array}{l} p \to q \\ q \to r \\ \hline p \to r \end{array}$$

も妥当な推論であり，**推移律** (transitive law) という．最後に，

$$\begin{array}{l} p \to q \\ \neg q \\ \hline \neg p \end{array}$$

[*4] このことを真理表によって確かめてみよう

は対偶の特性を用いた妥当な推論であり，**否定式** (modus tollens) と呼ばれる．

一方，

$$\begin{array}{l} p \to q \\ q \\ \hline p \end{array}$$

や

$$\begin{array}{l} p \to q \\ \neg p \\ \hline \neg q \end{array}$$

は誤った推論である*5．

3.2.2　証明の仕方

真と認められているいくつかの命題から出発し，妥当な推論によって，他の命題が真であることを示すことを**証明** (proof) という．より厳密に言えば次の通りである．有限個の**公理** (axiom) と呼ばれる命題と有限個の妥当な推論規則から成り立っている体系のことを**公理系** (axiomatic system) という．ここで，出発点となる公理は無前提に（ア・プリオリに）「真」と認められている命題のことを指す．これらの公理から推論規則を用いて導かれる命題のことを**定理** (theorem)，あるいは単に**命題** (proposition) という．そして，公理から定理に至る導出過程を示したものが証明である．

証明方法には，大きく分けて直接法と間接法の 2 通りがある．

■直接証明法　真と認められているいくつかの命題から出発して，これに妥当な推論を直接的に当てはめて，目的の命題が真であることを示す方法．例えば，次の例題で考える．

例題． p が真，q が真，$(p \wedge q) \to r$ が真のとき，r が真であることを証明せよ．

解答． p が真，q が真なので，$(p \wedge q)$ もまた真．よって，肯定式を用いて，

---

*5 これらが誤った推論であることを真理表によって確かめてみよう

$$\begin{array}{l} (p \wedge q) \to r \\ (p \wedge q) \\ \hline r \end{array}$$

より，r は真である．

■間接証明法　背理法 (proof by contradiction) とも呼ばれる．この証明法ではまず，証明しようとする目的の命題が偽であると仮定する．これを前提に加えて推論を進めていくことで，すでに真であることが分かっている命題が偽となることを証明する．そしてこの結果をもって最初の仮定自体が誤り，つまり目的の命題は真でなければならない，と結論づける．先と同じ例題を間接証明法で証明する．
解答． r は偽である，すなわち $\neg r$ は真であると仮定する．ところで

$$(p \wedge q) \to r \iff \neg r \to \neg(p \wedge q) \tag{3.7}$$

より，$(p \wedge q) \to r$ が真であるので，$\neg r \to \neg(p \wedge q)$ も真である．よって，肯定式を用いて，

$$\begin{array}{l} \neg r \to \neg(p \wedge q) \\ \neg r \\ \hline \neg(p \wedge q) \end{array}$$

より，$\neg(p \wedge q)$ は真である．ところで

$$\neg(p \wedge q) \iff \neg p \vee \neg q \tag{3.8}$$

であるので，p と q のいずれかは真ではない．しかしながら，「p が真，q が真」との前提に反する．ゆえに，r は真でなければならない．

練習問題 3.2

次の推論の妥当性を調べよ (矢野 1968: 28).

(1) $\begin{array}{ll} & p \vee q \\ & \neg p \\ \hline \therefore & q \end{array}$　(2) $\begin{array}{ll} & p \vee q \\ & p \to q \\ \hline \therefore & q \end{array}$　(3) $\begin{array}{ll} & p \wedge q \\ & \neg p \to q \\ \hline \therefore & \neg q \end{array}$

次の命題に対して間接証明を与えよ (矢野 1968: 31).

(1) もし $p \vee q$ が真で，$\neg q$ が真なら，p は真である.

(2) もし $p \leftrightarrow q$，$q \rightarrow \neg r$ が真で，しかも r が真であれば，$\neg p$ は真である.

第 4 章

集合 (1)

本章と次章では，現代数学の基盤となる**集合論** (set theory) の基礎を学ぶ[*1]．本章では，集合の基本的定義と各種の集合演算を導入する．

4.1　集合の基本

4.1.1　集合の記法

いくつかの対象（もの）をひとまとめにした「対象の集まり」のことを**集合** (set) という．ただし，数学的には，どのような対象をとってきても，それが集まりの中に入っているかどうかがはっきりと判定できることが求められる．逆に言えば，対象が集まりに入っているかどうかさえ分かれば，どのような対象を持つ集合も考えることができる[*2]．

集合を構成する対象のことを，その集合の**要素**もしくは**元** (element) という．普通，集合は大文字 $A, B, \ldots$ などで表され，要素は小文字 $a, b, \ldots$ などで表される．ある対象 a が集合 A の要素であるとき，a は A に属するといって

$$a \in A, \text{あるいは} \quad A \ni a \tag{4.1}$$

と書く．反対に，a が A の要素でないときは,

$$a \notin A, \text{あるいは} \quad A \not\ni a \tag{4.2}$$

[*1] 石田 (2017) の 2 章と部分的に重複している．

[*2] なお，集合概念の拡張として，集合への所属について曖昧さを許容するファジィ集合 (fuzzy set) という考え方も提案されている．ファジィ集合については石田 (2017) などを参照.

と書く．

集合の中身の書き方は大きく分けて 2 通りある．例えば，1 から 6 までの自然数によって構成される集合 A を記述する場合を考える．もっとも単純な記法は，集合の要素をすべて列挙する方法であり，

$$A = \{1, 2, 3, 4, 5, 6\} \tag{4.3}$$

と書く．このような記法を**外延的記法** (denotation, extension) とよぶ．集合の中身を記述する場合は必ず波括弧 { } を用いる．後に導入する順序対やベクトルの表記では，一般的に丸括弧 ( ) を用いる．丸括弧を用いる場合は順番に意味があるが，集合の記述の際の波括弧の場合，順番には意味がない．つまり，先の集合を

$$A = \{3, 1, 2, 5, 6, 4\} \tag{4.4}$$

と書いても同じである．

あるいは集合 A を

$$A = \{1 \text{ から } 6 \text{ までの自然数 }\} \tag{4.5}$$

さらに一般的には

$$A = \{x \mid x \text{ は自然数であり，かつ，} 1 \leq x \leq 6\} \tag{4.6}$$

などと記述することもできる．このような記法を**内包的記法** (connotation, intension) という．一般的に内包的記法は

$$A = \{x \mid p(x)\} \tag{4.7}$$

という構成で，ここで $p(x)$ は x についての条件を示す命題関数あるいは述語といわれるものである．「集合 A は $p(x)$ が真となるような x によって構成される」というのが，ここでの記述の意味である．命題関数と集合の関係性については，5.1.2 節でさらに詳しく論じる．

4.1.2　特性関数

ある対象 x が集合 A に属しているかどうかを判断する関数として，**定義関数**あるいは**特性関数** (characteristic function) というものを考えることができる．具体的に

は，集合 A の特性関数を φ_A と表記し

$$\varphi_A(x) = \begin{cases} 1 & (x \in A) \\ 0 & (x \notin A) \end{cases} \tag{4.8}$$

と定義する．すなわち，x を入力して，x が集合 A に属していれば 1 を，属していなければ 0 を出力する関数である．

集合 A の定義から対応する特性関数が定まる．反対に，ある特性関数が定まれば対応する集合を定義することができる．つまり，集合を定義することと特性関数を与えることは同じことである．

4.1.3　いろいろな集合

前節の集合 $A = \{1, 2, 3, 4, 5, 6\}$ のように，有限個の要素を含む集合のことを**有限集合** (finite set) という．有限集合の大きさは，集合に含まれる要素の個数によって示される．集合 A の要素数は $|A|$ で表される．

一方，自然数全体の集合を考える．これを $\mathbb{N}$ で表し

$$\mathbb{N} = \{x \mid x \text{ は自然数 }\} \tag{4.9}$$

と定義する．$\mathbb{N}$ の要素は無限個ある．このように，無限個の要素を含む集合を**無限集合** (infinite set) という．無限集合は外延的に

$$\mathbb{N} = \{1, 2, 3, \ldots\} \tag{4.10}$$

と表されることがある．自然数全体の集合 $\mathbb{N}$ と同様にして，整数全体の集合として

$$\mathbb{Z} = \{x \mid x \text{ は整数 }\} \tag{4.11}$$

有理数全体の集合として

$$\mathbb{Q} = \{x \mid x \text{ は有理数 }\} \tag{4.12}$$

そして，実数全体の集合として

$$\mathbb{R} = \{x \mid x \text{ は実数 }\} \tag{4.13}$$

を定義する．これらも無限集合である[*3]．なお本ノートでは以後，$\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ をここで定義した意味で使うことにする．

さて，ここまでは何らかの要素によって構成される集合を考えてきたが，要素を 1 つも含まない集合というものもここで定義しておく．要素を 1 つも含まない集合のことを**空集合** (empty set) といい，$\emptyset$ で表す．当然のことながら，どのような対象 a をもってきても

$$a \notin \emptyset \tag{4.14}$$

が成立する．また，定義上 $|\emptyset| = 0$ である．

4.1.4　部分集合

2 つの集合

$$P = \{1, 3, 5, 7, 9\} \tag{4.15}$$

$$Q = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \tag{4.16}$$

を考える．ここで，集合 P のそれぞれの要素はすべて集合 Q の要素でもある．このように，集合 P の要素が必ず集合 Q の要素であるとき，P を Q の**部分集合** (subset) という．また，このとき Q は P を包含する，あるいは P は Q に包含されるとも表現し，

$$P \subset Q, \text{ あるいは, } Q \supset P \tag{4.17}$$

と表す．部分集合の定義を論理記号を用いて書き直せば，すべての対象 x について

$$x \in P \Longrightarrow x \in Q \tag{4.18}$$

が成り立つとき，$P \subset Q$ ということである．つまり，x が P の要素であるときは必ず，その x が Q の要素でもあるということである[*4]．

[*3] 無限集合は，無限個同士の比較となるので，単純に要素の数によって大きさを定義することができない．そこで，**濃度** (cardinality) という概念が導入されることになるが，ここではこれ以上フォローしない．

[*4] なお，P と Q が部分集合関係にないことを，とくに $P \not\subset Q$ という記号で表すことがある．

定義より，任意の集合 P はそれ自身の部分集合でもある．また，空集合 $\emptyset$ はあらゆる集合の部分集合である．すなわち，

$$P \subset P, \tag{4.19}$$

$$\emptyset \subset P. \tag{4.20}$$

また，部分集合の定義から，集合 P, Q, R について，

$$(x \in P \Longrightarrow x \in Q) \wedge (x \in Q \Longrightarrow x \in R) \Longrightarrow (x \in P \Longrightarrow x \in R) \tag{4.21}$$

であるので，

$$(P \subset Q) \wedge (Q \subset R) \Longrightarrow P \subset R \tag{4.22}$$

となる．つまり，部分集合関係について推移律が成り立つ．

2 つの集合 P, Q について，P の要素は必ず Q の要素であり，同時に，Q の要素は必ず P の要素であるとき，P と Q は**等しい** (equal)，もしくは**相等**であるといい，

$$P = Q \tag{4.23}$$

と表す．相等の定義を論理記号を用いて書き直せば，

$$P = Q \Longleftrightarrow (P \subset Q) \wedge (Q \subset P) \tag{4.24}$$

となる．これは，すべての x について

$$(x \in P \Longrightarrow x \in Q) \wedge (x \in P \Longleftarrow x \in Q) \tag{4.25}$$

つまり，すべての x について

$$(x \in P) \Longleftrightarrow (x \in Q) \tag{4.26}$$

であることを示しており，相等においては，P と Q の構成要素が完全に一致する．この性質は，集合論を公理論的に構成する場合には，「外延性の公理」という最も基本的な公理として設定される．

なお，P と Q が等しくないことを，$P \neq Q$ という記号で表すことがある．また，$P \subset Q$ であるが $P \neq Q$ のとき，とくに P は Q の真部分集合と呼ぶことがある[*5]．

[*5] 矢野 (1968) のように，部分集合の記号を $\subseteq$ とし，真部分集合の記号を $\subset$ とすることで，両者を区別する書き方もあるが，本ノートでは部分集合にのみ記号を与える．

4.1.5　集合族とベキ集合

対象がその集まりに属しているかどうかが判明するのであれば，どのような対象の集まりでも集合として考察できるのであった．そこで，集合を要素とするような集合を考える．こうした集合の集合は，**集合族** (family of sets) と呼ばれる．

ある集合 A の部分集合となる集合をすべて集めた集合は集合族である．この集合族をとくに，**ベキ集合** (power set) といい $\mathcal{P}(A)$ と表す．つまり，ベキ集合の内包的定義は

$$\mathcal{P}(A) = \{B \mid B \subset A\} \tag{4.27}$$

である．

なお，集合 A のベキ集合の要素数は，集合 A が有限であるとすると，

$$|\mathcal{P}(A)| = 2^{|A|} \tag{4.28}$$

となる．つまり，要素数が 2 のベキに等しいことから，ベキ集合という名前が付いた．また，ベキ集合を 2^A という表記で表すこともある．

例えば，$A = \{a, b, c\}$ とすると，A の可能な部分集合は，

$$\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\} \tag{4.29}$$

の $2^3 = 8$ 個である．ゆえに，A のベキ集合 $\mathcal{P}(A)$ は

$$\mathcal{P}(A) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\} \tag{4.30}$$

である．くれぐれも注意すべきは，それぞれの部分集合は $\mathcal{P}(A)$ の要素であって，決して $\{a, b\} \subset \mathcal{P}(A)$ などという関係は成り立たない，ということである[*6]．成立するのは $a \in \{a, b\}$ であり，$\{a, b\} \in \mathcal{P}(A)$ なのであって，集合 $\{a, b\}$ と集合族 $\mathcal{P}(A)$ は，レベルが異なる集合であるということに注意してほしい．

練習問題 4.1

$A = \{1, 2, 3\}$ のベキ集合を求め，ベキ集合の要素間で成り立つ部分集合関係をすべて列挙せよ．

[*6] ただし $\{\{a, b\}\} \subset \mathcal{P}(A)$ は成立する．

4.2 集合間の演算

4.2.1 全体集合

集合を考えるとき，考察の対象となっているものの全体を限定して，その部分集合について考察する場合がある．このとき，考察の対象となる全体を**全体集合**，もしくは**普遍集合** (universal set) といい，U で表す[*7]．

全体集合 U の各部分集合について考えるとき，しばしば簡単な図を描く．このとき，全体集合 U を長方形で表し，部分集合 A を長方形の中の円内部のグレーで表す．このような図を**ベン図** (Venn diagram) という（図 4.1）．

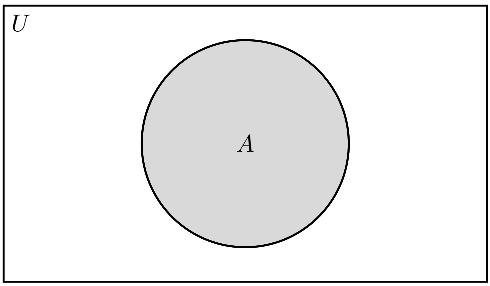

図 4.1　ベン図の例

4.2.2 和集合

全体集合 U の部分集合 A, B について，A または B に含まれるすべての要素の集合を，A と B の**和集合** (union)，または**合併集合**とよんで，

$$A \cup B \tag{4.31}$$

[*7] 全体集合は Ω で表される場合もある．

と表す[*8]．A と B の和集合 $A \cup B$ を内包的記法によって厳密に定義すれば,

$$A \cup B = \{x \mid (x \in A) \vee (x \in B)\} \tag{4.32}$$

となる．例えば,

$$A = \{a, b, c, d, e\}, \quad B = \{c, d, e, f, g, h\} \tag{4.33}$$

であれば,

$$A \cup B = \{a, b, c, d, e, f, g, h\} \tag{4.34}$$

となる．

全体集合 U の部分集合 A, B の和集合 $A \cup B$ をベン図で表せば図 4.2 のようになる．

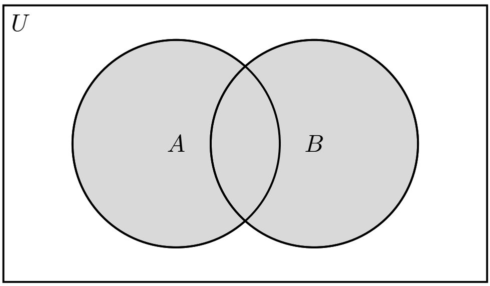

図 4.2　和集合 $A \cup B$ のベン図

4.2.3　積集合

全体集合 U の部分集合 A, B について，A と B の両方に含まれる要素の集合を，A と B の**積集合** (intersection)，または**共通集合**とよんで,

$$A \cap B \tag{4.35}$$

[*8] しばしば，A カップ B と発音される．

と表す*9．A と B の積集合 $A \cap B$ を内包的記法によって厳密に定義すれば,

$$A \cap B = \{x \mid (x \in A) \wedge (x \in B)\} \tag{4.36}$$

となる．例えば,

$$A = \{a, b, c, d, e\}, \quad B = \{c, d, e, f, g, h\} \tag{4.37}$$

であれば,

$$A \cap B = \{c, d, e\} \tag{4.38}$$

となる．

全体集合 U の部分集合 A, B の積集合 $A \cap B$ をベン図で表せば図 4.3 のようになる．

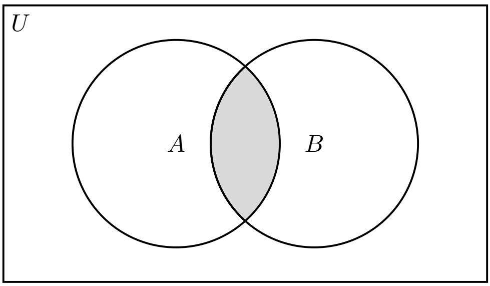

図 4.3　積集合 $A \cap B$ のベン図

とくに，集合 A, B の積集合 $A \cap B$ が空集合のとき，つまり

$$A \cap B = \emptyset \tag{4.39}$$

のとき，A と B は，**互いに素** (mutually disjoint)，あるいは**相互排反** (mutually exclusive) という（図 4.4）．

*9 しばしば，A キャップ B と発音される．

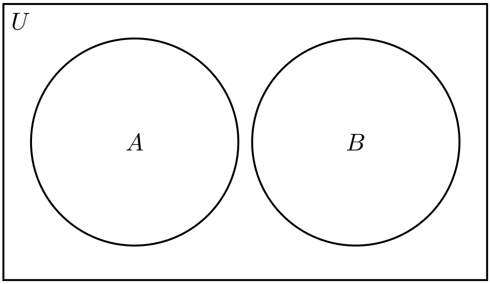

図 4.4　積集合 $A \cap B = \emptyset$ のベン図

4.2.4　補集合

全体集合 U の部分集合 A について，全体集合 U には含まれているが集合 A には含まれない要素の集合を，U に関する A の**補集合** (complementary set) とよんで，

$$A^c \tag{4.40}$$

と表す．肩に complementary の c を付ける．A の補集合 A^c を厳密に定義すれば，

$$A^c = \{x \mid (x \in U) \land \neg(x \in A)\} \tag{4.41}$$

である．例えば，

$$U = \{a, b, c, d, e, f, g\}, \quad A = \{a, b, c\} \tag{4.42}$$

であれば，

$$A^c = \{d, e, f, g\} \tag{4.43}$$

となる．

全体集合 U に関する A の補集合 A^c をベン図で表せば図 4.5 のようになる．

4.2.5　集合演算の基本法則

集合演算 $\cup, \cap, (\ )^c$ に関して，以下のような基本法則が成り立つ．

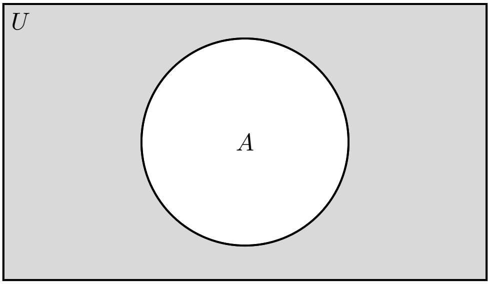

図 4.5　補集合 A^c のベン図

(1) 交換律

$$A \cup B = B \cup A \tag{4.44}$$

$$A \cap B = B \cap A \tag{4.45}$$

(2) 分配律

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \tag{4.46}$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \tag{4.47}$$

(3) 同一律

$$A \cap U = A \tag{4.48}$$

$$A \cup \emptyset = A \tag{4.49}$$

(4) 補元律

$$A \cup A^c = U \tag{4.50}$$

$$A \cap A^c = \emptyset \tag{4.51}$$

(5) ベキ等律

$$A \cup A = A \tag{4.52}$$

$$A \cap A = A \tag{4.53}$$

(6) 有界律

$$A \cup U = U \tag{4.54}$$

$$A \cap \emptyset = \emptyset \tag{4.55}$$

(7) 吸収律

$$A \cup (A \cap B) = A \tag{4.56}$$

$$A \cap (A \cup B) = A \tag{4.57}$$

(8) 結合律

$$A \cup (B \cup C) = (A \cup B) \cup C \tag{4.58}$$

$$A \cap (B \cap C) = (A \cap B) \cap C \tag{4.59}$$

(9) 対合律

$$(A^c)^c = A \tag{4.60}$$

(10) ド・モルガン律

$$(A \cup B)^c = A^c \cap B^c \tag{4.61}$$

$$(A \cap B)^c = A^c \cup B^c \tag{4.62}$$

これらの法則が成立することは，それぞれの集合演算の定義に立ち戻って証明することができる．より直感的には，ベン図を描き左辺と右辺の集合が等しくなることで確認することができる．例えば，(2) 分配律 $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ についてベン図を描くと，左辺右辺とも図 4.6 のようになることが確認できる．

また，それぞれの部分集合の特性関数を利用して，真理表のような表を描く方法もある．任意の $x \in U$ について，部分集合 A, B, C に属しているかどうかのパターンは全部で 8 通りある[*10]．それらそれぞれのパターンについて，集合演算の定義に基づき，x が $A \cup (B \cap C)$ に所属しているかどうかを確認する（表 4.1）．同様のことを $(A \cup B) \cap (A \cup C)$ についても行う（表 4.2）．

結局，$A \cup (B \cap C)$ の所属パターンと，$(A \cup B) \cap (A \cup C)$ の所属パターンが完全に等しいことから，2 つの集合は相等であることが分かる．

[*10] 集合族 $\{A, B, C\}$ のベキ集合 $\mathcal{P}(\{A, B, C\})$ の要素数の $2^3 = 8$ である．

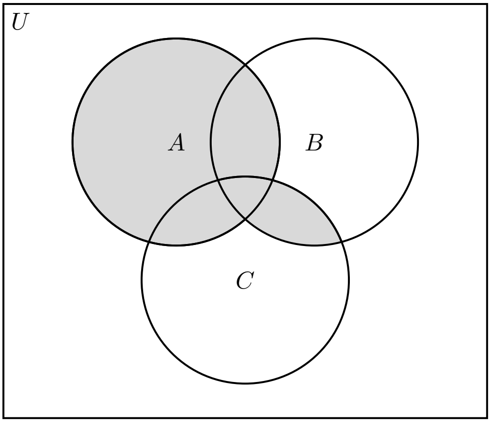

図 4.6　$A \cup (B \cap C)$ ならびに $(A \cup B) \cap (A \cup C)$ のベン図

表 4.1　$A \cup (B \cap C)$ の所属パターン

A	B	C	A	$\cup$	$(B$	$\cap$	$C)$
1	1	1	1	1	1	1	1
1	1	0	1	1	1	0	0
1	0	1	1	1	0	0	1
1	0	0	1	1	0	0	0
0	1	1	0	1	1	1	1
0	1	0	0	0	1	0	0
0	0	1	0	0	0	0	1
0	0	0	0	0	0	0	0

また，勘の良い読者は，ここでのやり方が，命題の同値を確認するやり方と，形式的には全く同じであったことに気づいただろう．この形式的な同一性は，ブール代数の考え方につながる．

表 4.2　$(A \cup B) \cap (A \cup C)$ の所属パターン

A	B	C	$(A$	$\cup$	$B)$	$\cap$	$(A$	$\cup$	$C)$
1	1	1	1	1	1	1	1	1	1
1	1	0	1	1	1	1	1	1	0
1	0	1	1	1	0	1	1	1	1
1	0	0	1	1	0	1	1	1	0
0	1	1	0	1	1	1	0	1	1
0	1	0	0	1	1	0	0	0	0
0	0	1	0	0	0	0	0	1	1
0	0	0	0	0	0	0	0	0	0

練習問題 4.2

集合演算の基本法則の (1) から (11) までを，ベン図を描く方法，所属のパターンを検討する方法，それぞれの方法で確認せよ．

第 5 章

集合 (2)

本章では，命題関数を導入し論理と集合の関係を確認するとともに，関係，関数の概念を導入する．

5.1　命題関数と集合

5.1.1　命題関数

数式

$$3x + 1 = 7 \tag{5.1}$$

は，x が整数をとるとすると，$x = 2$ のときのみが真で，それ以外の数の場合偽となる．このように，1 つの集合 M の任意の要素 x を含む何らかの意味のある文章 $p(x)$ であって，x に M のそれぞれの要素を代入するたびに真偽の判定ができる命題を，集合 M に関する**命題関数** (propositional function) という．また，$p(x)$ を**述語** (predicate)，x を述語の**引数** (argument) とも呼ぶ．引数は「ひきすう」と発音される．引数は有限であればいくつ設定してもよい．例えば，1 つの集合 M の任意の要素 x と y についての命題関数 $p(x, y)$ を考えることもできる．

実数集合 $\mathbb{R}$ に関して，数式

$$(x + 1)^2 = x^2 + 2x + 1 \tag{5.2}$$

は，$\mathbb{R}$ に属するすべての x について真である．

このように，集合 M に属するすべての x について，命題関数 $p(x)$ が真であると

き，この事実を

$$\forall x \in M, \quad p(x) \tag{5.3}$$

と書いて，

「M に属するすべての x に対して，$p(x)$ は真である」

あるいは「M に属する任意の x について，$p(x)$ が成り立つ」などと読むことにする．ここで，記号 $\forall$ は，**全称記号** (universal quantifier) もしくは**全称限量子**といわれるものであり，日常言語でいえば，「すべての (for all)」あるいは「任意の (for any)」に対応している[*1]．この書き方で先の命題を表せば，

$$\forall x \in \mathbb{R}, \quad (x+1)^2 = x^2 + 2x + 1 \tag{5.4}$$

となる．また，2 つの引数をもつ命題関数として，

$$\forall x \in \mathbb{R}, \forall y \in \mathbb{R}, \quad (x+y)^2 = x^2 + 2xy + y^2 \tag{5.5}$$

も成り立つ．こうした式は，**恒等式** (identity) と呼ばれる．

一方，実数集合 $\mathbb{R}$ に関して，数式

$$x^2 - 3x + 2 = 0 \tag{5.6}$$

は，$\mathbb{R}$ に属するすべての x について真とは限らない．実際，$x = 1$，$x = 2$ のときは真であるが，それ以外は偽である．このように M に属するある x に対して，命題関数 $p(x)$ が真であるとき，これを

$$\exists x \in M, \quad p(x) \tag{5.7}$$

と書いて

「M に属するある x に対して，$p(x)$ は真である」

あるいは「$p(x)$ が成り立つような M に属するある x が存在する」などと読む．ここで，記号 $\exists$ は**存在記号** (existential quantifier) もしくは**存在限量子**とよばれる．$\exists$ は

[*1] 英語の All の A をひっくり返したものと覚える．

日常言語でいえば,「ある〜が存在する (there exists)」や「いくつかの (for some)」に対応する[*2]. 全称記号と存在記号をあわせて**限定記号** (quantifier) もしくは**限量子**という.

存在記号を用いて，先の命題関数を表すと

$$\exists x \in \mathbb{R}, \quad x^2 - 3x + 2 = 0 \tag{5.8}$$

となる．このような式を**方程式** (equation) という.

全称記号を含む命題関数

$$\forall x \in M, \quad p(x) \tag{5.9}$$

の否定

$$\neg[\forall x \in M, \quad p(x)] \tag{5.10}$$

を考える．これは

「M に属するすべての x に対して，$p(x)$ は真であるということはない」

つまり,

「M に属するある x に対して，$\neg p(x)$ は真である」

ことを示している．したがって,

$$\neg[\forall x \in M, \quad p(x)] \Longleftrightarrow [\exists x \in M, \quad \neg p(x)] \tag{5.11}$$

である．ここで，$[\exists x \in M, \quad \neg p(x)]$ に該当する x を**反例** (counter example) という.

同様に,

$$\neg[\exists x \in M, \quad p(x)] \Longleftrightarrow [\forall x \in M, \quad \neg p(x)] \tag{5.12}$$

となることが示される.

これら 2 つを述語論理についてのド・モルガン律という.

[*2] 英語の Exist の E を反転させたものと覚える.

練習問題 5.1.1

式 5.12 の左辺と右辺の命題を言葉で説明せよ．

5.1.2　命題関数と集合の関係

ここまでの議論ですでに多くの人が気付いているように，論理と集合は命題関数を通して密接に関連する．

まず，1 つの論理的可能性の集合 U を考え，その任意の要素 x を含む命題関数 $p(x), q(x), r(x), \ldots$ を考える．このとき，ある命題関数 $p(x)$ に対して，$p(x)$ を真とするような要素 x の集合が一意的に定まる．これを命題関数 $p(x)$ の**真理集合**とよんで，$p(x)$ に対応させて P で示す．つまり，

$$P = \{x \mid (x \in U) \wedge p(x)\} \tag{5.13}$$

である．

命題関数 $p(x)$ と $q(x)$ の合成命題の真理集合が以下のようになることは，和集合，積集合，補集合の定義より自明であろう．すなわち，

$p(x) \vee q(x)$ の真理集合は $P \cup Q$
$p(x) \wedge q(x)$ の真理集合は $P \cap Q$
$\neg p(x)$ の真理集合は P^c

さて，いま命題関数 $p(x)$ が恒真命題だとする．恒真命題は，あらゆる論理的可能性について真となることであるから，$p(x)$ の真理集合は全体集合 U と一致する．一方，命題関数 $p(x)$ が恒偽命題だとする．この場合，あらゆる論理的可能性について偽となるのであるから，$p(x)$ の真理集合は空集合 $\emptyset$ と一致する．これらの考察をまとめておこう．

恒真命題の真理集合は U
恒偽命題の真理集合は $\emptyset$

さて，命題関数 $p(x)$ と $q(x)$ が論理的に同値である，つまり $p(x) \Longleftrightarrow q(x)$ となるとき，真理集合上の関係はどのようになっているだろうか．$p(x) \Longleftrightarrow q(x)$ は，双

条件文 $p(x) \leftrightarrow q(x)$ が恒真命題であることを示している．双条件文 $p(x) \leftrightarrow q(x)$ は，$p(x)$ と $q(x)$ がともに真であるとき真となり，$p(x)$ と $q(x)$ がともに偽であるとき真となる．つまり，

$$p(x) \leftrightarrow q(x) \Longleftrightarrow (p(x) \wedge q(x)) \vee (\neg p(x) \wedge \neg q(x)) \tag{5.14}$$

である[*3]．ゆえに，このとき双条件文 $p(x) \leftrightarrow q(x)$ の真理集合は図 5.1 のようになる．しかしながら，双条件文 $p(x) \leftrightarrow q(x)$ が恒真命題であるならば，双条件文 $p(x) \leftrightarrow q(x)$ の真理集合は全体集合 U に一致しなければならない．そして，このことが成り立つのは真理集合上で P と Q が等しいとき，そのときのみである（図 5.2）．

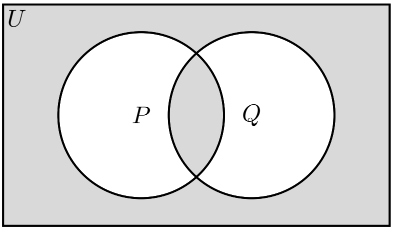

図 5.1　$p(x) \leftrightarrow q(x)$ の真理集合のベン図

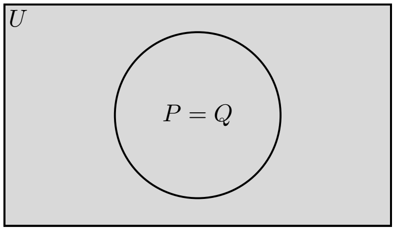

図 5.2　$p(x) \leftrightarrow q(x)$ が恒真命題の場合

次に，命題関数 $p(x)$ が命題関数 $q(x)$ を論理的に含意するとき，つまり $p(x) \Longrightarrow q(x)$ のとき，それぞれの真理集合間の関係がどのようになっているかを検討しよう．$p(x) \Longrightarrow q(x)$ が成立するということは，条件文 $p(x) \rightarrow q(x)$ が恒真命題となっていることを意味している．ところで，

$$p(x) \rightarrow q(x) \Longleftrightarrow \neg p(x) \vee q(x) \tag{5.15}$$

が成立する[*4]．これより，条件文 $p(x) \rightarrow q(x)$ の真理集合は $P^c \cup Q$ であることが分かる．そして，$p(x) \rightarrow q(x)$ が恒真命題であるということは，真理集合上の関係として

$$U = P^c \cup Q \tag{5.16}$$

[*3] 命題論理として，2 つの論理式の同値性を確認してみよう．

[*4] 命題論理として，2 つの論理式の同値性を確認してみよう．

が成立していることと同じことである．そして，この関係が成立するのは P が Q に包含されているとき，つまり $P \subset Q$，そのときだけである（図 5.3, 図 5.4 参照）.

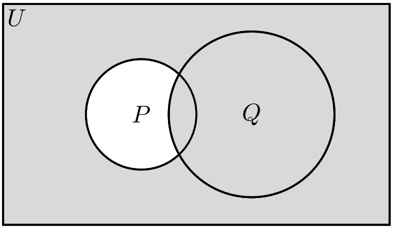

図 5.3　$P^c \cup Q (\neq U)$ のベン図

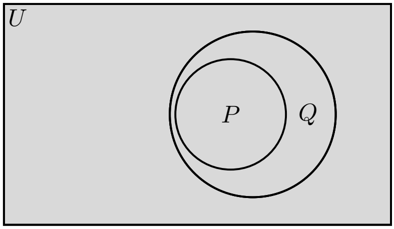

図 5.4　$P^c \cup Q (= U)$ のベン図

このことを標語的に言えば

$p(x) \Longrightarrow q(x)$ のとき，そのときに限り真理集合上で $P \subset Q$ が成立する

ということになる．

また，論理的含意と包含の関係から，3.1.4 節で登場した必要条件，十分条件，必要十分条件について，次のようにまとめることができる．

$P \subset Q$ のとき，そのときに限り，$q(x)$ は $p(x)$ の必要条件である
$P \subset Q$ のとき，そのときに限り，$p(x)$ は $q(x)$ の十分条件である
$P = Q$ のとき，そのときに限り，$p(x)$ と $q(x)$ はお互いの必要十分条件である

最後に，ここまでの考察の結果明らかになった，命題関数についての論理演算と真理集合についての集合演算の対応関係を表 5.1 にまとめておく．

ある集合上に定義された演算のシステムのことを，抽象的に**代数系** (algebraic system) というが，論理演算と集合演算は代数系として形式的に同じである．このような代数系を**ブール代数** (Boolean algebra) と呼ぶ*5.

練習問題 5.1.2

命題関数 $p(x), q(x)$ について，$p(x)$ が $q(x)$ の必要条件，$p(x)$ が $q(x)$ の十分条件，$p(x)$ が $q(x)$ の必要十分条件であるとき，それぞれの場合の真理集合の関係をベン図

*5 ブール代数の特性と応用については，石田 (2017) の 3 章以降を参照.

表 5.1　論理演算と集合演算の対応関係

論理演算	集合演算
$p(x)$ が真	$x \in P$
$p(x)$ が偽	$x \notin P$
$\vee$	$\cup$
$\wedge$	$\cap$
$\neg$	$\{\ \}^c$
恒真命題	U
恒偽命題	$\emptyset$
$\Longleftrightarrow$	$=$
$\Longrightarrow$	$\subset$

で表せ.

5.2　関係

5.2.1　直積集合

2 つの要素の順序の決まった組のことを**順序対** (ordered pair) とよび,

$$(x, y) \tag{5.17}$$

などと表記する．ここでの x を第 1 成分，y を第 2 成分とよぶ．順序対は集合 $\{\ \}$ とは異なり，順序に意味があるので，一般に (x, y) と (y, x) が常に等しいとは限らない．つまり，

$$(x, y) = (x', y') \tag{5.18}$$

とは

$$x = x' \quad \text{かつ} \quad y = y' \tag{5.19}$$

であることを意味する．

2つの集合 A，B の**直積集合** (Cartesian product)$A \times B$ とは，第1成分が集合 A の要素で，かつ第2成分が集合 B の要素であるような順序対のすべてによって構成される集合のことをいう．つまり，

$$A \times B = \{(x, y) \mid (x \in A) \wedge (y \in B)\} \tag{5.20}$$

である．ここで注意しておきたいのは，$A \times B$ はあくまで1つの集合を表す記号であるという点である．A と B がともに有限集合のとき，直積集合の要素数 $|A \times B|$ は

$$|A \times B| = |A| \times |B| \tag{5.21}$$

となる．

例えば，$A = \{1, 0\}$，$B = \{a, b, c\}$ とすると，

$$A \times B = \{(1, a), (1, b), (1, c), (0, a), (0, b), (0, c)\} \tag{5.22}$$
$$B \times A = \{(a, 1), (a, 0), (b, 1), (b, 0), (c, 1), (c, 0)\} \tag{5.23}$$

となる．このとき，$A \times B \neq B \times A$ である．一般に，$A \times B$ と $B \times A$ は常に等しくはならない．一方，要素数については $|A \times B| = |B \times A|$ である．この例では $2 \times 3 = 6$ である．

自分自身との直積は，$A \times A$ の代わりに A^2 とかく．つまり

$$A^2 = A \times A = \{(x, y) \mid (x \in A) \wedge (y \in A)\} \tag{5.24}$$

である．例えば，$A = \{1, 0\}$ とすると

$$A^2 = \{(1, 1), (1, 0), (0, 1), (0, 0)\} \tag{5.25}$$

である．

一般に，n 個の順序づけられた要素の組のことを，$\boldsymbol{n}$ **項組** (n–tuple) といい，

$$(x_1, x_2, \cdots, x_n) \tag{5.26}$$

で表す．また，n 個の集合 $A_1, A_2, \cdots, A_n$ について

$$A_1 \times A_2 \times \cdots \times A_n = \{(x_1, x_2, \cdots, x_n) \mid \forall i \in \{1, 2, \cdots, n\},\ \ x_i \in A_i\} \tag{5.27}$$

を，集合 $A_1, A_2, \cdots, A_n$ の直積集合という．とくに，n 個の A について

$$A^n = A \times A \times \cdots \times A \tag{5.28}$$

を A の n 次の直積集合という．

5.2.2　$\boldsymbol{n}$ 項関係

集合 $A=\{1,2,3\}$ と $B=\{2,3,4\}$ を考える．$x\in A$ と $y\in B$ について，$x\geq y$ となる順序対 (x,y) の集合を考えると

$$\{(2,2),(3,2),(3,3)\} \tag{5.29}$$

となる．この集合は直積集合 $A\times B$ の部分集合となっている．このような，2 つの集合 A,B の直積集合 $A\times B$ の任意の部分集合を，A から B への **2 項関係** (binary relation)，もしくは単に**関係** (relation) とよび，R で表す．つまり，

$$R\subset A\times B \tag{5.30}$$

である．先の例では，R は「x が y と等しいか，y より大きいという関係」を示している．$x\in A$ が $y\in B$ と R という関係にあることを

$$xRy \tag{5.31}$$

とかく．A から A への 2 項関係，つまり A^2 の部分集合をとくに $\boldsymbol{A}$ **上の 2 項関係**，もしくは $\boldsymbol{A}$ **上の関係**という．

一般に，n 個の集合 $A_1,A_2,\cdots,A_n$ からできる直積集合 $A_1\times A_2\times\cdots\times A_n$ の部分集合 R を $\boldsymbol{n}$ **項関係** (n–ary relation) という．とくに，A^n の部分集合 R を $\boldsymbol{A}$ **上の** $\boldsymbol{n}$ **項関係**という．

5.2.3　関係の性質

以下では，A 上の（2 項）関係 R について考える．R について次のような性質を定義する．

(1) **反射的**

$$\forall x\in A,\ \ xRx,\ \text{つまり},\ x\in A\Longrightarrow(x,x)\in R \tag{5.32}$$

(2) **対称的**

$$\begin{aligned}&xRy\Longrightarrow yRx,\\&(x,y)\in R\Longrightarrow(y,x)\in R\end{aligned} \tag{5.33}$$

(3) 反対称的

$$
\begin{aligned}
&xRy \land yRx \Longrightarrow x = y, \\
&(x, y) \in R \land (y, x) \in R \Longrightarrow x = y
\end{aligned}
\tag{5.34}
$$

(4) 推移的

$$
\begin{aligned}
&xRy \land yRz \Longrightarrow xRz, \\
&(x, y) \in R \land (y, z) \in R \Longrightarrow (x, z) \in R
\end{aligned}
\tag{5.35}
$$

(5) 比較可能性

$$
\begin{aligned}
&\forall x, y \in A, \quad xRy \lor yRx, \\
&\forall x, y \in A, \quad (x, y) \in R \lor (y, x) \in R
\end{aligned}
\tag{5.36}
$$

5.2.4　同値関係

関係の性質のなかで，(1) 反射的，(2) 対称的，かつ (4) 推移的であるような A 上の関係 R のことを**同値関係** (equivalence relation) という．R が同値関係であるときに，xRy であることをとくに

$$x \sim_R y \tag{5.37}$$

と表すことがある．

例えば，自然数集合 $\mathbb{N}$ 上に定義される $=$ という関係は同値関係である．というのも，

$$\forall x \in \mathbb{N}, \quad x = x, \quad x = y \Longrightarrow y = x, \quad (x = y) \land (y = z) \Longrightarrow x = z \tag{5.38}$$

がそれぞれ成立する．

また，整数集合 $\mathbb{Z}$ について，$x, y \in \mathbb{Z}$ を正の整数 m で割ったときの余りが等しいとき，x と y は m を法として合同といい

$$x \equiv y \pmod{m} \tag{5.39}$$

とかく．例えば，$1, 4, 7, 10, 13, \cdots$ は 3 を法として合同である．つまり，

$$1 \equiv 4 \equiv 7 \equiv 10 \equiv 13 \equiv \cdots \pmod{3} \tag{5.40}$$

が成り立つ．このような，$\mathbb{Z}$ 上の m を法とする合同関係は同値関係である．

R を A 上の同値関係とするとき，任意の $a \in A$ について

$$[a]_R = \{x \mid x \in A \wedge a \sim_R x\} \tag{5.41}$$

という集合を，a を含む**同値類** (equivalence class) とよび，a を同値類 $[a]_R$ の代表元という．また，

$$A/R = \{[a]_R \mid a \in A\} \tag{5.42}$$

という集合の集合（集合族）を，A の R による**商集合** (quotient set) もしくは**分割** (partition) という．

5.2.5　順序関係

関係の性質のなかで，(1) 反射的，(3) 反対称的，かつ (4) 推移的であるような A 上の関係 R のことを**順序関係** (order relation) という．R が順序関係であるときに，xRy を

$$x \preceq_R y \tag{5.43}$$

と表し，「y は x よりも R の意味で大きい」という．集合 A と A 上の順序関係 $\preceq_R$ の対 $(A, \preceq_R)$ を**順序集合** (ordered set) という．つまり，集合の要素間の一部について順序が定義されているような集合のことを，順序集合というわけである．

さらに，A 上の順序関係 $\preceq_R$ が (5) 比較可能性をも満たすとき，$\preceq_R$ を**全順序関係** (total order relation) といい，対 $(A, \preceq_R)$ を**全順序集合** (totally ordered set) という．

例えば，自然数集合 $\mathbb{N}$ と，$\mathbb{N}$ 上の数の大小関係 $\leq$ からなる対 $(\mathbb{N}, \leq)$ は全順序集合である．

一方，集合 $A = \{a, b, c\}$ のベキ集合

$$\mathcal{P}(A) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\} \tag{5.44}$$

と，$\mathcal{P}(A)$ 上の部分集合関係 $\subset$ による $(\mathcal{P}(A), \subset)$ は順序集合であるが，$\{a\}$ と $\{b\}$ の間などは関係 $\subset$ を定義できないので，全順序集合ではない．

5.3　関数

5.3.1　写像と関数

写像，もしくは関数とは，関係の特別なものである．集合 X から Y への関係 M が

$$\forall x \in X, \quad \exists y \in Y, \quad xMy \tag{5.45}$$

を満たし，かつ $x, x' \in X, y, y' \in Y$ について $xMy, x'My'$ のとき，

$$y \neq y' \Longrightarrow x \neq x' \tag{5.46}$$

を満たすとき, この関係 M を X から Y への**写像** (mapping) もしくは**関数** (function) という[*6]．関数の定義をわかりやすく言い直すと次のようになる．

> 集合 X の任意の要素 x に対して，集合 Y のある要素 y をただ 1 つ対応させる規則 f を，X から Y への関数という．

関数は f, g, φ, ψ などの記号が当てられ，X から Y への関数 f を

$$f : X \to Y \tag{5.47}$$

と表す[*7]．$x \in X$ に対応する Y の要素 y を，x の f による**像** (image) といい，

$$y = f(x) \tag{5.48}$$

と表す．

関数 $f : X \to Y$ において，X を**定義域** (domain) という．定義域全体によってできる像を**値域** (range) といい，$f(X)$ とかく[*8]．$f(X)$ は Y の部分集合である，つまり，$f(X) \subset Y$ である．

---

*6 写像と関数は厳密に言えば異なる概念であるが，本講義では同一の概念として扱う．

*7 ここでの $\to$ は論理演算子の条件法の意味ではない．

*8 厳密には，値域は

$$f(X) = \{y \in Y \mid \exists x \in X, f(x) = y\} \tag{5.49}$$

と定義される．

5.3.2　関数の性質

関数 $f: X \to Y$ についての重要な性質を述べる．

X の任意の要素 x に対して，Y のある要素がそれぞれ一対一対応しているとき，言い換えれば

$$\forall x, x' \in X, \quad x \neq x' \Longrightarrow f(x) \neq f(x') \tag{5.50}$$

を満たすとき，また対偶をとって言い直せば,

$$\forall x, x' \in X, \quad f(x) = f(x') \Longrightarrow x = x' \tag{5.51}$$

を満たすとき，関数 $f: X \to Y$ は**単射** (injection) であるという．

また，Y の任意の要素に必ず対応する X の要素があるとき，言い換えれば

$$\forall y \in Y, \quad \exists x \in X, \quad f(x) = y \tag{5.52}$$

が成り立つとき，さらに言い換えれば

$$f(X) = Y \tag{5.53}$$

が成立するとき，関数 $f: X \to Y$ は**全射** (surjection) であるという．

最後に，関数 $f: X \to Y$ が全射でかつ単射であるとき，**全単射** (bijection) であるという．f が全単射であるとき,

> $\forall x \in X$ について，$y = f(x)$ となるただ 1 つの $y \in Y$ が存在し，かつ $\forall y \in Y$ について，$y = f(x)$ となるただ 1 つの $x \in X$ が存在する

ことが成り立つ[*9]．このことを言い換えれば，X の各要素と Y の各要素が完全に一対一対応になっているということである．

図 5.5 は単射，全射，全単射の概念図である．

---

*9

証明．関数の基本性質より，$y = f(x)$ かつ $y' = f(x')$ とすると,

$$f(x) \neq f(x') \Longrightarrow x \neq x'. \tag{5.54}$$

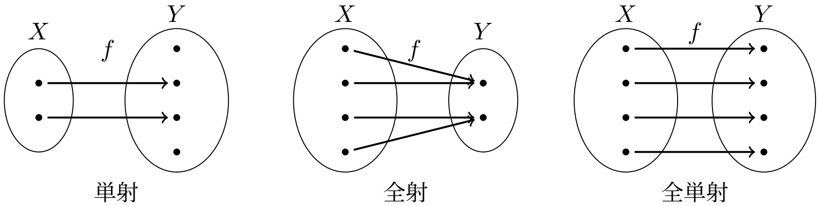

図 5.5　単射，全射，全単射の概念図

ここで，実数集合 $\mathbb{R}$ から $\mathbb{R}$ への関数を考えよう．関数 $f:\mathbb{R}\to\mathbb{R}$ を $f(x)=2^x$ とする．この関数は $\mathbb{R}$ から $\mathbb{R}$ への単射であるが，$f(x)>0$ であるので，全射ではない (図 5.6)[*10].

一方，関数 $f:\mathbb{R}\to\mathbb{R}$ を $f(x)=x$ とする．この関数は全単射である（図 5.6）．この関数のように，一般に集合 A に対し，

$$\forall x\in A, f(x)=x \tag{5.57}$$

を満たす関数 $f:A\to A$ を，A 上の**恒等関数** (identity function) と呼び I_A で表す．

5.3.3　合成関数，逆関数

2 つの関数 $f:X\to Y$，$g:Y\to Z$ が与えられているとき，X の任意の要素 x について，

$$(g\circ f)(x)=g(f(x)) \tag{5.58}$$

で定義される関数 $g\circ f:X\to Z$ を f と g の**合成関数** (composite) という．合成関数 $g\circ f$ は，x をまず f によって $f(x)\in Y$ に写し，さらに $f(x)$ を g によって

いま全射の定義 $\forall y\in Y,\exists x\in X, y=f(x)$ より，

$$\forall y,y'\in Y,\quad (y=f(x))\wedge(y'=f(x'))\wedge(f(x)\neq f(x')\Longrightarrow x\neq x') \tag{5.55}$$

となる．一方，単射の定義より，

$$\forall x,x'\in X,\quad (y=f(x))\wedge(y'=f(x'))\wedge(x\neq x'\Longrightarrow f(x)\neq f(x')). \tag{5.56}$$

となる． □

[*10] ただし，$f(x)=2^x$ を $f:\mathbb{R}\to(0,\infty)$ とすると全単射である．

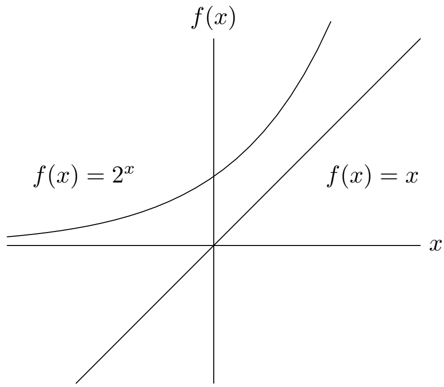

図 5.6　$f(x) = 2^x, f(x) = x$ のグラフ

$g(f(x)) \in Z$ に写す関数である．

一方，2 つの関数 $f : X \to Y$，$g : Y \to X$ に関して，

$$(g \circ f = I_X) \wedge (f \circ g = I_Y) \tag{5.59}$$

が成り立つとき，すなわち，

$$(\forall x \in X,\ \ g(f(x)) = x) \wedge (\forall y \in Y,\ \ f(g(y)) = y) \tag{5.60}$$

が成立するとき，関数 g を f の**逆関数** (inverse function) といい，$f^{-1} : Y \to X$ と表す．関数 f の逆関数 f^{-1} はどのような関数についても存在するわけではなく，f が全単射であるとき，そのときのみに存在する．つまり，関数 $f : X \to Y$ に関して，f が全単射であることと，f の逆関数 $f^{-1} : Y \to X$ が存在することは同値である．また，そのような逆関数はただ 1 つ存在する（証明は A.1 参照）．

例えば，関数 $f : \mathbb{R} \to \mathbb{R}$，$f(x) = 2x$ は全単射であり，ゆえに逆関数が存在する．逆関数 $f^{-1} : \mathbb{R} \to \mathbb{R}$ は具体的には，$f^{-1}(y) = (1/2)y$ である．というのも，

$$f^{-1}(f(x)) = (1/2)(2x) = x,\ \ f(f^{-1}(y)) = 2((1/2)y) = y \tag{5.61}$$

が成り立つからである．

第 6 章

確率 (1)

本章と次章では，データ分析・統計学を学ぶときに欠かせない確率論の基礎を学ぶ．一言で言うと，確率論とは不確実性を取り扱うための数学モデルである．

6.1　標本空間と事象

6.1.1　標本空間

確率 (probability) とは，ある出来事が起こる「確からしさ」を数量的に定めたものである．確率論の基礎になるのが集合を使った起こりうる出来事の表現である．

コインやサイコロを投げる，あるいは壷の中にいくつか入っている小玉の一つを取り出す，あるいは東京ドームをぶん回して中の人 1 人を取り出すといった操作のことを確率論では**試行** (trial) という．ある試行を行ったときに生じるすべての可能な結果の集合のことを**標本空間** (sample space) あるいは全事象といい，Ω で表す．また，標本空間の任意の要素を**標本点** (sample point) といって ω で表す．定義から $\omega \in \Omega$ である．結果がある区間の実数をとりうる場合，標本空間はその区間のすべての点の集合になる．図 6.1 は標本空間の模式図である．

■例（コイン投げ試行の標本空間）　コインを投げたときの裏表の標本空間は

$$\Omega = \{\text{裏}, \text{表}\}$$

である．コインを 2 枚同時に投げたときの標本空間は

$$\Omega = \{(\text{裏}, \text{裏}), (\text{裏}, \text{表}), (\text{表}, \text{裏}), (\text{表}, \text{表})\}$$

である．

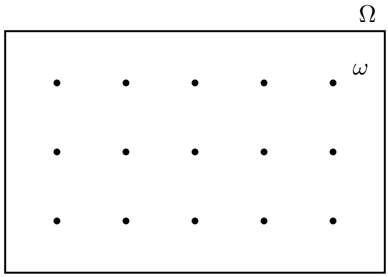

図 6.1　標本空間

■例（サイコロ投げ試行の標本空間）　サイコロを投げたときの目の標本空間は

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

である．

6.1.2　事象

標本空間の部分集合のことを**事象** (event) という．事象は確率的に起こりうる出来事を意味する．標本点をただ 1 つ含む事象のことを**根元事象** (elementary event) とよび，2 つ以上の標本点を含む事象のことを複合事象という．標本点を含まない事象のことを**空事象** (empty event) といい，$\emptyset$ で表す．ある事象を A で表すと $\emptyset \subset A \subset \Omega$ である．

■例（コイン投げ試行の事象）　コイン投げ試行において考えうる事象は

$$\{\,裏, 表\,\}, \qquad \{\,裏\,\}, \qquad \{\,表\,\}, \qquad \emptyset$$

の 4 つであり，それぞれ「裏か表が出る」,「裏が出る」,「表が出る」,「裏も表も出ない」である．このうち $\{\,裏\,\}, \{\,表\,\}$ は根元事象であり，$\{\,裏, 表\,\}$ は複合事象，$\emptyset$ は空事象である．

■例（サイコロ投げ試行の事象）　サイコロ投げ試行において考えうる事象は全部で $2^6 = 64$ 通りある．例えば，事象 $\{1, 2, 3\}$ は「3 以下の目が出る」という事象であり，$\{2, 4, 6\}$ は「偶数の目が出る」という事象である．

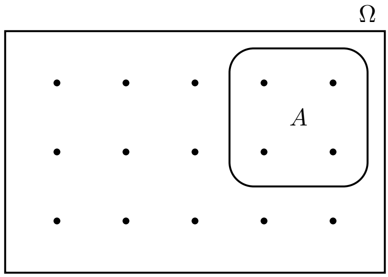

図 6.2　事象の例

6.1.3　事象の演算

ここで，サイコロ投げ試行の事象として

$$A = \{1, 2, 3\}, \quad B = \{2, 4, 6\}$$

を例とする．「A と B のうち，少なくとも一方が起こる事象」のことを A と B の**和事象** (union of events) とよび $A \cup B$ で表す．演算としては集合演算の和集合と同じであり，$A \cup B$ の場合

$$A \cup B = \{1, 2, 3, 4, 6\}$$

となる．「A と B が同時に起こる事象」を A と B の**積事象** (intersection of events) とよび $A \cap B$ で表す．演算としては集合演算の積集合と同じであり，$A \cap B$ の場合

$$A \cap B = \{2\}$$

となる．$A \cap B = \emptyset$ のとき A と B は**互いに排反** (mutually exclusive) であるという．最後に,「A が起こらないという事象」を A の**補事象** (complementary event) とよび，A^c で表す．集合演算における補集合と同じであり，例の場合

$$A^c = \{4, 5, 6\}, \quad B^c = \{1, 3, 5\}$$

である．

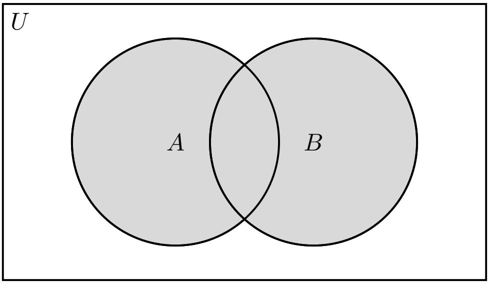

図 6.3　和事象 $A \cup B$ のベン図

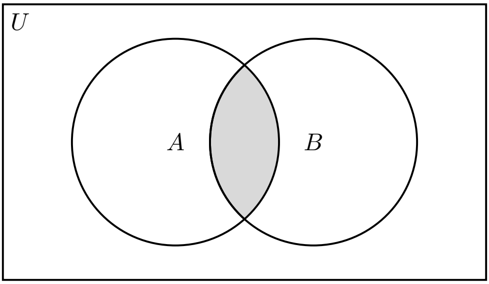

図 6.4　積事象 $A \cap B$ のベン図

それぞれの事象演算を**ベン図** (Venn diagram) で表せば図 6.3–6.5 のようになる．灰色部分が該当する演算の結果の集合を示す．

なお，集合演算 $\cup, \cap, (\ )^c$ に関して成り立つ基本定理は 4 章ですでに導入した．

6.2　確率の定義

確率 (probability) とは，事象の起こりやすさを定量的に示す測度 (measurement) であり，事象 A の確率を $P(A)$ と表す．つまり，確率とは集合を引数とする集合関数の一種である．

確率を具体的にどのように定義するかということは，これまで一つの論争点であった．ここでは，ラプラスによる古典的定義，頻度による定義，そしてコルモゴロフに

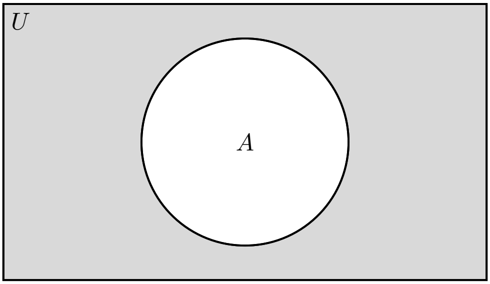

図 6.5　補事象 A^c のベン図

よる公理主義的定義を紹介する．本講義では，直感的には古典的定義と頻度による定義，数学的には公理主義的定義に立って議論を行う．

定義 6.1 (**確率の古典的定義**)**.** ある試行について，標本空間の要素数が n であり，それぞれの根元事象は同程度に確からしく起こるとする．ある事象 A についてその要素数が r のとき，A の確率 $P(A)$ を

$$P(A) = \frac{r}{n} \tag{6.1}$$

と定義する．

■例（サイコロ投げの確率）　ゆがみのないサイコロのそれぞれの目（根元事象）は，同程度に確からしく起こると考えてよい．このとき，確率の古典的定義によれば，偶数の目が出る確率は $3/6 = 0.5$ である．

この定義には直感的には分かりやすいが，根元事象が同程度に確からしい場合にしか定義できないという問題がある．

定義 6.2 (**確率の頻度による定義**)**.** ある試行を n 回行ったとき，ある事象 A が r 回起こったとする．n を限りなく大きくしていくとき，r/n が一定の値 p に近づくならば，A の確率を

$$P(A) = p = \lim_{n\to\infty} \frac{r}{n} \tag{6.2}$$

と定義する[*1].

■**例（コイン投げの確率）**　ゆがみのないコインを投げて表が出るかどうかをチェックする．試行回数を n，表が出る回数を r とすると，試行回数が多くなればなるほど r/n は $1/2$ に近づくことが確認できる．

n の極限は数学的には定義しうるものの，経験的には観察不可能であるという問題がある．

コルモゴロフによる確率の公理主義的定義は，確率の経験的内容に関する議論を留保したままで，数学的取り扱いについての形式的側面を公理化したものであり，この定義の登場によってより体系的に確率を数学的に取り扱うことが可能になった．

公理 6.1 (**確率の公理**)**.** 標本空間 Ω における各事象 A について，次の *3* つの条件を満たす実数 $P(A)$ が存在するとき，$P(A)$ を事象 A が起こる確率という．

(1) すべての事象 A について，$0 \leq P(A) \leq 1$.

(2) $P(\Omega) = 1$.

(3) $A_1, A_2, A_3, \cdots$ が互いに排反のとき

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = P(A_1) + P(A_2) + P(A_3) + \cdots .$$

6.3　確率の性質

6.3.1　確率の基本定理

公理から直ちに導かれる確率の基本定理を確認する．

定理 6.1 (**確率の基本定理**)**.**

(1) $P(\emptyset) = 0$.

(2) $P(A^c) = 1 - P(A)$.

[*1] $\lim_{n\to\infty} f(n)$ は極限を取るという操作を表す．数学的には厳密な定義が必要であるが，ここでは，n を限りなく大きくしていったときに，$f(n)$ が限りなく近づいていく値，とラフに理解すればよい．

(3) Ω が n 個の根元事象 $E_i(i=1,2,\cdots,n)$ から成り立っているならば，

$$\sum_{i=1}^{n} P(E_i) = 1.$$

(4) $A \subset B \Longrightarrow P(A) \leq P(B)$.

証明. (1) 恒等律もしくは支配律より $\emptyset \cap \Omega = \emptyset$, $\emptyset \cup \Omega = \Omega$ である．ゆえに，

$$P(\emptyset \cup \Omega) = P(\Omega) = P(\emptyset) + P(\Omega)$$

である．$P(\Omega) = 1$ なので，$P(\emptyset) = 0$ である．
(2) 矛盾律より $A \cap A^c = \emptyset$，また排中律より $A \cup A^c = \Omega$ なので

$$P(A \cup A^c) = P(\Omega) = P(A) + P(A^c)$$

より，$P(A^c) = 1 - P(A)$ を得る．
(3) 根元事象は定義より互いに排反であるので

$$P(E_1 \cup E_2 \cup \cdots \cup E_n) = P(E_1) + P(E_2) + \cdots + P(E_n)$$

である．前提条件より $\Omega = E_1 \cup E_2 \cup \cdots \cup E_n$ なので，定理が成立する．
(4) 集合的に A が B に含まれる，つまり A が B の部分集合であることを $A \subset B$ と表す．$A \subset B$ であれば，$B = A \cup (A^c \cap B)$ かつ $A \cap (A^c \cap B) = \emptyset$ となる（ベン図を描いて確認せよ）．ゆえに，

$$P(B) = P(A) + P(A^c \cap B)$$

となる．確率の公理 (1) より，$P(A^c \cap B) \geq 0$ なので，定理が成立する． □

次の定理は**加法定理**として知られている．

定理 6.2 (**加法定理**)**.** 任意の事象 A, B について

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{6.3}$$

証明. ベン図を検討するか，もしくは集合演算基本公式を適用することで

$$A = (A \cap B^c) \cup (A \cap B)$$
$$B = (A^c \cap B) \cup (A \cap B)$$
$$A \cup B = (A \cap B^c) \cup (A^c \cap B) \cup (A \cap B)$$

かつ，$(A \cap B^c), (A^c \cap B), (A \cap B)$ が互いに排反であることが分かる．よって，

$$\begin{aligned} P(A) &= P(A \cap B^c) + P(A \cap B) \\ P(B) &= P(A^c \cap B) + P(A \cap B) \\ P(A \cup B) &= P(A \cap B^c) + P(A^c \cap B) + P(A \cap B) \end{aligned}$$

である．ここから，

$$\begin{aligned} P(A \cup B) &= P(A \cap B^c) + P(A^c \cap B) + P(A \cap B) \\ &= P(A \cap B^c) + P(A \cap B) + P(A^c \cap B) + P(A \cap B) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

を得る． □

A, B が互いに排反のときは $A \cap B = \emptyset$ なので

$$P(A \cup B) = P(A) + P(B)$$

となる．**事象が相互排反のときは，事象の和の確率は確率の和である．**

6.3.2　古典的定義の導出

ある試行について，標本空間の要素数が n であり，それぞれの根元事象 E_i は同程度に確からしく起こるとする．

このとき，$P(E_1) = P(E_2) = \cdots = P(E_n) = \alpha$ であり，定理 6.1 より，

$$\sum_{i=1}^{n} P(E_i) = n\alpha \Longleftrightarrow \alpha = \frac{1}{n}$$

となる．つまり，$P(E_i) = 1/n$ である．

さらに，根元事象の和で構成される複合事象 A について，その要素数が r のとき，A の確率は

$$P(A) = r\alpha = \frac{r}{n}$$

である．これで確率の古典的定義を公理論的に導出することができた．

■例（コイン投げの確率）　　コイン投げにおける表と裏の出現が同様に確からしいとき，公理より以下のように確率を与えることができる．

$$P(\{\text{ 裏}, \text{表 }\}) = 1,\quad P(\{\text{ 裏 }\}) = P(\{\text{ 表 }\}) = \frac{1}{2},\quad P(\emptyset) = 0.$$

■例（サイコロ投げの確率）　　サイコロの目の出現が同様に確からしいとき，公理よりあたられるサイコロ投げの事象の確率の一部である．

$$P(\{1\}) = P(\{2\}) = \cdots = P(\{6\}) = \frac{1}{6},\quad P(\{1,2,3\}) = P(\{4,5,6\}) = \frac{1}{2}.$$

練習問題 6.3

サイコロの目の出現が同様に確からしいとき，以下の確率を求めよ．

(1) 1 から 6 のいずれかの目が出る確率
(2) 1 から 6 のいずれの目も出ない確率
(3) 3 以下の目か 6 のいずれかが出る確率
(4) 4 以上の目か奇数のいずれかが出る確率

6.4　条件付き確率と独立性

6.4.1　条件付き確率

ある事象 A が起こったときに，さらに事象 B が起こる確率を知りたいとする．このような確率を A が起きたという条件のもとでの B の**条件付き確率** (conditional probability) といい，$P(B|A)$ で表す．

定義 6.3 (**条件付き確率**)**.** 事象 A が起きたという条件のもとでの事象 B の条件付き確率は

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \tag{6.4}$$

で定義される．ただし，事象 A は $P(A) > 0$ であるとする．

この定義は，もともと標本空間 Ω 上で割り当てられた確率 $P(A)$ と $P(A \cap B)$ の比を維持したままで，A を縮小された標本空間と見なして事象 B の確率を再割り当てするという操作を意味している[*2]．

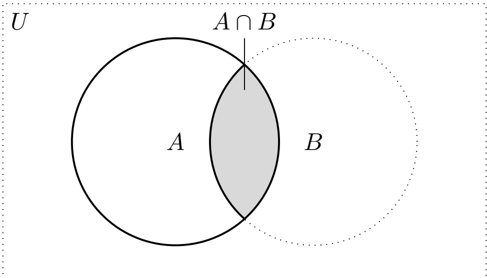

図 6.6　条件付き確率 $P(B|A)$ についてのベン図 (A を標本空間として $A \cap B$ に割り当てた確率測度)

■例（トランプの絵札）　トランプをよく切って 1 枚を引いたときそれがハートであったということがわかっている．このときさらにそのカードが絵札（11, 12, 13）である確率が知りたいとする．「ハート」という事象を A,「絵札」という事象を B として条件付き確率を求める．確率の古典的定義より

$$P(A) = 13/53, \quad P(B) = 12/53, \quad P(A \cap B) = 3/53$$

である．ゆえに条件付き確率 $P(B|A)$ は

$$P(B|A) = \frac{3/53}{13/53} = \frac{3}{13}$$

となるが，これは A を標本空間として絵札が出る古典的確率を直接求めた値に等しい.

条件 B のもとでの A の条件付き確率も同様に求められ,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{6.5}$$

[*2] 詳しくは，P・G・ホエール，1978,『入門数理統計学』培風社，pp. 14-15 を参照のこと．

となる．式 (6.4)，(6.5) の分母を払うと，

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B) \tag{6.6}$$

を得る．この式をとくに**乗法定理**とよぶことがある．

6.4.2　独立

条件付き確率において，事象 A が起こったことが，B が起きる確率に何の影響も与えないことがある．このとき，

$$P(B|A) = P(B) \tag{6.7}$$

となる．これを式 (6.4) に代入すると

$$P(A \cap B) = P(A)P(B) \tag{6.8}$$

を得る．式 (6.7)，(6.8) が成り立っているとき，事象 A と B は**独立** (independent) であるという．

独立の概念を n 個の事象に敷衍することができる．事象 $A_1, A_2, \cdots, A_n$ について，

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2)\cdots P(A_n) \tag{6.9}$$

が成り立つとき，事象 $A_1, A_2, \cdots, A_n$ は独立であるという．

■例（トランプの絵札）　さきほどのトランプの絵札 B の条件付き確率の例では，

$$P(B|A) = 3/13 \neq 12/53 = P(B)$$

であるので，統計的に独立ではない．これは「ジョーカー」の 1 枚があるためであって，もとの B の確率よりも，ハート A を条件としたときの条件付き確率 $P(B|A)$ のほうが，ジョーカーを排除したために絵札である確率が若干上昇する．ジョーカーを除いた 52 枚のトランプで同様の試行を行う場合，

$$P(A) = 13/52, \quad P(B) = 12/52, \quad P(A \cap B) = 3/52$$

であり，$P(A \cap B) = P(A)P(B)$ が成り立つので，この場合「絵札が出ること」と「ハートが出ること」は独立の事象である．

■例（コインを繰り返し投げる）　ゆがみのないコインを 10 回投げたときに，10 回連続で表が出る確率はどれくらいだろうか．t 回目に表が出るという事象を A_t と表す．確率を求めたい事象は $A_1 \cap A_2 \cap \cdots \cap A_{10}$ である．コインを 10 回投げたときの表裏の組み合わせのパターンは $2^{10} = 1024$ パターンある．どの組み合わせのパターンになるかは「同様に確からしい」と考えられる場合，確率の古典的定義から，

$$P(A_1 \cap A_2 \cap \cdots \cap A_{10}) = \frac{1}{2^{10}}$$

である．一方，事象 A_t の確率 $P(A_t)$ はそれぞれ $1/2$ なので，

$$P(A_1)P(A_2)\cdots P(A_{10}) = \left(\frac{1}{2}\right)^{10}$$

なので，

$$P(A_1 \cap A_2 \cap \cdots \cap A_{10}) = P(A_1)P(A_2)\cdots P(A_{10})$$

であり，それぞれのコイン投げの結果は独立である．1 回 1 回の結果が独立となる試行のことを，独立試行ともいう．

練習問題 6.4

　サイコロの目の出現が同様に確からしく，コイン投げにおける表と裏の出現が同様に確からしいとき，以下の確率を求めよ．

(1) サイコロ投げにおいて「奇数が出る」という条件のもとで 1 の目が出る確率
(2) コイン投げとサイコロ投げを独立に行ったとき，表が出たことを条件としてサイコロの 6 の目が出る確率
(3) サイコロを独立に 3 回投げたときに 1 の目が 3 回とも出る確率
(4) サイコロを独立に 3 回投げたときに目の合計が 4 になる確率

6.5　最低限知っておくべきこと

　以上の確率論の基礎知識は，今後推測統計を学んでいく上で最低限知っておくべき知識である．そのなかでも，推測統計に関して頻出する知識は以下の 3 点である．

1. 排反な事象のいずれかが起こる確率は，それぞれの**確率の和** (確率の公理 3)
2. 独立な事象が同時に起こる確率は，それぞれの**確率の積** (独立の定義)
3. 復元無作為抽出は独立試行

第 7 章

確率 (2)

引き続き，推測統計学に必須の確率論の知識を導入する.

7.1　確率変数

標本空間のそれぞれの根元事象に適当な数値を対応させた変数 X を考える．このとき，この変数 X がある特定の値 x とる確率が対応する根元事象の確率として定まる．このような変数を**確率変数** (random variable) という．X は変数であり試行を行う前は値は定まっていない．一方，x は試行後に定まった値を表すのがこの分野の慣例である[*1].

確率変数 X がとびとびの値 $x_1, x_2, \cdots, x_k, \cdots$ しかとらないとき，X を**離散確率変数** (descrete random variable) という．以下，とりうる値は有限個であると仮定して話を進めよう．値が x_i をとるときの確率を

$$P(X = x_i) = p_i \qquad (i = 1, 2, \cdots, k)$$

というように表す.

[*1] 確率変数 X は，試行を行う前，あるいは試行結果が明らかになる前の不確実な変数であり，値が定まっていないものである．なかなかイメージしづらいかもしれないが，スロットやルーレットが回っている状態を思い浮かべてもよいだろう．個人的には X は，プルプルと小刻みに動いているイメージである．これを具現化するために，いちいち $(((X)))$ と表記してもよいが，めちゃめちゃ読みにくくなると思う．諸君も確率論を極めた暁には，文中の確率変数が独りでにプルプルし始めるであろう.

■**例（サイコロの目を値としてとる確率変数）**　サイコロの目をとる確率変数 X を考えると，それぞれの確率は

$$P(X=i)=\frac{1}{6} \qquad (i=1,2,3,4,5,6)$$

である．

7.2　離散型確率分布

確率 p_i は，値 x_i の関数であるから，これを明示化して

$$P(X=x_i)=f(x_i) \qquad (i=1,2,\cdots,k) \tag{7.1}$$

とも表す．このとき $f(x_i)$ は**確率関数** (probability function) といわれる．確率の性質から確率関数は

$$f(x_i)\geq 0 \qquad (i=1,2,\cdots,k) \tag{7.2}$$

$$\sum_{i=1}^{k} f(x_i)=f(x_1)+f(x_2)+\cdots+f(x_k)=1 \tag{7.3}$$

を満たす（総和についての解説は次章）．この確率関数によって定まる x と $f(x)$ との関係を**離散型確率分布** (probability distribution of discrete type) とよぶ．

■**例（サイコロの目の確率分布）**　表 7.1 と図 7.1 はサイコロの目の確率分布，表 7.2 と図 7.2 は 2 個のサイコロの目の和の確率分布を示している．

表 7.1　サイコロの目の確率分布

x	1	2	3	4	5	6
$f(x)$	1/6	1/6	1/6	1/6	1/6	1/6

7.3　連続型確率分布

確率変数 X が連続の値をとるとき，X を**連続確率変数** (continuous random variable) という．連続確率変数の場合，1 つ 1 つの値ごとではなく，値のとる範囲に

表 7.2　2 個のサイコロの目の和の確率分布

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$

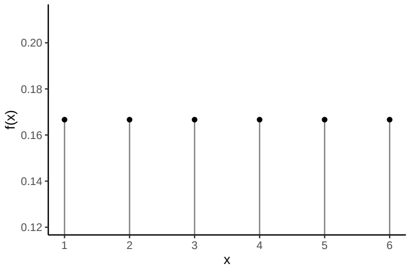

図 7.1　サイコロの目の確率分布

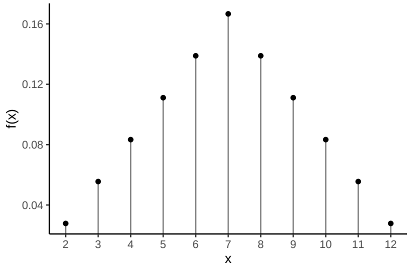

図 7.2　2 個のサイコロの目の和の確率分布

応じて確率が決まると考える．確率変数 X がある範囲の値をとる確率が，関数 $f(x)$ によって

$$P(a \leq X \leq b) = \int_a^b f(x)dx \tag{7.4}$$

と表され，かつ

$$\text{すべての } x \text{ について}, f(x) \geq 0 \tag{7.5}$$

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{7.6}$$

を満たすとき，X は**連続型確率分布** (probability distribution of continuous type) をもつという．このとき関数 $f(x)$ を，**確率密度関数** (probability density function; PDF)，あるいは単に密度関数という．

定義より，確率変数がただ 1 つの値 a をとる確率は

$$P(a \leq X \leq a) = \int_a^a f(x)dx = 0$$

である．連続型の場合は，値の範囲にのみ意味があることに注意すること．

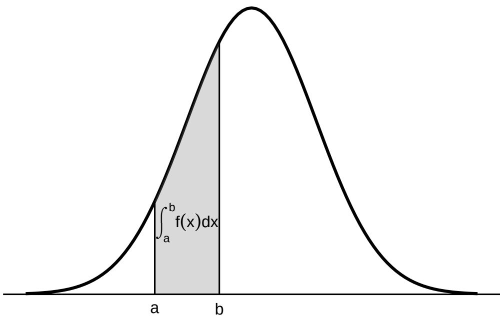

図 7.3　確率密度関数の例

さて，ここでいきなり積分記号 $\int$ が出てきて，戸惑った人もいるかもしれない．主に確率論で用いる積分については 10 章で解説する．ここでは最低限，定積分 $\int_a^b f(x)dx$ は，離散型確率分布のある範囲の確率を和 $\sum_i f(x_i)$ で表すことに対応しており，図形的には「連続関数 $f(x)$ と x 軸（$y=0$ の水平線）との間のにおける a から b までの範囲の面積」を示すことを理解しておけばよい（図 7.3 参照）．

7.4　期待値

7.4.1　期待値の定義

記述統計のところでデータに対して求めた平均や分散を確率変数について求める一般的な操作を**期待値** (expectation) という．

定義 7.1. 確率変数 X の任意の関数を $\varphi(X)$ とするとき，$\varphi(X)$ の**期待値**を離散確率変数の場合

$$\begin{aligned} E[\varphi(X)] &= \varphi(x_1)f(x_1) + \varphi(x_2)f(x_2) + \cdots + \varphi(x_k)f(x_k) \\ &= \sum_{i=1}^{k} \varphi(x_i)f(x_i) \end{aligned} \tag{7.7}$$

連続確率変数の場合

$$E[\varphi(X)] = \int_{-\infty}^{\infty} \varphi(x) f(x) dx \tag{7.8}$$

と定義する．

7.4.2　平均と分散

$\varphi(X) = X$ とするとき，確率変数 X の**平均** (mean) が定義される．離散確率変数の場合

$$\mu = E(X) = \sum_{i=1}^{k} x_i f(x_i), \tag{7.9}$$

連続確率変数の場合，

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx \tag{7.10}$$

となる．この確率変数の平均を「期待値」と呼ぶケースが多い．

$\varphi(X) = (X - \mu)^2$ とするとき，確率変数 X の**分散** (variance) が定義される．離散確率変数については

$$\begin{aligned} \sigma^2 = V(X) &= E[(X - \mu)^2] \\ &= (x_1 - \mu)^2 f(x_1) + (x_2 - \mu)^2 f(x_2) + \cdots + (x_k - \mu)^2 f(x_k) \\ &= \sum_{i=1}^{k} (x_i - \mu)^2 f(x_i) \end{aligned} \tag{7.11}$$

また，連続確率変数については

$$\begin{aligned} \sigma^2 = V(X) &= E[(X - \mu)^2] \\ &= \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \end{aligned} \tag{7.12}$$

と定義される．分散のルートをとったものを確率変数 X の**標準偏差** (standard deviation) といい，σ で表す．すなわち，$\sigma = \sqrt{\sigma^2}$ である．

期待値の性質については，次章の総和を参照のこと．

7.5　二項分布

確率分布には様々なタイプのものがあるが，以下では具体例として，もっとも基本的な離散型確率分布である二項分布を導入する．

7.5.1　ベルヌーイ試行とベルヌーイ分布

結果が 2 種類しかない試行を**ベルヌーイ試行** (Bernoulli trials) という．多くの場合，2 つの結果は「成功（S)」と「失敗（F)」と名付けられる．ゆえにベルヌーイ試行の標本空間は

$$\Omega = \{S, F\}$$

である．「コイン投げ試行」などは典型的なベルヌーイ試行である．ベルヌーイ確率変数 X を，事象 $\{S\}$ に対して 1，$\{F\}$ に対して 0 と定める．このとき，ベルヌーイ試行における確率関数は,

$$P(X = x) = f(x) = p^x(1-p)^{1-x}, \qquad x = 0, 1 \tag{7.13}$$

と定義される．定義より

$$f(x) = \begin{cases} p & (x = 1) \\ 1 - p & (x = 0) \end{cases}$$

である．p はベルヌーイ試行の成功確率であり，$1 - p$ は失敗確率である．このように定義される確率分布を**ベルヌーイ分布**という．

■ベルヌーイ分布の平均

$$E(X) = 1 \times p + 0 \times (1 - p) = p$$

■ベルヌーイ分布の分散

$$V(X) = (1 - p)^2 p + (0 - p)^2 (1 - p) = p(1 - p)$$

7.5.2　順列の数と組合せの数

二項分布の導入の前に，順列の数と組合せの数について定義しておく．

5 人が参加するコンテストにおいて，3 位までの順位を付けることを考える．このとき，可能な順位付けは，まず，5 人のうち 1 人が 1 位になるのは 5 通り，残り 4 人のうち 1 人が 2 位になるのは 4 通り，残り 3 人のうち 1 人が 3 位になるのは 3 通り，なので全部で

$$5 \times 4 \times 3 = 60 \text{ 通り}$$

になる．

ここで自然数 n に対してその**階乗** (factorial) を

$$n! = \prod_{k=1}^{n} k = n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1 \tag{7.14}$$

と定義しておく．$n = 0$ のとき，$0! = 1$ と定義する．

さて，一般に，n 個のものから r 個を取り出して並べるときの**順列の数** (permutation) は，

$$\begin{aligned} {}_nP_r &= n \times (n-1) \times (n-2) \times \cdots \times (n-(r-1)) \\ &= \frac{n!}{(n-r)!} \end{aligned}$$

となる．

順位を付けずに組合せだけを考える場合，例えば 5 人のうち 3 人だけを合格とするような場合，3 人の人が選ばれて順位が付けられたときの順列の数は $3 \times 2 \times 1 = 6$ なので，これを先の順列の数から割れば組合せの数がわかる．つまり

$$\frac{5 \times 4 \times 3}{3 \times 2 \times 1} = 10 \text{ 通り}$$

になる．

一般に，n 個のものから r 個を取り出すときの**組合せの数** (combination) は，

$$
\begin{aligned}
{}_nC_r = \begin{pmatrix} n \\ r \end{pmatrix} &= \frac{{}_nP_r}{r!} \\
&= \frac{n!}{r!(n-r)!}
\end{aligned}
$$

となる．

7.5.3　二項分布

同じ成功確率 p をもち，かつ独立なベルヌーイ試行 (つまり，各試行の成功・失敗事象が独立) を n 回繰り返す．各回のベルヌーイ試行の確率変数を X_i $(i = 1, 2, \cdots, n)$ と表記する．先ほどと同じく X_i は 0,1 の値をとる．このとき，

$$
X = X_1 + X_2 + \cdots + X_n
$$

とすると，X もまた確率変数であり，値は 0 から n までの整数値をとり，「n 回の試行のうち成功した回数」を表す．

例えば，1 回の試行の成功確率が 0.3 のとき，5 回試行を行って 3 回成功する確率を求めよう．成功失敗の生起順序は関係ないのであるから，5 回試行を行って 3 回成功する場合の数は組合せの数として求まり，

$$
{}_5C_3 = \frac{5!}{3!2!} = 10
$$

である．5 回試行を行って 3 回成功する確率は，独立であることに注意すると，

$$
0.3 \times 0.3 \times 0.3 \times 0.7 \times 0.7 = 0.013
$$

である．この確率を場合の数だけ足し合わせると $10 \times 0.013 = 0.13$ という確率を得る．

一般に，成功確率が p であるベルヌーイ試行を n 回行うとき，x 回成功する確率は

$$
P(X = x) = f(x) = {}_nC_x p^x (1-p)^{n-x}, \qquad x = 0, 1, \cdots, n \tag{7.15}
$$

で与えられる．この確率分布を**二項分布** (binomial distribution) という．この二項分布は試行回数 n と成功確率 p を定めれば，形状が一意に決定される．このように，

分布の形を定める定数のことを**分布のパラメータ** (parameters)，もしくは母数という．パラメータ n, p をもつ二項分布をとくに

$$Bi(n, p)$$

と表すことにする．また，ある確率変数 X の分布が二項分布 $Bi(n, p)$ になるとき，「X は $Bi(n, p)$ に従う」と言い方をして，

$$X \sim Bi(n, p) \tag{7.16}$$

と示すことがある．$Bi(1, p)$ という特殊ケースが先のベルヌーイ分布である．

ところで，「二項定理」を用いると，

$$\sum_{x=0}^{n} {}_nC_x p^x (1-p)^{n-x} = (p + (1-p))^n = 1 \tag{7.17}$$

となり，確かに確率分布であることが確かめられる．

二項確率変数は，独立なベルヌーイ確率変数の和であるから，確率変数の和についての定理（次章も参照）より以下が成立する．

■二項分布の平均

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_n) = np$$

■二項分布の分散

$$V(X) = V(X_1) + V(X_2) + \cdots + V(X_n) = np(1-p)$$

ところで，二項分布に従う確率変数 X が独立したベルヌーイ確率変数の和であるという事実から，以下のような定理もほぼ自明に成立する．

定理 7.1. 確率変数 X が $Bi(n, p)$ に従い，確率変数 Y が $Bi(m, p)$ に従っており，かつ，X と Y が独立であるなら，

$$X + Y \sim Bi(n + m, p).$$

このように元の分布のそれらの和の分布が同じ種類の分布であるとき，そうした種類の分布は**再生的** (reproductive) であるという．

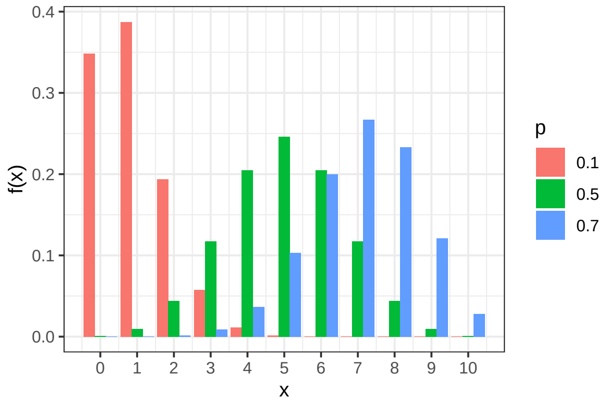

図 7.4　二項分布の例（$n = 10, p = 0.1, 0.5, 0.7$）

練習問題 7.5

$n = 5, 10, 20$ それぞれの場合の，$p = 0.3, 0.5, 0.7$ の二項分布の棒グラフを作成してレポートにまとめよ．`R` や `Excel` を使ってもよいし，手書きでもよい．Word か pdf で作成して LUNA 上で提出せよ．

第 8 章

総和

本章では，統計において欠かすことのできない計算法である「総和」について解説していく．

8.1　総和の定義

n 個の有限数列 $a_1, a_2, \cdots, a_n$ をすべて足したものを**総和** (summation) といって

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \cdots + a_n \tag{8.1}$$

と表す．記号 $\sum$ は Summation, Sum の頭文字 S にあたるギリシア文字シグマの大文字であって，総和という操作を表す．$\sum$ の右には総和の対象である数列の一般項が記される．ここで i は一般化された**添え字** (index) であり，数列の要素番号の変数である．添え字は，一般的には index の i が用いられるが，他の変数と区別がつけば他の記号でもかまわない[*1]．$\sum$ の下は添え字 i が 1 から始まることを示し，$\sum$ の上は添え字番号が n で終わることを示している．添え字の最初と最後は，文脈上明らかな場合はしばしば省略される．

■例（1–10 の数列の総和）　10 個の数 $a_1 = 1, a_2 = 2, \cdots, a_{10} = 10$ の総和は

$$\sum_{i=1}^{10} a_i = 1 + 2 + \cdots + 10 = 55 \tag{8.2}$$

[*1] 一般的には，i から始まって j, k を順番に使うことが多い．

である．これは添え字 i を直接使って，$a_i = i$ なので，

$$\sum_{i=1}^{10} i = 1 + 2 + \cdots + 10 = 55 \tag{8.3}$$

と書いてもよい．

8.2　総和の基本公式

(1) $ca_1, ca_2, \cdots, ca_n$，ただし c を定数とすると，

$$\begin{aligned}\sum_{i=1}^{n} ca_i &= ca_1 + ca_2 + \cdots + ca_n \\ &= c(a_1 + a_2 + \cdots + a_n) = c\sum_{i=1}^{n} a_i.\end{aligned} \tag{8.4}$$

(2) 2 つの数列 $a_1, a_2, \cdots, a_n$ と $b_1, b_2, \cdots, b_n$ について

$$\begin{aligned}\sum_{i=1}^{n}(a_i + b_i) &= (a_1 + b_1) + (a_2 + b_2) + \cdots + (a_n + b_n) \\ &= (a_1 + a_2 + \cdots + a_n) + (b_1 + b_2 + \cdots + b_n) = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} b_i.\end{aligned} \tag{8.5}$$

(3) $a_i = c$ (定数) となる数列 $a_1, a_2, \cdots, a_n$ について

$$\begin{aligned}\sum_{i=1}^{n} c &= \sum_{i=1}^{n} a_i \\ &= a_1 + a_2 + \cdots + a_n \\ &= \underbrace{c + c + \cdots + c}_{n\ 個} = nc.\end{aligned} \tag{8.6}$$

とくに，$\sum_{i=1}^{n} 1 = n$ である．

練習問題 8.2

次の式を展開せよ．

(1) $\displaystyle\sum_{i=1}^{5} 2i$　　(2) $\displaystyle\sum_{i=1}^{n}(a_i - b_i)$

(3) $\displaystyle\sum_{i=1}^{n} 3a_i + 2b$　　(4) $\displaystyle\sum_{i=1}^{n}(a_i + b_i)^2$

8.3　平均

8.3.1　平均の定義

記述統計学の枠組みで 1 変数のデータ $x_1, x_2, \cdots, x_n$ が得られたとする．このときデータの**平均** (mean) は

$$\begin{aligned}\bar{x} &= \frac{1}{n}\sum_{i=1}^{n} x_i \\ &= \sum_{i=1}^{n}\frac{1}{n}x_i \\ &= \frac{1}{n}x_1 + \frac{1}{n}x_2 + \cdots + \frac{1}{n}x_n \\ &= \frac{x_1 + x_2 + \cdots + x_n}{n} \end{aligned} \tag{8.7}$$

と定義される．つまり，平均は各値に $1/n$ の重みを等しく与えた総和である．

なお，離散確率変数 X の平均 (期待値) は

$$\mu = E(X) = \sum_{i=1}^{k} x_i f(x_i), \tag{8.8}$$

である．つまり，確率変数の実現値にその値が実現する確率の重みを与えた総和が平均である．

8.3.2　平均の性質

さらに，データの平均の性質について見ていこう．定義より，

$$\sum_{i=1}^{n} x_i = n\bar{x} \tag{8.9}$$

つまり，データの総和は平均にデータサイズ n をかけたものに変換できる．左辺は n 個の値の情報で構成されるが，右辺は n と $\bar{x}$ という 2 つの値の情報に縮約されている．このように，平均はデータの情報縮約の 1 つの方法である．

各値と平均との差 $x_1 - \bar{x}, x_2 - \bar{x}, \cdots, x_n - \bar{x}$ の総和（偏差和ともいう）は，平均 $\bar{x}$ が定数であることに注意すると，

$$\begin{aligned} \sum_{i=1}^{n}(x_i - \bar{x}) &= \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \bar{x} \\ &= n\bar{x} - n\bar{x} = 0 \end{aligned} \tag{8.10}$$

である．これはどのような変数のデータでも成り立つ．偏差和はつねに 0 である．逆に言うと，偏差和が常に 0 になるように定義した代表値が平均であるとも言える．

これを別の角度から見てみよう．いま，偏差和全体を，非負になる偏差 $x_j - \bar{x}$ と負になる偏差 $x_k - \bar{x}$ に分けたとすると，

$$\sum_{i=1}^{n}(x_i - \bar{x}) = \sum_{j}(x_j - \bar{x}) + \sum_{k}(x_k - \bar{x}) = 0 \tag{8.11}$$

より，

$$\sum_{j}(x_j - \bar{x}) = \sum_{k}(\bar{x} - x_k) \tag{8.12}$$

つまり，平均より大きい値についての平均からの距離の総和と，平均より小さい値についての平均からの距離の総和は常に等しい．逆に言うと，代表値より大きい値についての距離の総和と，代表値より小さい値についての距離の総和を等しくするように定義された代表値が平均であるとも言える．

偏差和と平均の関係を図示したのが図 8.1 である．データを値の昇順に左から右に並べる（青い棒グラフ）．それぞれの棒の底辺の長さを 1，高さをデータの値とすると，棒グラフの面積は $\sum_{i=1}^{n} 1 \cdot x_i = \sum_{i=1}^{n} x_i$ である．赤い横線は平均値 $\bar{x}$ を示す．

次に，平均からの偏差部分を赤い棒グラフで示す．平均より小さい値についての平均からの距離の総和 $\sum_{k}(\bar{x} - x_k)$ と，平均より大きい値についての平均からの距離の総和 $\sum_{j}(x_j - \bar{x})$ は常に等しい．そこで，平均より小さくて足りない部分を平均より大きくて出っ張った部分からもってきて「均等にならしたもの」が左下の長方形であ

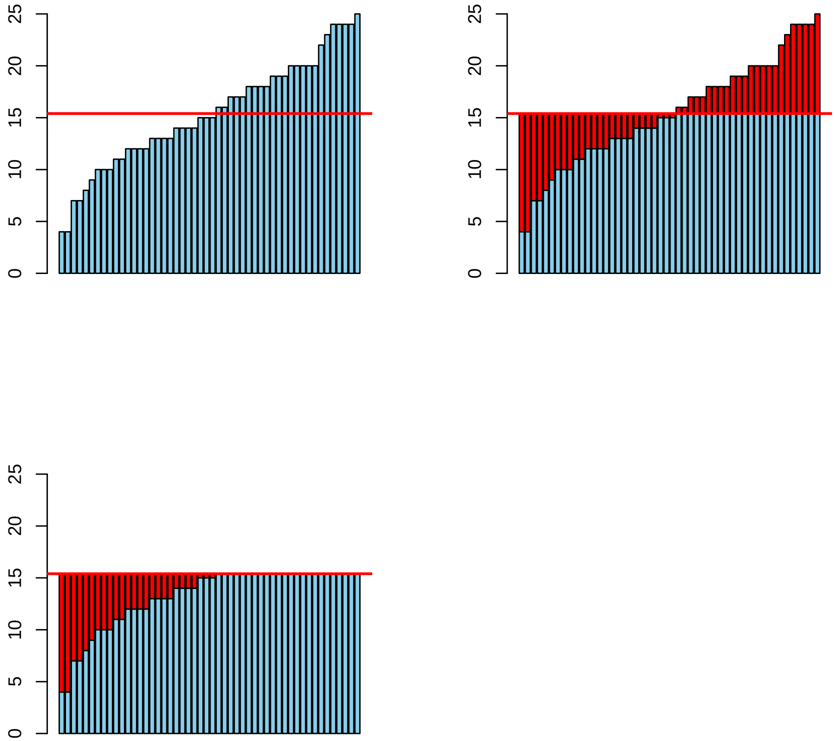

図 8.1　偏差和と平均

る．長方形の底辺は $\sum_{i=1}^{n} 1 = n$，高さは平均 $\bar{x}$ なので，面積は $n\bar{x}$ で，失われた面積はないので

$$\sum_{i=1}^{n} x_i = n\bar{x}$$

である．ここから分かるとおり，「平均」とはまさに「平らに均す」ものである．

確率の平均についても，$\sum_{i=1}^{n} f(x_i) = 1$ に注意して，平均からの偏差の期待値をとると，

$$\begin{aligned} E(X-\mu) &= \sum_{i=1}^{n}(x_i-\mu)f(x_i) \\ &= \sum_{i=1}^{n} x_i f(x_i) - \mu \sum_{i=1}^{n} f(x_i) \\ &= \mu - \mu = 0 \end{aligned} \tag{8.13}$$

となって常に 0 となることが分かる．

8.4　分散

8.4.1　分散の定義

1 変数のデータ $x_1, x_2, \cdots, x_n$ が得られたとする．このときデータの**分散** (variance) は

$$\begin{aligned} s_x^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 \\ &= \frac{(x_1-\bar{x})^2 + (x_2-\bar{x})^2 + \cdots + (x_n-\bar{x})^2}{n} \end{aligned} \tag{8.14}$$

と定義される．つまり，分散は平均からの偏差の平方の総和（偏差平方和）を n で割ったものである[*2]．分散のルートをとったものを**標準偏差** (standard deviation) といい，s_x で表す．つまり，

$$s_x = \sqrt{s_x^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}. \tag{8.15}$$

[*2] 推測統計の文脈では，一般に $n-1$ で割る「不偏分散」が用いられる．

離散確率変数 X の分散は

$$\begin{aligned}\sigma^2 &= E[(X-\mu)^2] \\ &= \sum_{i=1}^{n}(X-\mu)^2 f(x_i) \\ &= (x_1-\mu)^2 f(x_1) + (x_2-\mu)^2 f(x_2) + \cdots + (x_n-\mu)^2 f(x_n)\end{aligned} \tag{8.16}$$

と定義される．離散確率変数 X の標準偏差は，$\sigma = \sqrt{\sigma^2}$ である．

8.4.2　分散の特性

先ほどの偏差和はプラス部分とマイナス部分が打ち消し合って，常に 0 となるので，ばらつきの指標として用いることができない．そこで，偏差の 2 乗をとることで，平均からの各値の距離の平均をばらつきの指標としたものが分散である．

分散の定義はさらに以下のように変形することができる．

$$\begin{aligned}s_x^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x} + \bar{x}^2) \\ &= \frac{1}{n}\sum_{i=1}^{n}x_i^2 - 2\bar{x}\frac{1}{n}\sum_{i=1}^{n}x_i + \bar{x}^2\frac{1}{n}\sum_{i=1}^{n}1 \\ &= \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \bar{x}^2\end{aligned} \tag{8.17}$$

つまり，分散は x_i^2 の平均と x_i の平均の 2 乗に分解できる．

離散確率変数 X の分散についても同様に以下のように変形できる．

$$\begin{aligned}\sigma^2 &= E[(X-\mu)^2] \\ &= E(X^2) - E(X)^2\end{aligned} \tag{8.18}$$

8.5　一次変換と標準化

8.5.1　一次変換

データ x_i，もしくは離散確率変数 X を a 倍して b を足す操作を一次変換という．データ x_i について

$$\forall i \in \{1, 2, \cdots, n\}, y_i = ax_i + b \tag{8.19}$$

と変換する．変換後の平均 $\bar{y}$ と分散 s_y^2 について以下のことが成立する．

(1) 平均

$$\begin{aligned}\bar{y} &= \frac{1}{n}\sum_{i=1}^{n} y_i \\ &= \frac{1}{n}\sum_{i=1}^{n}(ax_i + b) \\ &= a\frac{1}{n}\sum_{i=1}^{n} x_i + \frac{1}{n}\sum_{i=1}^{n} b = a\bar{x} + b\end{aligned} \tag{8.20}$$

(2) 分散

$$\begin{aligned}s_y^2 &= \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})^2 \\ &= \frac{1}{n}\sum_{i=1}^{n}(ax_i + b - (a\bar{x} + b))^2 \\ &= a^2\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = a^2 s_x^2\end{aligned} \tag{8.21}$$

これらの性質は離散確率変数 X の一次変換 $Y = aX + b$ についても成り立つ．

$$\mu_y = a\mu_x + b \tag{8.22}$$

$$\sigma_y^2 = a^2\sigma_x^2 \tag{8.23}$$

8.5.2　標準化

データの一次変換において，

$$a = \frac{1}{s_x}, \qquad b = -\frac{\bar{x}}{s_x}$$

とする特殊な変換を考える．

$$\forall i \in \{1, 2, \cdots, n\}, z_i = ax_i + b = \frac{x_i - \bar{x}}{s_x} \tag{8.24}$$

とすると，一次変換の特性より，

$$\bar{z} = a\bar{x} + b = \frac{\bar{x} - \bar{x}}{s_x} = 0 \tag{8.25}$$

$$s_z^2 = a^2 s_x^2 = \frac{s_x^2}{s_x^2} = 1 \tag{8.26}$$

$$s_z = \sqrt{s_z^2} = 1 \tag{8.27}$$

つまり，変換後の平均は常に 0，分散と標準偏差は常に 1 となる．

このような変換を**標準化** (standardization) といい，z_i を**標準得点** (standard score) という．標準得点は z 得点ともいわれる．

標準得点を変形すると，

$$z_i = \frac{x_i - \bar{x}}{s_x} \Longleftrightarrow x_i = \bar{x} + z_i s_x \tag{8.28}$$

となる．つまり，標準得点はデータのある値 x_i が，平均から標準偏差の何倍分ずれているかを示す値となっている．たとえば，$z_i = \pm 1$ は平均から値がちょうど標準偏差 1 つ分だけプラスもしくはマイナスに離れていることを示す．

なお，標準化は確率変数についても行うことができる．確率変数 X が平均 μ_x，分散 σ_x^2 をもつとき，確率変数 Z を

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

と定義すれば，Z について

$$\mu_z = 0, \quad \sigma_z = 1$$

が成立する．Z を標準化確率変数という．

練習問題 8.5

(1) 離散確率変数 X の一次変換 $Y = aX + b$ について，$\mu_y = a\mu_x + b, \sigma_y^2 = a^2\sigma_x^2$ が成り立つことを示せ.

(2) 離散確率変数 X の標準化 $Z = (X - \mu)/\sigma$ について，$\mu_z = 0, \sigma_z = 1$ が成り立つことを示せ.

第 9 章

微分

実数値関数の挙動を調べる際に用いられる手法が微分である．ここでは，基本的な関数の導関数の求め方，合成関数の微分，極値の求め方を簡単に確認した上で，統計手法において用いられる微分法を確認する．

以下では，区間，極限，連続といった基本的概念は（なんとなくであれ）了解のものとして話を進める．基本的概念について確認したい場合は付録 A.2 を参照.

9.1　微分係数

関数 $y = f(x)$ がある区間 D において定義されているとし，$a, b \in D$ をその区間に属する異なる 2 点とする．このとき，

$$f(b) - f(a) \tag{9.1}$$

は x の値が a から b まで変化するときの，関数の値の変化量を表している．そして，

$$\frac{f(b) - f(a)}{b - a} \tag{9.2}$$

は x の変化量に対する関数の変化量の比を示している．これを，x が a から b まで変化するときの関数 $f(x)$ の**平均変化率** (average rate of change) という[*1]．図形的にいえば，関数 $y = f(x)$ のグラフ上の 2 点 $P(a, f(a)), Q(b, f(b))$ を通る直線の傾きを表している．

x の変化量を

$$\Delta x = h = b - a \tag{9.3}$$

[*1] ニュートン商 (Newton quotient) と呼ばれることもある．

とおいて，それに対する関数 $y = f(x)$ の変化量を

$$\Delta y = f(a+h) - f(a) = f(a+\Delta x) - f(a) \tag{9.4}$$

とおく．このとき，平均変化率は

$$\frac{\Delta y}{\Delta x} = \frac{f(a+h) - f(a)}{h} \tag{9.5}$$

と表すこともできる．

さて，$b = a + h$ であるので，$b \to a$ と $h \to 0$ は同値である．$b \to a$ すなわち $h \to 0$ のとき，平均変化率の有限の極限値が存在するならば，この極限値を関数 $f(x)$ の $x = a$ における**微分係数** (differential coefficient) といい，$f'(a)$ で表す．すなわち，

$$f'(a) = \lim_{b \to a} \frac{f(b) - f(a)}{b - a} = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \tag{9.6}$$

である．ここで，$f'(a)$ は 1 つの定数である．微分係数 $f'(a)$ が存在するとき，関数 $f(x)$ は $x = a$ において**微分可能** (differentiable) であるという．また，関数 $f(x)$ が区間 D のすべての点で微分可能であれば，つまり，

$$\forall a \in D, f'(a) \text{ が存在する} \tag{9.7}$$

が成り立てば，$f(x)$ は**区間 $\boldsymbol{D}$ で微分可能** (differentiable in interval D) であるという．

微分可能と連続の関係について，以下のことが成り立つ．すなわち，関数 $f(x)$ は $x = a$ において微分可能ならば，$f(x)$ は $x = a$ において連続である．したがって，$f(x)$ が区間 D で微分可能ならば，$f(x)$ は D で連続である（詳細は付録 A.3）．

つまり，「微分可能 $\Longrightarrow$ 連続」であるので，逆命題である「連続 $\Longrightarrow$ 微分可能」は常に成立するとは限らないことに注意する．ごくごくおおざっぱな視覚的イメージでいうと，区間 D において微分可能な関数は，D において連続であるばかりでなく，なめらかな（カクカクしていない）関数である必要がある．

練習問題 9.1

次の関数 $y = f(x)$ の $x = a$ における微分係数を求めよ（矢野・田代 (1993: 70) より）．

(1)　x　(2)　x^2　(3)　x^3　(4)　k (定数)
(5)　$\dfrac{1}{x}$ $(a \neq 0)$

9.2　導関数

関数 $y = f(x)$ が，ある区間 D において微分可能であるとする．このとき，$\forall a \in D$ について微分係数 $f'(a)$ が定まるのであった．ここで，定数 a を独立変数 x に置き換えると，変数 x に応じて値 $f'(x)$ をとる関数を得ることができる．この関数 $f'(x)$ を $f(x)$ の**導関数** (derivative, derived function) と呼ぶ．しつこく言えば，**微分係数 $f'(a)$ は 1 つの定数であるが，導関数 $f'(x)$ は x を独立変数とする関数である．**

関数 $y = f(x)$ の導関数は，また

$$y', \qquad \frac{dy}{dx}, \qquad \frac{d}{dx}f(x)$$

とも表される．とくに dy/dx や $df(x)/dx$ は導関数の古典的な表記であり，

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \tag{9.8}$$

に由来する．ただし，$\Delta y/\Delta x$ は x の変化量と y の変化量の比を示す分数であるが，**dy/dx は分数ではなく全体で導関数を表す記号である**．実際，dy/dx は「ディーワイ分のディーエックス」と読むのではなく「ディーワイディーエックス」と読むのが慣習である．

関数 $f(x)$ から，その導関数 $f'(x)$ を求めることを，**$f(x)$ を x で微分する** (differentiate $f(x)$ with respect to x) という．

9.3　導関数の基本定理

まず始めに，定数関数 $f(x) = c$ の導関数を検討する．この関数の平均変化率は

$$\frac{f(x+h) - f(x)}{h} = \frac{c - c}{h} = 0 \tag{9.9}$$

なので

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = 0 \tag{9.10}$$

である．次に，$f(x)=x$ を考えよう．

$$\frac{f(x+h)-f(x)}{h}=\frac{x+h-x}{h}=1 \tag{9.11}$$

なので

$$f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=1 \tag{9.12}$$

である．また，$f(x)=x^2$ のとき，

$$\frac{f(x+h)-f(x)}{h}=\frac{x^2+2hx+h^2-x^2}{h}=2x+h \tag{9.13}$$

なので

$$f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=2x \tag{9.14}$$

である．これらの結果を

$$(c)'=0, \qquad (x)'=1, \qquad \left(x^2\right)'=2x$$

とも書くことにする．このように，ある関数の導関数について 1 つずつ定義通りに求めていってもよいが，普通はよく知られた以下のような法則を用いて解いていく．

定理 9.1. $f(x)$, $g(x)$ が区間 D で微分可能な関数ならば，その区間において以下のことが成立する

(1) $\{kf(x)\}'=kf'(x)$　　（ただし，k は定数）

(2) $\{f(x)+g(x)\}'=f'(x)+g'(x)$

(3) $\{f(x)-g(x)\}'=f'(x)-g'(x)$

(4) $\{f(x)g(x)\}'=f'(x)g(x)+f(x)g'(x)$

(5) $\left\{\dfrac{f(x)}{g(x)}\right\}'=\dfrac{f'(x)g(x)-f(x)g'(x)}{\{g(x)\}^2}$　　（ただし，$g(x)\neq 0$ のとき）

証明． (1)

$$\{kf(x)\}'=\lim_{h\to 0}\frac{kf(x+h)-kf(x)}{h}=k\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=kf'(x) \tag{9.15}$$

(2)

$$\{f(x)+g(x)\}' = \lim_{h\to 0}\frac{\{f(x+h)+g(x+h)\}-\{f(x)+g(x)\}}{h} \tag{9.16}$$

$$= \lim_{h\to 0}\frac{f(x+h)-f(x)}{h} + \lim_{h\to 0}\frac{g(x+h)-g(x)}{h} = f'(x)+g'(x) \tag{9.17}$$

(3) についても上と同様に証明される．

(4)

$$\{f(x)g(x)\}' = \lim_{h\to 0}\frac{f(x+h)g(x+h)-f(x)g(x)}{h} \tag{9.18}$$

$$= \lim_{h\to 0}\frac{f(x+h)g(x+h)-f(x)g(x+h)+f(x)g(x+h)-f(x)g(x)}{h} \tag{9.19}$$

$$= \lim_{h\to 0}\left\{\frac{f(x+h)-f(x)}{h}g(x+h)+f(x)\frac{g(x+h)-g(x)}{h}\right\} \tag{9.20}$$

ここで，微分可能と連続の関係より，関数 $g(x)$ は連続である．ゆえに

$$\lim_{h\to 0} g(x+h) = \lim_{(x+h)\to x} g(x+h) = g(x) \tag{9.21}$$

である．$h \to 0$ のとき，$g(x+h) \to g(x)$ なので

$$\{f(x)g(x)\}' = f'(x)g(x)+f(x)g'(x) \tag{9.22}$$

が成立する．

(5)

$$\left\{\frac{f(x)}{g(x)}\right\}' = \lim_{h\to 0}\frac{1}{h}\left\{\frac{f(x+h)}{g(x+h)}-\frac{f(x)}{g(x)}\right\} \tag{9.23}$$

$$= \lim_{h\to 0}\frac{f(x+h)g(x)-f(x)g(x+h)}{hg(x+h)g(x)} \tag{9.24}$$

$$= \lim_{h\to 0}\frac{1}{g(x+h)g(x)}\left\{\frac{f(x+h)-f(x)}{h}g(x)-f(x)\frac{g(x+h)-g(x)}{h}\right\} \tag{9.25}$$

$$= \frac{1}{\{g(x)\}^2}\{f'(x)g(x)-f(x)g'(x)\} \tag{9.26}$$

□

定理 9.2. 任意の整数 $n \in \mathbb{Z}$ について

$$(x^n)' = nx^{n-1} \tag{9.27}$$

が成立する．ただし，n が負のときには，$x \neq 0$ であると仮定する．

証明. (i) n が正の整数，(ii) n が 0，(iii) n が負の整数のときの 3 つに場合分けして証明する．

(i)　n は正の整数とする．関数 $f(x) = x^n$ の平均変化率は，

$$\frac{f(x+h) - f(x)}{h} = \frac{(x+h)^n - x^n}{h} \tag{9.28}$$

である．ここで，$u = x + h$ とおくと

$$\frac{f(x+h) - f(x)}{h} = \frac{u^n - x^n}{u - x} \tag{9.29}$$

である．ところで

$$u^n - x^n = (u - x)(\underbrace{u^{n-1} + u^{n-2}x + u^{n-3}x^2 + \cdots + u^2x^{n-3} + ux^{n-2} + x^{n-1}}_{n\text{ 個の項}}) \tag{9.30}$$

が成立し（気になる人は自ら確かめよ），$h \to 0$ のとき $u \to x$ であるので，

$$\begin{aligned}
\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} &= \lim_{u \to x} \frac{u^n - x^n}{u - x} \\
&= \lim_{u \to x} (u^{n-1} + u^{n-2}x + u^{n-3}x^2 + \cdots + u^2x^{n-3} + ux^{n-2} + x^{n-1}) \\
&= x^{n-1} + x^{n-2}x + x^{n-3}x^2 + \cdots + x^2x^{n-3} + xx^{n-2} + x^{n-1} \\
&= nx^{n-1}
\end{aligned} \tag{9.31}$$

となり，$(x^n)' = nx^{n-1}$ が成立する．

(ii)　$n = 0$ とする．$x^0 = 1$ かつ $(1)' = 0$ である．一方，$0 \cdot x^{-1} = 0$ なので定理の等式は成立する．

(iii)　n は負の整数とする．また $x \neq 0$ とする．$n = -m$ とおくと

$$x^n = x^{-m} = \frac{1}{x^m} \tag{9.32}$$

と書き換えられる．ここで m は正の整数である．右辺と左辺を x について微分すると

$$(x^n)' = \left(\frac{1}{x^m}\right)' \tag{9.33}$$

である．このとき，右辺について定理 9.1 の (5) を適用すると，

$$\begin{aligned}\left(\frac{1}{x^m}\right)' &= \frac{(1)'x^m - 1\cdot(x^m)'}{(x^m)^2} \\ &= \frac{-mx^{m-1}}{x^{2m}} = -\frac{m}{x^{m+1}} = -mx^{-(m+1)}\end{aligned} \tag{9.34}$$

となる．$m=-n$ より，$(x^n)' = nx^{n-1}$ が成立する． □

練習問題 9.3

次の関数を微分せよ（矢野・田代 (1993: 73) より）．

(1) x^2-4x　(2) x^3+3x　(3) $(3x-2)(x+5)$

(4) $\dfrac{1}{x}+x^2$　(5) $(x^3+x^2)(x+1)$　(6) $\dfrac{5x-3}{4x+1}$

(7) $\dfrac{1}{x^2}$　(8) $\dfrac{1}{x^3}$　(9) $\left(x+\dfrac{1}{x}\right)^2$

9.4　合成関数の微分

次に導関数を求める際にしばしば用いられるテクニックである合成関数の微分を導入する．詳細は付録 A.4 節参照．

定理 9.3 (合成関数の微分)**.** 関数 $y=f(x)$ が区間 I において微分可能で，その値域が区間 J に含まれているとする．このとき，$z=g(y)$ が区間 J において微分可能であれば，合成関数

$$z = (g\circ f)(x) = g(f(x)) \tag{9.35}$$

は区間 I において微分可能であり，

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} = g'(f(x))f'(x) \tag{9.36}$$

である．

例えば，関数 $z = (x^3 + 2x^2 + x)^4$ は $y = x^3 + 2x^2 + x$ とおけば，定理 9.3 より，

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} = 4y^3(3x^2 + 4x + 1) = 4(x^3 + 2x^2 + x)^3(3x^2 + 4x + 1)$$

と計算できる．

練習問題 9.4

次の関数を微分せよ（矢野・田代 (1993: 83) より）．

(1) $\left(x^3 + 2\right)^2$　(2) $\left(x^2 + 1\right)^2(3x - 1)^4$　(3) $\dfrac{1}{(x + 3)^2}$

9.5　極点を求める

ここまで，導入した微分法は関数の挙動，とくに極点を求める際に必要になってくる．ここでは，最低限の知識を導入する．微分と関数の挙動の詳細については，付録 A.5 節を参照.

関数 $f(x)$ の定義域内の 1 点 c とその十分近くの任意の x について，$f(c) \geq f(x)$ が成立するとき，もう少し厳密に言うと，$a < c < b$ を満たす a と b を c の十分近くにとるとき

$$\forall x \in (a, b), \qquad f(c) \geq f(x) \tag{9.37}$$

が成立するならば，c は f の**極大点** (local maximum point)，$f(c)$ は f の**極大値** (local maximum value) という．一方，

$$\forall x \in (a, b), \qquad f(c) \leq f(x) \tag{9.38}$$

が成立するならば，c は f の**極小点** (local minimum point)，$f(c)$ は f の**極小値** (local minimum value) という．極大点と極小点を合わせて**極値点**，極大値と極小値を合わせて**極値**という．

さらに，

$$\forall x \in (a,c) \cup (c,b), \qquad f(c) > f(x) \tag{9.39}$$

が成立するならば，c を**狭義の極大点**，$f(c)$ を**狭義の極大値**ともいう．また，

$$\forall x \in (a,c) \cup (c,b), \qquad f(c) < f(x) \tag{9.40}$$

が成立するならば，c を**狭義の極小点**，$f(c)$ を**狭義の極小値**ともいう．

関数 $f(x)$ の定義域 D 内の 1 点 c について，

$$\forall x \in D, \qquad f(c) \geq f(x) \tag{9.41}$$

が成立するとき，c を D における f の**最大点** (global maximum point)，$f(c)$ を D における**最大値** (global maximum value) という．同様に，

$$\forall x \in D, \qquad f(c) \leq f(x) \tag{9.42}$$

が成立するとき，c を D における f の**最小点** (global minimum point)，$f(c)$ を D における**最小値** (global minimum value) という．

以下の定理を証明なしで導入する（詳細は付録 A.5 節を参照）．

定理 9.4. *$f(x)$ が区間 D で定義された関数，c を D の 1 つの端点ではない点とする．このとき，f は c において微分可能であり，かつ c が f の極値点であれば，*

$$f'(c) = 0 \tag{9.43}$$

である．

定理 9.5. 関数 $f(x)$ が区間 D で連続で，かつ D の端点を取り除いた区間（これを D の内部という）で微分可能であるとする．そのとき，

(1)　D の内部でつねに $f'(x) > 0$ ならば，f は D において単調に増加する．

(2)　D の内部でつねに $f'(x) < 0$ ならば，f は D において単調に減少する．

(3)　D の内部でつねに $f'(x)=0$ ならば，f は D において定数である．

ある区間で定義された関数 $y=f(x)$ の導関数 $y'=f'(x)$ が，その区間においてさらに微分可能であれば，その導関数を**二次導関数** (second derivative) といって，

$$f''(x), \qquad y'', \qquad \frac{d^2y}{dx^2}, \qquad \frac{d^2}{dx^2}f(x)$$

などと表す．

定理 9.6. 関数 $f(x)$ が区間 D で第 *2* 次導関数 $f''(x)$ をもち，f'' が連続であるとする．このとき，D の内部の点である a について，

$$f'(a)=0, f''(a)>0 \text{ ならば, } a \text{ は } f \text{ の狭義の極小点であり,}$$
$$f'(a)=0, f''(a)<0 \text{ ならば, } a \text{ は } f \text{ の狭義の極大点である.}$$

9.6　平均を導出する

統計に関連した微分法の応用例を示そう．

一変数のデータ $x_1, x_2, \cdots, x_n$ が得られたとき，ある値 m からの偏差平方和

$$f(m)=\sum_{i=1}^{n}(x_i-m)^2 \tag{9.44}$$

を最小にするように値 m をとり，これをデータの代表値としたい．導関数の基本定理と合成関数の微分を用いて，

$$\begin{aligned} f'(m) &= -2\sum_{i=1}^{n}(x_i-m) \\ &= -2\left[\sum_{i=1}^{n}x_i-nm\right] \end{aligned} \tag{9.45}$$

$$f''(m)=2n>0 \tag{9.46}$$

ゆえに，

$$f'(m)=-2\left[\sum_{i=1}^{n}x_i-nm\right]=0 \tag{9.47}$$

が狭義の極小点である．これを m について解くと

$$m = \frac{1}{n}\sum_{i=1}^{n} x_i = \bar{x} \tag{9.48}$$

を得る．$m < \bar{x} \Longrightarrow f'(m) < 0, m > \bar{x} \Longrightarrow f'(m) > 0$ なので，m は $f(m)$ の最小点でもある．

つまり，平均値は偏差平方和を最小にするデータの代表値である．これが，代表的なパラメータ推定法である**最小二乗法** (least square method) のもっとも単純な例である．回帰係数の最小二乗推定は，この考えを多変数に拡張したものであり，多変数関数の微分法である偏微分によって最小化問題を解くことになる．

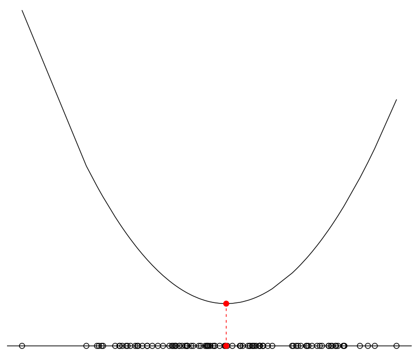

図 9.1　偏差平方和を最小にする値としての平均

第 10 章

積分

本章では，連続確率分布の考え方を理解するために必要な，必要最低限の積分の知識を導入する．

10.1　区分求積法

放物線 $y = x^2$ と x 軸および直線 $x = 1$ によって囲まれた図形の面積 S を求めたい．このために，この図形の長方形の和によって近似することを考える．

区間 $[0, 1]$ を n 等分して，長方形の底辺を作る．そして，高さは底辺の左端における x の値に対応する $f(x) = x^2$ とする．例えば，$n = 8$ のとき，8 つ（実質的には 7 つ）の長方形があり，底辺は 1/8，高さは左から $0^2, (1/8)^2, (2/8)^2, (3/8)^2, (4/8)^2, (5/8)^2, (6/8)^2, (7/8)^2$ となる（図 10.1）．この長方形の面積の合計が問題の面積 S の下からの近似になる．一般に n 等分したときの面積 S の下からの近似を s_n とすると，

$$s_n = \frac{1}{n}\left\{0^2 + \left(\frac{1}{n}\right)^2 + \left(\frac{2}{n}\right)^2 + \cdots + \left(\frac{n-1}{n}\right)^2\right\} \tag{10.1}$$

$$= \frac{1}{n^3}\left\{1^2 + 2^2 + \cdots + (n-1)^2\right\} \tag{10.2}$$

$$= \frac{1}{6n^2}(n-1)(2n-1) = \frac{1}{6}\left(1 - \frac{1}{n}\right)\left(2 - \frac{1}{n}\right) \tag{10.3}$$

第 2 式から第 3 式は自然数の平方和の公式 $\sum_{k=1}^{n} k^2 = (1/6)n(n+1)(2n+1)$ を用いた．

次に，同じように区間 $[0, 1]$ を n 等分して，長方形の底辺を作る．そして，高さは底辺の右端における x の値に対応する $f(x) = x^2$ とする．例

えば，$n = 8$ のとき，8 つの長方形があり，底辺は 1/8，高さは左から $(1/8)^2, (2/8)^2, (3/8)^2, (4/8)^2, (5/8)^2, (6/8)^2, (7/8)^2, (8/8)^2$ となる（図 10.2）．この長方形の面積の合計は面積 S の上からの近似になる．一般に n 等分したときの面積 S の上からの近似を S_n とすると，

$$S_n = \frac{1}{n}\left\{\left(\frac{1}{n}\right)^2 + \left(\frac{2}{n}\right)^2 + \cdots + \left(\frac{n}{n}\right)^2\right\} \tag{10.4}$$

$$= \frac{1}{n^3}\left\{1^2 + 2^2 + \cdots + n^2\right\} \tag{10.5}$$

$$= \frac{1}{6n^2}(n+1)(2n+1) = \frac{1}{6}\left(1 + \frac{1}{n}\right)\left(2 + \frac{1}{n}\right) \tag{10.6}$$

である．s_n と S_n，そして真の面積 S の関係を考えると，n が有限であれば明らかに

$$s_n < S < S_n \tag{10.7}$$

である．ところで，s_n と S_n の $n \to \infty$ のときの極限値をとると，

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} \frac{1}{6}\left(1 - \frac{1}{n}\right)\left(2 - \frac{1}{n}\right) = \frac{1}{3} \tag{10.8}$$

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{1}{6}\left(1 + \frac{1}{n}\right)\left(2 + \frac{1}{n}\right) = \frac{1}{3} \tag{10.9}$$

ゆえに

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} S_n = \frac{1}{3}$$

であるので，n を大きくしていったとき，s_n と S_n は S をはさんで，ただ 1 つの値 1/3 に近づく．このことから，$S = 1/3$ である（と定義する）．このように面積を求める方法を**区分求積法**という．以下では，区分求積法を一般化した方法としての定積分を導入する．

10.2 定積分の定義

一般的に f を区間 $[a, b]$ で定義された**連続関数**とする．ここで区間 $[a, b]$ を

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = b \tag{10.10}$$

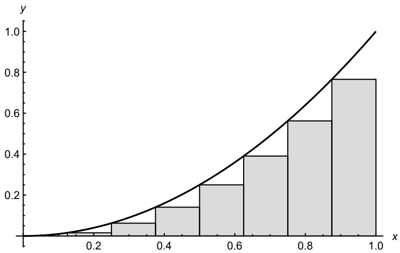

図 10.1　面積 S への下からの近似（$n=8$）

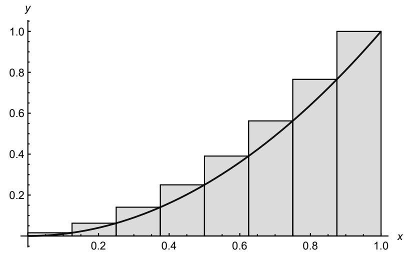

図 10.2　面積 S への上からの近似（$n=8$）

を満たす有限個の点列 $x_0, x_1, x_2, \cdots, x_n$ で分割することを考える．**分割** (partition) という操作を一般的に

$$P = (x_0, x_1, x_2, \cdots, x_n) \tag{10.11}$$

というように書く．P によって分割された $[a,b]$ の小区間はそれぞれ

$$[x_0, x_1], [x_1, x_2], \cdots, [x_{n-1}, x_n]$$

となる．分割は必ずしも等分割である必要はない．

区間 $[a,b]$ について，分割 P によって得られる各々の小区間 $[x_{k-1}, x_k]$ $(k = 1, 2, \cdots, n)$ から任意に点 c_k を選択して，$f(c_k)$ をそれぞれ $(x_k - x_{k-1})$ とかけて長方形の面積を得る．こうしてできる k 個の長方形の面積を足し合わせたものを，分割 P に関する f の**リーマン和** (Riemann sum) と呼び，ここでは記号 $R(P,f)$ で表す[*1]．すなわち，

$$R(P,f) = \sum_{k=1}^{n} f(c_k)(x_k - x_{k-1}) \tag{10.12}$$

である[*2]．ところで，小区間 $[x_{k-1}, x_k]$ における f の最大値と最小値を

$$\max_{c_k} f(c_k) = M_k, \quad \min_{c_k} f(c_k) = m_k \tag{10.13}$$

[*1] しがないリーマンが，デスクに向かって面積を手計算している図をイメージすればよい．

[*2] リーマン和は点列 $C = (c_1, c_2, \cdots, c_n)$ の取り方によっても変化するので，より厳密には $R(P,C,f)$ などと書くべきであるが，ここでは簡単のため C を省略する．

と書くことにする[*3]．このとき，それぞれの k について $f(c_k) = M_k$ となるように c_k をとったときのリーマン和を

$$U(P, f) = \sum_{k=1}^{n} M_k(x_k - x_{k-1}) \tag{10.14}$$

と表して，これを特に**上方和** (upper Riemann sum) という．また，それぞれの k について $f(c_k) = m_k$ となるように c_k をとったときのリーマン和を

$$L(P, f) = \sum_{k=1}^{n} m_k(x_k - x_{k-1}) \tag{10.15}$$

と表して，これを特に**下方和** (lower Riemann sum) という．明らかに，

$$L(P, f) \leq R(P, f) \leq U(P, f) \tag{10.16}$$

である[*4]．また，

$$\max_k M_k = M, \quad \min_k m_k = m \tag{10.17}$$

とすると，

$$\sum_{k=1}^{n} m(x_k - x_{k-1}) = m\left(\sum_{k=1}^{n} x_k - \sum_{k=0}^{n-1} x_k\right) \tag{10.18}$$

$$= m\left(x_n + \sum_{k=1}^{n-1} x_k - \sum_{k=1}^{n-1} x_k - x_0\right) = m(b-a) \tag{10.19}$$

となり，同様に

$$\sum_{k=1}^{n} M(x_k - x_{k-1}) = M(b-a) \tag{10.20}$$

が成り立つ．ゆえに，

$$m(b-a) \leq L(P, f), \quad U(P, f) \leq M(b-a) \tag{10.21}$$

[*3] 例えば，$\max_{c_k} f(c_k)$ は $f(c_k)$ について c_k を変えていったときに最大になる $f(c_k)$ の値，という意味である．

[*4] $\forall k \in \{1, 2 \cdots, n\}, m_k \leq f(c_k) \leq M_k$ であるので，各辺に $(x_k - x_{k-1}) > 0$ をかけると，$\forall k \in \{1, 2 \cdots, n\}, m_k(x_k - x_{k-1}) \leq f(c_k)(x_k - x_{k-1}) \leq M_k(x_k - x_{k-1})$ となる．これを k について総計をとると $L(P, f) \leq R(P, f) \leq U(P, f)$ を得る．

である.

　さて, 分割 $P = (x_0, x_1, \cdots, x_n)$ において, 分割されたそれぞれの小区間 $[x_k - x_{k-1}]$ の長さ $x_k - x_{k-1}$ の中で最大のものを $|P|$ と表すことにする．すなわち，

$$\max_k (x_k - x_{k-1}) = |P| \tag{10.22}$$

とする．このとき，すべての小区間の長さ $x_k - x_{k-1}$ が 0 に近づくように，分割を細かくしていくことを，簡単に $|P| \to 0$ と表すことができる．そして，$|P| \to 0$ のときのリーマン和 $R(P, f)$ の極限値を，区間 $[a, b]$ における f の**定積分** (definite integral) と呼んで

$$\int_a^b f \quad \text{または} \quad \int_a^b f(x)d(x) \tag{10.23}$$

と表す．すなわち，

$$\int_a^b f(x)dx = \lim_{|P| \to 0} R(P, f) \tag{10.24}$$

である[*5]．a をこの定積分の**下端** (lower end)，b を**上端** (upper end) という．**定積分は 1 つの定数であって x の関数ではないことに注意せよ．**定積分内部の x は，$\sum_{i=1}^n$ 記号における添字 i と同様の意味を持つものにすぎないので，これを t に置き換えても他の記号に置き換えても意味は同じである.

　区間 $[a, b]$ において $f(x) \geq 0$ であれば，定積分 $\int_a^b f(x)dx$ は，$y = f(x)$ のグラフと x 軸および，2 つの直線 $x = a, x = b$ によって囲まれる図形の面積を表すと定義するのである.

10.3　定積分の基本的性質

　ここでは，定積分のいくつかの性質を証明なしで導入する．詳細は付録 A.6 を参照.

[*5] より厳密には，関数 f について，区間 $[a, b]$ の分割 P によって異なる上方和のすべての集合の下限と，下方和のすべての集合の上限が一致するとき，関数 f は区間 $[a, b]$ において**積分可能** (integrable) であるといい，そのときの値を定積分という．実数の連続性の性質から，関数 f が $[a, b]$ で連続であれば，$[a, b]$ で積分可能であることが知られている.

定理 10.1. 区間 $[a,b]$ で連続な関数 f について

$$m(b-a) \leq \int_a^b f \leq M(b-a). \tag{10.25}$$

ここまで，われわれは実際の区間を考えてきたので，上端と下端について $a<b$ を仮定してきた．ここで，定積分の計算を一般化するために，$a \geq b$ のとき，

$$\int_a^b f = -\int_b^a f, \qquad \int_a^a f = 0 \tag{10.26}$$

と定義する．このとき，以下の定理が成立する．

定理 10.2. 関数 f が積分可能である区間 D において，任意の $a,b,c \in D$ について，

$$\int_a^b f = \int_a^c f + \int_c^b f \tag{10.27}$$

が成り立つ.

関数 f をある区間 D において連続な関数とする．関数 f は D において積分可能であるので，ある $a \in D$ を下端として，任意の $x \in D$ を上端とする定積分が存在する．これは変数 x の関数と見なすことができるので，この関数を $G(x)$ とおくと

$$G(x) = \int_a^x f(t)dt \tag{10.28}$$

と定義される．このとき，以下の定理が成立する．

定理 10.3. 上記の仮定の下で，関数 $G(x)$ は区間 D で微分可能で，

$$G'(x) = f(x) \tag{10.29}$$

が成立する.

10.4　原始関数の定義

$f(x)$ を区間 D で定義された関数とする．このとき $F(x)$ が同じ区間 D で定義された関数であって，

$$F'(x) = f(x) \tag{10.30}$$

が成り立つとき，関数 $F(x)$ を $f(x)$ の**原始関数** (primitive function) または**不定積分** (indefinite integral) とよぶ．$F(x)$ が $f(x)$ の原始関数であれば，C を任意の定数とするとき，$F(x)+C$ もまた $f(x)$ の原始関数の 1 つとなる．というのも

$$(F(x)+C)' = F'(x) = f(x) \tag{10.31}$$

だからである．逆に言えば，区間 D で f が原始関数を持つならば，任意の定数部分 C を除くと一意に定まる．たとえば,

$$\left(\frac{1}{3}x^3\right)' = x^2 \tag{10.32}$$

なので，x^2 の原始関数は

$$\frac{1}{3}x^3 + C \tag{10.33}$$

と表される．定理 10.3 における $G(x)$ も f の 1 つの原始関数になっている．また，ここから f が連続関数であれば必ず原始関数を持つ，ということが言える．

10.4.1　微分積分学の基本定理

定理 10.4 (**微分積分学の基本定理**)**.** *$f(x)$ を区間 D で定義された連続関数，$F(x)$ を $f(x)$ の 1 つの原始関数とする．このとき，D に属する任意の a,b について,*

$$\int_a^b f = F(b) - F(a) \tag{10.34}$$

が成り立つ.

証明. 与えられた，a,b に対して

$$G(x) = \int_a^x f \tag{10.35}$$

とおくと，定理 10.3 より，関数 G もまた f の 1 つの原始関数である．ゆえに，何らかの定数 C が存在し,

$$G(x) = F(x) + C \tag{10.36}$$

が成り立たなければならない．ここで，$x=a$ とすると

$$G(a) = F(a) + C, \tag{10.37}$$

ここで,

$$G(a) = \int_a^a f = 0 \tag{10.38}$$

なので，$C = -F(a)$ を得る．したがって,

$$G(x) = \int_a^x f = F(x) - F(a) \tag{10.39}$$

となる．ここで $x=b$ とすると,

$$\int_a^b f = F(b) - F(a) \tag{10.40}$$

を得る． □

この定理によって，積分可能な関数 $f(x)$ について，a から b の範囲で定積分を計算するためには，関数 $f(x)$ の不定積分 $F(x)$ を導き出し，公式に当てはめて計算すればよいことが分かる．公式の右辺 $F(b)-F(a)$ はしばしば $[F(x)]_a^b$ と表される．つまり,

$$\int_a^b f(x)dx = [F(x)]_a^b = F(b) - F(a). \tag{10.41}$$

例えば，$f(x) = x^2$ において，下端 0，上端 1 の定積分を計算する．原始関数の 1 つは $F(x) = x^3/3$ なので

$$\int_0^1 x^2 dx = \left[\frac{1}{3}x^3\right]_0^1 = \frac{1}{3}\cdot 1 - \frac{1}{3}\cdot 0 = \frac{1}{3} \tag{10.42}$$

となる．

10.5　不定積分の計算

この節では，連続関数の不定積分の具体的な計算方法を概観する[*6]．さらなる積分計算のテクニックについては，付録 A.7 を参照.

[*6] ここでは原始関数という用語は使わず不定積分で統一する.

一般に関数 $f(x)$ の不定積分を

$$\int f \quad \text{または} \quad \int f(x)dx$$

と表す．$F(x)$ を $f(x)$ の 1 つの不定積分とすると，

$$\int f(x)dx = F(x) + C \tag{10.43}$$

である．このとき定数 C を**積分定数** (constant of integration) とよぶ．関数 f の不定積分を求めることを，f を**積分する** (integrate) という．

不定積分は微分の逆計算であるので，基本的な関数の積分方法は微分の公式から求められる．まずは，定理 9.1 の (1), (2), (3) より，定数倍，和，差の積分について以下の定理が成り立つ．

定理 10.5.

(1) $\displaystyle\int kf(x)dx = k\int f(x)dx$　　（ただし，k は定数）

(2) $\displaystyle\int \{f(x)+g(x)\}dx = \int f(x)dx + \int g(x)dx$

(3) $\displaystyle\int \{f(x)-g(x)\}dx = \int f(x)dx - \int g(x)dx$

また，定理 9.2 より，関数 $x^n(n \in \mathbb{Z})$ の積分について，次の定理が成立する．

定理 10.6. $n \neq -1$ を満たす任意の整数 $n \in \mathbb{Z}$ について，

$$\int x^n dx = \frac{1}{n+1}x^{n+1} + C \tag{10.44}$$

が成立する．ただし，$n \leq -2$ のときは，$x \neq 0$ であると仮定する．

証明. 条件の下で，

$$\left(\frac{1}{n+1}x^{n+1} + C\right)' = x^n \tag{10.45}$$

が成立する．□

練習問題 10.5

次の不定積分を求めよ（松坂 (1989-91: 1008, 1010) より）．

(1) $\displaystyle\int(3x^2-4x-2)dx$　　(2) $\displaystyle\int(2x+x^3)dx$

(3) $\displaystyle\int(x-1)(2x-3)dx$　　(4) $\displaystyle\int\frac{1}{x^2}dx$

(5) $\displaystyle\int(-x^{-6})dx$

10.6　確率密度関数と累積分布関数

確率変数 X が連続の値をとるとき，X を**連続確率変数** (continuous random variable) という．連続確率変数の場合，1 つ 1 つの値ごとではなく，値のとる範囲に応じて確率が決まると考える．確率変数 X がある範囲の値をとる確率が，関数 $f(x)$ によって

$$P(a \leq X \leq b) = \int_a^b f(x)dx \tag{10.46}$$

と表され，かつ

$$\text{すべての } x \text{ について}, f(x) \geq 0 \tag{10.47}$$

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{10.48}$$

を満たすとき，X は**連続型確率分布** (probability distribution of continuous type) をもつという．このとき関数 $f(x)$ を，**確率密度関数** (probability density function; PDF)，あるいは単に密度関数という．

定義より，確率変数がただ 1 つの値 a をとる確率は

$$P(a \leq X \leq a) = \int_a^a f(x)dx = 0$$

である．連続型の場合は，値の範囲にのみ意味があることに注意すること．

また，確率変数 X がある値 x 以下の値をとる確率を示す関数 $F(x)$ を，X の**累積分布関数** (cumulative distribution function: CDF)，あるいは単に分布関数といって

$$F(x) = \int_{-\infty}^{x} f(z)dz = P(X \leq x) \tag{10.49}$$

と定義する．定義より $0 \leq F(x) \leq 1$ である．$F(x)$ が微分可能であるとすると，微分積分の基本定理から，密度関数と分布関数の間には次の関係が成り立つ．

$$F'(x) = \frac{dF(x)}{dx} = f(x), \qquad F(b) - F(a) = \int_a^b f(z)dz. \tag{10.50}$$

密度関数の定義より，$F'(x) = f(x) \geq 0$ なので，分布関数は単調非減少関数である．

■例：一様分布　区間 $[0,1]$ 上の**一様分布** (uniform distribution) の確率密度関数は

$$f(x) = \begin{cases} 1 & (x \in [0,1]) \\ 0 & (\text{otherwise}) \end{cases} \tag{10.51}$$

である．以下，区間 $[0,1]$ に限定して話を進める．$f(x)$ の原始関数は,

$$\int f(x)dx = \int 1dx = x + C \tag{10.52}$$

である．ここから，$F(x) = x + C$ であるが，$x = 0$ のとき，$F(0) = 0 = 0 + C$ ゆえに $C = 0$ なので，区間 $[0,1]$ における分布関数は $F(x) = x$ である．

$$\int_0^1 f(x)dx = F(1) - F(0) = 1 - 0 = 1$$
$$\int_{0.25}^{0.75} f(x)dx = F(0.75) - F(0.25) = 0.75 - 0.25 = 0.5$$

などである．確率密度関数と累積分布関数のグラフ上の関連を図 10.3 に示す．

■例：標準正規分布　　平均 0 分散 1 の標準正規分布の確率密度関数と累積分布関数はそれぞれ,

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \tag{10.53}$$

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left\{-\frac{z^2}{2}\right\} dz \tag{10.54}$$

である．覚える必要はない．推測統計学でよく用いる範囲としては

$$\int_{-1.96}^{1.96} f(x)dx = F(1.96) - F(-1.96) = 0.975 - 0.025 = 0.95$$
$$\int_{-2.576}^{2.576} f(x)dx = F(2.576) - F(2.576) = 0.995 - 0.005 = 0.99$$

がある．確率密度関数と累積分布関数のグラフ上の関連を図 10.4 に示す．

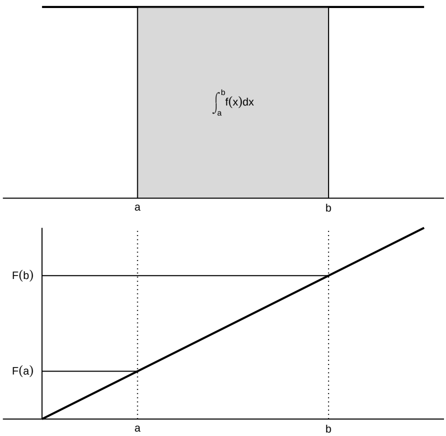

図 10.3　一様分布における確率密度関数と累積分布関数

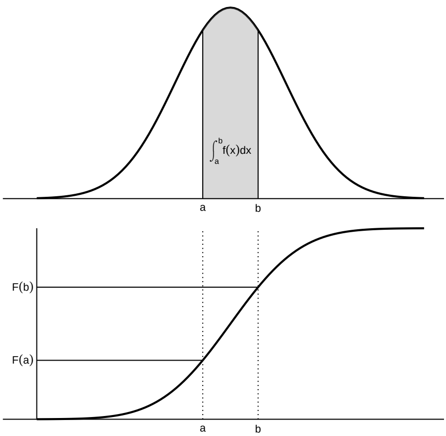

図 10.4　正規分布における確率密度関数と累積分布関数

第 11 章

指数と対数

この章では指数と対数の扱い方，その統計などでの使われ方を確認する．足し算がもとになる総和とは対照的に，指数はかけ算に関連し，対数は数の大小関係を保持したままかけ算を足し算にする．この性質が統計分析においてしばしば重要になる．

11.1 指数

11.1.1 指数法則

$a, b > 0$ なる実数 a, b と任意の実数 x, y について以下が成立する．

(1) $a^x a^y = a^{x+y}$	(2) $(a^x)^y = a^{xy}$
(3) $(ab)^x = a^x b^x$	(4) $a^{-x} = \dfrac{1}{a^x}$
(5) $\dfrac{a^x}{a^y} = a^{x-y}$	(6) $\left(\dfrac{a}{b}\right)^x = \dfrac{a^x}{b^x}$

これらの性質，とくに (1)–(3) を，（実数に拡張された）指数法則という．これらの性質は累乗の自然な性質から得られる．また，

$$a^0 = 1, \qquad a^{\frac{1}{2}} = \sqrt{a}, \qquad a^{\frac{m}{n}} = \sqrt[n]{a^m} = (\sqrt[n]{a})^m \quad (m, n \in \mathbb{Z}, n > 0)$$

が成立することにも注意しておく．

関連して，プロダクト記号も導入する．数列 $x_1, x_2, \cdots, x_n$ のすべての要素を掛け合わせることを総乗 (product) といって，

$$\prod_{i=1}^{n} x_i = x_1 \times x_2 \times \cdots \times x_n \tag{11.1}$$

と表す．$\prod$ はプロダクトの p に対応するギリシア文字のパイである．とくに，

$$\prod_{i=1}^{n} a^{x_i} = a^{x_1} \times a^{x_2} \times \cdots \times a^{x_n} = a^{\sum_{i=1}^{n} x_i}, \tag{11.2}$$

$$\prod_{i=1}^{n} a = a \times a \times \cdots \times a = a^n. \tag{11.3}$$

11.2　指数関数と対数関数

11.2.1　指数関数

$a > 0$ なる定数と実数値変数 x によって

$$f(x) = a^x \tag{11.4}$$

という関数を定義する．これを a を底とする**指数関数** (exponential function) という．指数の特性より，$f(x) > 0$．$a = 1$ のとき $f(x) = 1^x = 1$ となって定数関数となるので，$a \neq 1$ を仮定する．$a > 1$ のとき単調増加関数，$a < 1$ のとき単調減少関数となる．例として，$f(x) = 2^x, f(x) = (1/2)^x = 2^{-x}$ のグラフを示す (図 11.1).

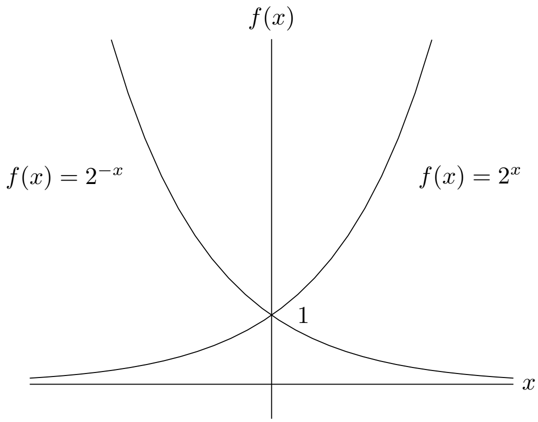

図 11.1　$f(x) = 2^x, f(x) = 2^{-x}$ のグラフ

11.2.2　対数関数

対数は指数の逆関数として定義される[*1].

$$x = a^y \iff y = \log_a x \tag{11.5}$$

という関係が成り立つとき，y を a を底とする x の**対数** (logarithm) という．底 a は正の実数で $a \neq 1$ でなければならない．**対数関数** (logarithm function)$f(x) = \log_a x$ は区間 $(0, \infty)$ で定義される．$a > 1$ のとき単調増加関数，$a < 1$ のとき単調減少関数となる (図 11.2, 11.3).

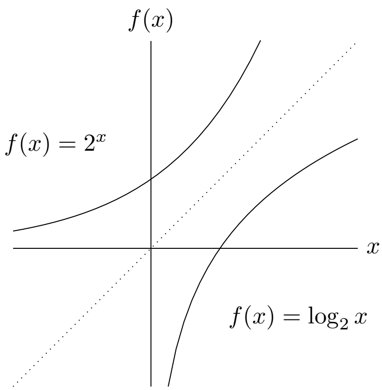

図 11.2　$f(x) = 2^x, f(x) = \log_2 x$ のグラフ

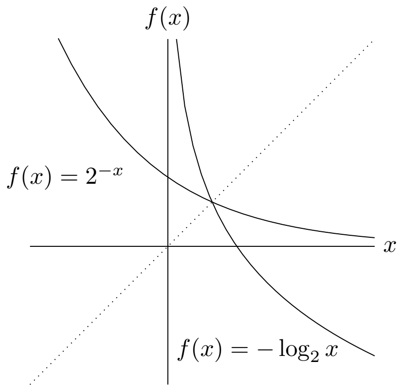

図 11.3　$f(x) = (1/2)^x, f(x) = \log_{1/2} x$ のグラフ

11.2.3　対数の特性

対数の特性として以下が成立する.

定理 11.1. 対数の底について $a > 0, b > 0, a \neq 1, b \neq 1$，また対数関数の引数は正であるとする．このとき以下が成立する.

[*1] $a > 0$ のとき，関数 $f(x) = a^x$ は $f : \mathbb{R} \to (0, \infty)$ において全単射であるので，付録 A.1 より，唯一の逆関数 $f^{-1} : (0, \infty) \to \mathbb{R}$ が存在する.

(1) $\log_a 1 = 0, \quad \log_a a = 1$

(2) $\log_a x_1 x_2 = \log_a x_1 + \log_a x_2$

(3) $\log_a \dfrac{x_1}{x_2} = \log_a x_1 - \log_a x_2$

(4) $\log_a x^p = p \log_a x$ 　　(p は任意の実数)

(5) $\log_b x = \dfrac{\log_a x}{\log_a b}$

証明. (1)

$$a^0 = 1 \Longleftrightarrow \log_a 1 = 0, \quad a^1 = a \Longleftrightarrow \log_a a = 1 \tag{11.6}$$

(2) $u = \log_a x_1, v = \log_a x_2$ とすると，

$$a^u = x_1, \quad a^v = x_2 \tag{11.7}$$

したがって

$$x_1 x_2 = a^u a^v = a^{u+v} \tag{11.8}$$

ここから

$$\log_a x_1 x_2 = u + v = \log_a x_1 + \log_a x_2. \tag{11.9}$$

(3) $u = \log_a x_1, v = \log_a x_2$ とすると，

$$\frac{x_1}{x_2} = \frac{a^u}{a^v} = a^{u-v} \tag{11.10}$$

ここから

$$\log_a \frac{x_1}{x_2} = u - v = \log_a x_1 - \log_a x_2. \tag{11.11}$$

(4) $u = \log_a x$ とすると，$x = a^u$ なので

$$x^p = (a^u)^p = a^{up} \tag{11.12}$$

ここから

$$\log_a x^p = up = p \log_a x. \tag{11.13}$$

(5) $u = \log_a b, v = \log_b x$ とすると，

$$a^u = b, \quad b^v = x \tag{11.14}$$

したがって

$$x = b^v = (a^u)^v = a^{uv} \tag{11.15}$$

ここから

$$\log_a x = uv = \log_a b \log_b x \tag{11.16}$$

より，両辺を $\log_a b$ で割ることで命題を得る． □

統計分析に関連する対数の特性で特に重要なのが (2) である．この公式は「積の対数は対数の和」になることを示している．とくに，底が $a > 1$ のとき対数関数は単調増加関数になる．単調増加関数とは，定義域集合における順序関係を値域集合においても保存する，ということでもある[*2]．つまり，対数関数は積のかたまりを，相対的な大小関係はそのままに，和のかたまりに分解することができる．数値計算において，積の計算は桁数が大きくなり人間にとってもコンピュータにとっても扱いが難しい．そこで，数値計算問題においてはしばしば「対数をとる」という操作が重要になる．

対数の特性より，とくに総乗について以下が成立する．

$$\log_a \left(\prod_{i=1}^{n} x_i \right) = \sum_{i=1}^{n} \log_a x_i \tag{11.17}$$

$$\log_a \left(\prod_{i=1}^{n} a^{x_i} \right) = \sum_{i=1}^{n} x_i \tag{11.18}$$

$$\log_a \left(\prod_{i=1}^{n} a \right) = n \tag{11.19}$$

[*2] 単調増加の定義については，付録 A.5.2 を参照．

11.3　対数関数の導関数

次に対数関数の導関数を求めよう．対数関数の導関数を考察するために，平均変化率を求めると，$x > 0$ のとき

$$\frac{\log_a(x+h) - \log_a x}{h} = \frac{1}{h}\log_a\left(\frac{x+h}{x}\right) = \frac{1}{x}\frac{x}{h}\log_a\left(1+\frac{h}{x}\right) \tag{11.20}$$

$$= \frac{1}{x}\log_a\left(1+\frac{h}{x}\right)^{\frac{x}{h}} \tag{11.21}$$

と変形することができる．ここで，$t = x/h$ とおいて，極限値

$$\lim_{h\to 0}\frac{1}{x}\log_a\left(1+\frac{h}{x}\right)^{\frac{x}{h}} = \lim_{t\to\infty}\frac{1}{x}\log_a\left(1+\frac{1}{t}\right)^t \tag{11.22}$$

を考える．ここで，$(1+1/t)^t$ は $h \to 0$，$t \to \infty$ のとき収束して，

$$\lim_{t\to\infty}\left(1+\frac{1}{t}\right)^t = 2.7182818285\ldots \tag{11.23}$$

という 1 つの実数に収束することが知られている．この実数を e と表すと，対数関数 $\log_a x$ の導関数は,

$$(\log_a x)' = \frac{1}{x}\log_a e \tag{11.24}$$

となる．特に e を底とする場合,

$$(\log_e x)' = \frac{1}{x} \tag{11.25}$$

である．e を底とする対数は数学的に重要であり，**自然対数** (natural logarithm) といわれ，単に log あるいは ln と表される．以下，本ノートで単に log とある場合は，すべて自然対数を示すと約束する．このとき，定数 e は**自然対数の底** (base of natural logarithm) といわれる[*3]．

[*3] **ネイピア数** (Napier’s number) といわれることもある．

ところで，自然対数の底は複利計算と密接に関連する．元本 A，年利 r，1 年間における利子の繰り入れ期間を m として，t 年後の元利合計を計算すると，

$$V = A\left(1+\frac{r}{m}\right)^{mt} \tag{11.26}$$
$$= A\left[\left(1+\frac{r}{m}\right)^{\frac{m}{r}}\right]^{rt} \tag{11.27}$$
$$= A\left[\left(1+\frac{1}{w}\right)^{w}\right]^{rt} \tag{11.28}$$

となる．ただし，$w = m/r$．$m \to \infty$ のとき $w \to \infty$ なので，連続的な利子の繰り入れを仮定したときの t 年後の元利合計は

$$\lim_{m\to\infty} V = Ae^{rt} \tag{11.29}$$

と近似できる[*4]．

次に，e を底とする指数関数

$$y = e^x \tag{11.30}$$

の導関数を求める．指数関数の特性より，定義域 $(-\infty, \infty)$ において$e^x > 0$ であるので，両辺の対数をとれば,

$$\log y = x \tag{11.31}$$

である．この両辺を x について微分すれば,

$$(\log y)' = 1 \Longleftrightarrow \frac{1}{y}y' = 1 \tag{11.32}$$

であるので，結局

$$(e^x)' = e^x \tag{11.33}$$

となる．これらの結果を定理としてまとめておこう.

定理 11.2. e を自然対数の底とするとき，以下のことが成立する.

[*4] さらに詳細は，A. C. チャン，2010,『現代経済学の数学基礎〈上〉』シーエーピー出版，10 章などを参照.

$$(1) \qquad (\log_e x)' = \frac{1}{x} \qquad (\text{ただし } x > 0 \text{ のとき})$$

$$(2) \qquad (e^x)' = e^x$$

さらに，定理 9.2 を一般化して累乗の微分について以下の定理が成立する．

定理 11.3. α を任意の実数とするとき，定義域 $(0, \infty)$ において

$$(x^\alpha)' = \alpha x^{\alpha-1} \tag{11.34}$$

が成立する．

証明. $y = x^\alpha$ の両辺の対数をとって，$\log y = \alpha \log x$ とした上で，両辺を x について微分すれば，

$$\frac{1}{y}y' = \alpha\frac{1}{x} \Longleftrightarrow y' = \alpha\frac{y}{x} = \alpha x^{\alpha-1}. \tag{11.35}$$

□

練習問題 11.3

次の関数を微分せよ（矢野・田代 (1993: 85) より）．

(1) $\log x^3 \quad (x > 0)$　　(2) $x \log x \quad (x > 0)$　　(3) $\log(x^2 + 1)$

(4) e^{-2x}　　(5) $(x - 1)e^x$　　(6) $\log(1 + e^x)$

11.4　対数関数に関連する積分

対数関数に関連する積分法則を確認する．

定理 11.2 より以下の公式が成立する．

定理 11.4.

$$(1) \int \frac{1}{x}dx = \log|x| + C \qquad (\text{ただし，} x \neq 0 \text{ のとき})$$

$$(2) \int e^x dx = e^x + C$$

証明. (2) は自明である．(1) に対して証明を与えておく．$x > 0$ のとき，定理 11.2 より

$$(\log x)' = \frac{1}{x} \tag{11.36}$$

であるので

$$\int \frac{dx}{x} = \log x + C. \tag{11.37}$$

$x < 0$ のとき，関数 $\log(-x)$ を合成関数の微分法を用いて微分すると，$u = -x$ として，

$$(\log(-x))' = \frac{d}{du} \log u \cdot \frac{du}{dx} = \frac{1}{u} \cdot (-1) = \frac{1}{x} \tag{11.38}$$

なので

$$\int \frac{dx}{x} = \log(-x) + C. \tag{11.39}$$

絶対値の記号を用いてまとめると (1) を得る． □

定理 11.3 より，以下の定理が成立する．

定理 11.5. α を $\alpha \neq -1$ を満たす任意の実数とするとき，定義域 $(0, \infty)$ において

$$\int x^\alpha dx = \frac{1}{\alpha+1} x^{\alpha+1} + C \tag{11.40}$$

が成立する．

練習問題 11.4

次の不定積分を求めよ（松坂 (1989-91: 1012) より）．

(1) $\displaystyle\int 3\sqrt{x}dx \quad (x > 0)$　　(2) $\displaystyle\int \frac{(\sqrt{x}-1)^2}{x}dx \quad (x > 0)$

(3) $\displaystyle\int \frac{(x+1)^2}{x^2}dx$　　(4) $\displaystyle\int (4e^x + 5)dx$

11.5　二項分布のパラメータの推定

成功確率が p であるベルヌーイ試行を独立に n 回行うとき，x 回成功する確率は二項分布 $Bi(n,p)$ に従い，その場合の確率関数は

$$P(X=x)=f(x)={}_nC_x p^x(1-p)^{n-x}, \qquad x=0,1,\cdots,n \tag{11.41}$$

で与えられる．いま実際に，試行を行い x_o 回成功という結果を得た．これをデータとして二項分布のパラメータ p を推定したい．これが点推定の問題である．

ひとつの考え方として，結果 x_o を所与として，確率関数におけるパラメータ p の値を変えていったとき，結果 x_o を出力する確率 $P(X=x_o)$ が最も高くなる p を「最も尤もらしい」値として採用することを考えよう．これは，**最尤推定法** (maximum likelihood estimation method) の一つの例である．

このことを調べるために，確率関数 $f(x)$ を p を引数とする関数（尤度関数）と見なす．すなわち，

$$L(p)={}_nC_{x_o}p^{x_o}(1-p)^{n-x_o}, \tag{11.42}$$

この関数について，p の定義域における最大値を求めることが問題である．いま単純化のために $p\in(0,1)$ とする．まず，尤度関数を対数変換すると

$$\log L(p)=\log{}_nC_{x_o}+x_o\log p+(n-x_o)\log(1-p) \tag{11.43}$$

となる（対数尤度関数）．

このとき p について対数尤度関数を最大化する 1 階の必要条件は，p で微分して 0 と等しくなればよいので，

$$\frac{d\log L}{dp}=\frac{x_o}{p}-\frac{n-x_o}{1-p}=0$$

より，

$$\frac{x}{p}=\frac{n-x_o}{1-p}\Longleftrightarrow p=\frac{x_o}{n}$$

である[*5]．

[*5] 細かくは，2 階の十分条件，つまり 2 階微分して極点を代入したときに負となるという条件も確認

■コイン投げの確率の最尤推定　コインを 10 回投げたとき 7 回表が出た．表の出現確率である母数 p の尤度関数は

$$L(p) = {}_{10}C_7 p^7 (1-p)^3 = 120 p^7 (1-p)^3$$

であり（図 11.4），対数尤度関数は

$$\log L(p) = \log 120 + 7 \log p + 3 \log(1-p)$$

である（図 11.5）．p の最大化問題を解くと，

$$p = \frac{7}{10} = 0.7$$

を得る．

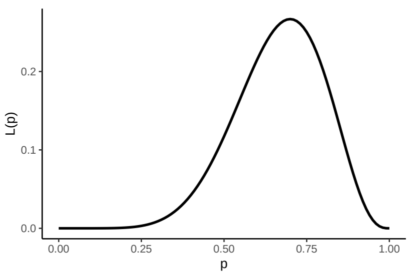

図 11.4　尤度関数 $L(p)$

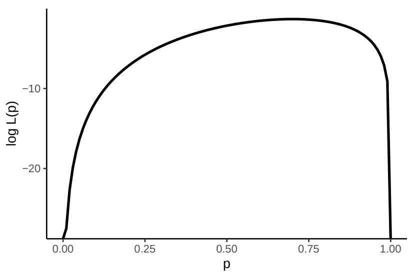

図 11.5　対数尤度関数 $\log L(p)$

しなければならない．

$$\left.\frac{d^2 \log L}{dp^2}\right|_{p=\frac{x_o}{n}} = -\frac{n^3}{(n-x_o)x_o} < 0$$

なので，$0 < p < 1$ のとき，$p = x_o/n$ は極大値である．

第 12 章

ベクトル

高度な統計手法，データサイエンスの方法をさらに学ぶためには，線形代数の知識が必須になる．ここではその道筋を示すために，岡太 (2008: 2 章) を主に参考にしながら，ごくごく初歩レベルのベクトルの知識と計算方法を導入し，共分散・相関係数といった統計概念との関連を示す．さらに次章では，行列について同様に初歩的知識を導入する．

12.1　ベクトルの基本

n 個の数 $a_1, a_2, \cdots, a_n$ を順番に縦に並べたものを n 次元 (列) ベクトル ((column) vector) といって，

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \tag{12.1}$$

と表す．ベクトルは一般的に太字のアルファベットの小文字で表し，その要素（成分）は添え字つきの太字でないアルファベットの小文字で表す．この文脈で太字でないアルファベットはベクトルではない数（スカラー）を表すものとする．

集合とは異なり，ベクトルには要素の順番に意味があるので，順番が入れ替わると一般に異なるベクトルになる．

列ベクトルを横に倒したものを行ベクトル (row vector) という．行と列を入れ替

える操作を**転置** (transpose) といって $\mathbf{a}^{\mathrm{T}}$ で表す[*1]．すなわち，

$$\mathbf{a}^{\mathrm{T}} = [a_1 \ a_2 \ \cdots \ a_n]. \tag{12.2}$$

基本は列ベクトルであるが，表記上スペースを取るので列ベクトルを

$$\mathbf{a} = [a_1 \ a_2 \ \cdots \ a_n]^{\mathrm{T}}$$

と表すことがある．

12.2　ベクトルの演算

2 つの n 次元列ベクトル

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \tag{12.3}$$

は，すべての i 番目の要素が等しいとき，つまり，

$$\forall i \in \{1, 2, \cdots, n\}, a_i = b_i \tag{12.4}$$

のとき，そのときに限り等しいとする．

すべての要素が 0 であるベクトルを**零ベクトル** (zero vector) といって

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{12.5}$$

と表す．

2 つの n 次元列ベクトル $\mathbf{a}, \mathbf{b}$ の和は

$$\mathbf{a} + \mathbf{b} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix} \tag{12.6}$$

[*1] テキストによってはダッシュで $\mathbf{a}'$ と表す場合や ${}^t\mathbf{a}$ と表す場合もある．

であり，n 次元列ベクトル $\mathbf{a}$ とスカラー α の積（スカラー倍）は，

$$\alpha\mathbf{a} = \alpha\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} \alpha a_1 \\ \alpha a_2 \\ \vdots \\ \alpha a_n \end{bmatrix} \tag{12.7}$$

である．ここから，2 つの n 次元列ベクトル $\mathbf{a}, \mathbf{b}$ の差は

$$\mathbf{a} - \mathbf{b} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + (-1)\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 - b_1 \\ a_2 - b_2 \\ \vdots \\ a_n - b_n \end{bmatrix} \tag{12.8}$$

となる．

ベクトルのスカラー倍と和について，以下の法則が成り立つ．

(1) 交換法則

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} \tag{12.9}$$

(2) 結合法則

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}) \tag{12.10}$$

(3) 分配法則 1

$$(\alpha_1 + \alpha_2)\mathbf{a} = \alpha_1\mathbf{a} + \alpha_2\mathbf{a} \tag{12.11}$$

(4) 分配法則 2

$$\alpha(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b} \tag{12.12}$$

いくつかのベクトル（たとえば，$\mathbf{a}, \mathbf{b}, \mathbf{c}$）について，スカラー倍したベクトルの和（たとえば，$\mathbf{d} = \alpha\mathbf{a} + \beta\mathbf{b} + \gamma\mathbf{c}$）からなるベクトルをベクトルの一次結合もしくは線形結合 (linear combination) という．

練習問題 12.2

$$\mathbf{a} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ 4 \\ -5 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

であるとき，次の式を計算せよ（矢野・田代 (1993: 3) より）．

(1) $\mathbf{a} + 3\mathbf{b}$　　(2) $\mathbf{a} - \mathbf{b} + 2\mathbf{c}$　　(3) $4\mathbf{a} + 2\mathbf{b} - 6\mathbf{c}$

12.3　ベクトルの幾何学的表現

ベクトルを直交座標上の点の座標を表すものとして，原点から点までの矢印を引くことで，ベクトルを幾何学的に表現できる[*2]．人間が理解できるのはおそらく 2 次元（平面）と 3 次元（空間）座標であろう．TeX で 3 次元の絵を描くのは面倒なため，ここでは 2 次元座標上のみ例示する．

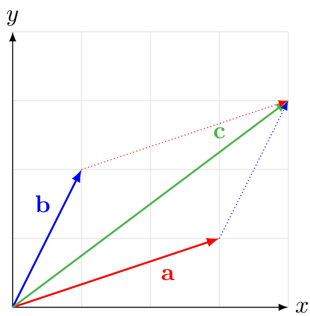

図 12.1　平面上のベクトルの和

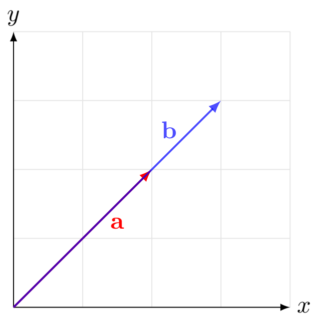

図 12.2　平面上のベクトルのスカラー倍

図 12.1 は，ベクトル $\mathbf{a} = [3\ 1]^{\mathrm{T}}$ と $\mathbf{b} = [1\ 2]^{\mathrm{T}}$ の和が，2 つのベクトルが作る平行四辺形の対角線であるベクトル $\mathbf{c} = [4\ 3]^{\mathrm{T}}$ となることを示している．

[*2] 点 P の原点 O に関する位置ベクトルともいう．

図 12.2 は，ベクトル $\mathbf{a} = [2\ 2]^{\mathrm{T}}$ の 1.5 倍がベクトル $\mathbf{b} = [3\ 3]^{\mathrm{T}}$ となることを示している．

一般に，

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{12.13}$$

を 2 次元ベクトル空間における**単位ベクトル** (unit vector) とすると，2 次元空間上の任意のベクトルは，単位ベクトルの線形結合で表すことができる．

なお，注意点として，ベクトルを幾何学的に表現する場合は，方向と長さの等しいベクトルは，その位置に関係なく同じベクトルであると見なす．

12.4　内積とノルム

12.4.1　定義

2 つの n 次元ベクトル $\mathbf{a}, \mathbf{b}$ の**内積** (inner product) を

$$\mathbf{a} \cdot \mathbf{b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \tag{12.14}$$

と定義する[*3]．

定義より以下の法則が成り立つ．

(1) **交換法則**

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a} \tag{12.15}$$

(2) **分配法則**

$$(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c} \tag{12.16}$$

(3) **定数倍**

$$(\alpha \mathbf{a}) \cdot \mathbf{b} = \alpha (\mathbf{a} \cdot \mathbf{b}) \tag{12.17}$$

[*3] ここでは，詳説しないが外積という概念もある．

n 次元ベクトル $\mathbf{a}$ について自分自身との内積の正の平方根を**ノルム** (norm) といって

$$\|\mathbf{a}\| = \sqrt{\mathbf{a}\cdot\mathbf{a}} = \sqrt{\sum_{i=1}^{n} a_i^2} = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2} \tag{12.18}$$

と表す．幾何学的にはノルムは原点からベクトル $\mathbf{a}$ に対応する点までのユークリッド距離を表す．

12.4.2　幾何学的意味

ベクトル $\mathbf{a}, \mathbf{b}$ とベクトル $\mathbf{a}-\mathbf{b}$（$\mathbf{b}$ の終点を始点として移動したもの）を辺とする三角形を考え，$\mathbf{a}$ と $\mathbf{b}$ のなす角を θ とする（図 12.3）．このとき余弦定理[*4]によって，

$$\begin{aligned} 2\|\mathbf{a}\|\|\mathbf{b}\|\cos\theta &= \|\mathbf{a}\|^2 + \|\mathbf{b}\|^2 - \|\mathbf{a}-\mathbf{b}\|^2 \\ &= \sum_{i=1}^{n} a_i^2 + \sum_{i=1}^{n} b_i^2 - \sum_{i=1}^{n} a_i^2 + 2\sum_{i=1}^{n} a_i b_i - \sum_{i=1}^{n} b_i^2 \\ &= 2\sum_{i=1}^{n} a_i b_i \end{aligned} \tag{12.19}$$

なので，結局

$$\mathbf{a}\cdot\mathbf{b} = \|\mathbf{a}\|\|\mathbf{b}\|\cos\theta \tag{12.20}$$

となる．

ここで，$\mathbf{b}$ から $\mathbf{a}$ 上へ垂直に引かれた線の足を終点とするベクトルを $\mathbf{b}'$ とする（図 12.4）．これを正射影（直交射影）ベクトルという．コサインの定義より

$$\cos\theta = \frac{\|\mathbf{b}'\|}{\|\mathbf{b}\|} \tag{12.21}$$

なので，$\mathbf{b}'$ のノルムは

$$\|\mathbf{b}'\| = \|\mathbf{b}\|\cos\theta \tag{12.22}$$

であり，内積 $\mathbf{a}\cdot\mathbf{b}$ は $\mathbf{a}$ のノルムと正射影ベクトル $\mathbf{b}'$ のノルムの積であるといえる．

[*4] ここでは詳細説明しませんが，気になる人はググってみてください．三角関数のところで出てくる定理です．

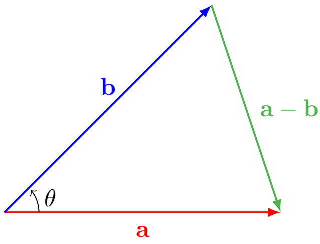

図 12.3　ベクトル $\mathbf{a}, \mathbf{b}$ のなす角

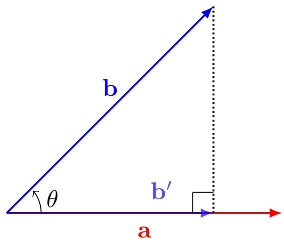

図 12.4　内積の幾何学的表現

内積の幾何学的表現から，ただちに次のことが分かる．$\mathbf{a}, \mathbf{b}$ のノルムを所与としたとき，$\mathbf{a}, \mathbf{b}$ のなす角が直角であるとき $\cos(\pi/2) = 0$ なので内積も 0 である．これを**直交** (orthogonal) という．この概念は高度な統計解析を理解する際に重要なものとして出会うことになる．

直交の状態から $\mathbf{b}$ が $\mathbf{a}$ に近づき同じ方向に向くほどプラスに大きくなり，同一線上で同じ方向を向いているとき最大となる．逆に反対方向に向くほどマイナスに小さな値になり，同一線上で反対方向を向いているとき最少になる．

練習問題 12.4

$$\mathbf{a} = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 5 \\ 4 \\ -5 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$ であるとき，次の式を計算せよ．

(1) $\|\mathbf{a}\|^2$　　(2) $\mathbf{a} \cdot \mathbf{b}$　　(3) $\mathbf{b} \cdot \mathbf{c}$

12.5　共分散と相関係数

ここで，記述統計の範囲で共分散と相関係数の概念を確認しよう．

2 変数のデータを考える．例えば，それぞれの個人に身長と体重を尋ねたとすると，それぞれの個人について観測値の組 $(\text{身長}_i, \text{体重}_i)$ が求まる．これを対象者分集めたものが 2 変数データである．

2 つの変数の結びつきを調べる上でもっとも基本となるのが，**共分散** (covariance)

である．変数 x と y の共分散を s_{xy} で表すと，

$$s_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \tag{12.23}$$
$$= \frac{(x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + \cdots + (x_n - \bar{x})(y_n - \bar{y})}{n}$$

と定義される．1 変数の分散は非負の値しか取らないが，共分散はマイナスにもなることに注意．2 つの変数の関連が強いほど共分散はプラスもしくはマイナス方向に大きくなり，関連がなけれなないほど 0 に近づく．

x と y の標準得点について共分散をとったものを**相関係数** (correlation coefficient) という．

$$z_i = \frac{x_i - \bar{x}}{s_x}, \qquad w_i = \frac{y_i - \bar{y}}{s_y}$$

とすると，相関係数 r_{xy} は

$$r_{xy} = \frac{1}{n}\sum_{i=1}^{n} z_i w_i \tag{12.24}$$
$$= \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right) = \frac{s_{xy}}{s_x s_y} \tag{12.25}$$

と定義される．

相関係数は -1 から 1 の間の値をとり，$|r_{xy}| = 1$ のときそのときに限り，$y_i = ax_i + b$ という線形の関係がすべての i について成り立っている．そして，$|r_{xy}|$ が 1 に近づけば近づくほど，線形の関係に近づく．詳細は付録 A.8 節を参照．

■例（身長と体重）　x を身長，y を体重とする．データは表 12.1，散布図（観測値ベクトルをプロットした図）は図 12.5 である．

表 12.1　身長と体重データ

個人 i	1	2	3	4	5	6	7	8	9	10
身長 x	152.8	150.1	182	163.2	167.3	160.2	164.9	161.4	179.9	172.2
体重 y	56.3	52.1	85.6	66.8	74.2	58.1	61.9	55.1	70.5	64.1

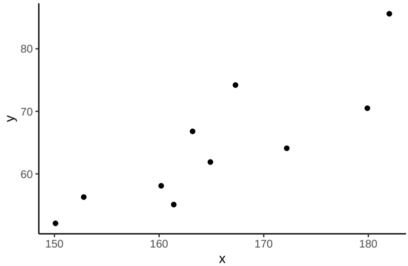

図 12.5　身長と体重の散布図

$$s_{xy} = 81.313, \qquad s_x = 9.882, \qquad s_y = 9.684$$
$$r_{xy} = \frac{81.313}{9.882 \times 9.684} = 0.850$$

身長と体重の関連はかなり強いと言える．

12.6　共分散と相関係数のベクトルによる理解

2 変数の n 個のデータ $x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_n$ について，それぞれの値の平均値を $\bar{x}, \bar{y}$ としたとき，各要素と平均の差で構成されるベクトル $\mathbf{x}, \mathbf{y}$ を考える（この処理をセンタリングということがある）．すなわち，

$$\mathbf{x} = \begin{bmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{bmatrix}, \qquad \mathbf{y} = \begin{bmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ \vdots \\ y_n - \bar{y} \end{bmatrix}. \tag{12.26}$$

このとき，$\mathbf{x}, \mathbf{y}$ のそれぞれのノルムは，

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} = \sqrt{n} s_x, \qquad \|\mathbf{y}\| = \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2} = \sqrt{n} s_y \tag{12.27}$$

となるので，それぞれの変数の標準偏差の $\sqrt{n}$ 倍である．

また，$\mathbf{x}, \mathbf{y}$ の内積は

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) = ns_{xy} \tag{12.28}$$

であるので，共分散の n 倍になる．

最後に，内積をそれぞれのノルムを掛けたもので割れば，$\cos\theta$ なので，

$$\cos\theta = \frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} = \frac{ns_{xy}}{\sqrt{n}s_x\sqrt{n}s_y} = r_{xy} \tag{12.29}$$

つまり，変数 x と y の相関係数は，ベクトル $\mathbf{x}$ の $\mathbf{y}$ のなす角の余弦と見なすことができる．ここから，$-1 \leq r_{xy} \leq 1$ の特性が直ちに言える．また，$y_i = ax_i + b$ という線形の関係が成り立っているとき，

$$\mathbf{y} = \begin{bmatrix} ax_1 + b - (a\bar{x} + b) \\ ax_2 + b - (a\bar{x} + b) \\ \vdots \\ ax_n + b - (a\bar{x} + b) \end{bmatrix} = \begin{bmatrix} a(x_1 - \bar{x}) \\ a(x_2 - \bar{x}) \\ \vdots \\ a(x_n - \bar{x}) \end{bmatrix} = a\mathbf{x} \tag{12.30}$$

であり，$\|\mathbf{y}\| = \sqrt{a^2\mathbf{x} \cdot \mathbf{x}} = |a|\|\mathbf{x}\|$ に注意しておくと，

$$r_{xy} = \cos\theta = \frac{a\mathbf{x} \cdot \mathbf{x}}{|a|\|\mathbf{x}\|\|\mathbf{x}\|} = \frac{a}{|a|} = \begin{cases} 1 & (a > 0) \\ -1 & (a < 0) \end{cases} \tag{12.31}$$

であることがわかる．

第 13 章

行列

前章に続いて，線形代数への入り口として，岡太 (2008: 4 章) を主に参考にしながら，初歩の初歩レベルの行列の知識と計算方法を導入し，最終的には分散共分散行列，相関行列の理解を目指す．

13.1　行列の基本

$m \times n$ 個の数 $a_{11}, a_{12}, \cdots, a_{1n}, a_{21}, a_{22}, \cdots, a_{2n}, a_{m1}, a_{m2}, \cdots, a_{mn}$ を順番に並べたものを $m \times n$ 次元**行列** (matrix) といって，

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{13.1}$$

と表す．行列は一般的にアルファベットの大文字で表し，その要素（成分）は添え字つきのアルファベットの小文字で表す．ここでは，行列をアルファベットの太字の大文字で表す．紙幅の節約のため，一般的な要素を代表させて行列を

$$\mathbf{A} = [a_{ij}] \tag{13.2}$$

と表すこともある．行列の横方向の配列が**行** (row) で上から第 1 行などと数える．縦方向の配列が**列** (columun) で左から第 1 列などと数える．

社会学などで普通扱うクロスセクショナル・データを行列で表すと，行がケース（回答者）で列が変数を表すことが一般的である．このようなデータ形式を**整然デー**

タ (tidy data) という[*1].

行列はまた

$$\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_n], \qquad \mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \tag{13.3}$$

というように，列ベクトルを並べた行ベクトルと見なすことができる．逆に，行ベクトルを並べた列ベクトルと見なすこともできる．

行列の行と列を入れ替える操作を**転置** (transpose) といって $\mathbf{A}^{\mathrm{T}}$ で表す．すなわち，

$$\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_n] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \tag{13.4}$$

$$\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} \mathbf{a}_1^{\mathrm{T}} \\ \mathbf{a}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{a}_n^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} \tag{13.5}$$

である．ここで $m \times n$ 行列を転置すると $n \times m$ 行列になり，行数と列数が入れ替わることに注意すること．

13.2　行列の演算

2 つの $m \times n$ 次元行列 $\mathbf{A} = [a_{ij}], \mathbf{B} = [b_{ij}]$ は，すべての ij 番目の要素が等しいとき，つまり，

$$\forall i \in \{1, 2, \cdots, n\}, \forall j \in \{1, 2, \cdots, m\}, a_{ij} = b_{ij} \tag{13.6}$$

のとき，そのときに限り等しいとする．

[*1] Hadley Wickham & Garrett Grolemund, *R for Data Science*, `https://r4ds.had.co.nz/`.

すべての要素が 0 である行列を**零行列** (zero matrix) といって

$$\mathbf{O} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \tag{13.7}$$

と表す．とくに行数列数を明示するときは $\mathbf{O}_{m\times n}$ などと表す.

行数 m と列数 n が等しい行列を**正方行列** (square matrix) という．$n \times n$ 次元正方行列のなかでも，行列の対角要素 a_{ii} が 1 で非対角要素 $a_{ij}(i \neq j)$ が 0 の行列を $n \times n$ 次元**単位行列** (identity matrix) と呼ばれ

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{13.8}$$

と表す．とくに次元数を明示するときは $\mathbf{I}_n$ と表す.

2 つの $m \times n$ 次元行列 $\mathbf{A} = [a_{ij}], \mathbf{B} = [b_{ij}]$ の和は

$$\begin{aligned} \mathbf{a} + \mathbf{b} &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix} \end{aligned} \tag{13.9}$$

であり，$m \times n$ 次元行列 $\mathbf{A} = [a_{ij}]$ とスカラー α の積（スカラー倍）は,

$$\alpha\mathbf{A} = \alpha \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{bmatrix} \tag{13.10}$$

である．ここから，2 つの $m \times n$ 次元行列 $\mathbf{A} = [a_{ij}], \mathbf{B} = [b_{ij}]$ の差は

$$\mathbf{A} - \mathbf{B} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + (-1) \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} - b_{11} & a_{12} - b_{12} & \cdots & a_{1n} - b_{1n} \\ a_{21} - b_{21} & a_{22} - b_{22} & \cdots & a_{2n} - b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} - b_{m1} & a_{m2} - b_{m2} & \cdots & a_{mn} - b_{mn} \end{bmatrix} \tag{13.11}$$

となる．

ベクトルと同じく行列のスカラー倍と和について，以下の法則が成り立つ．

(1) 交換法則

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \tag{13.12}$$

(2) 結合法則

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) \tag{13.13}$$

(3) 分配法則 1

$$(\alpha_1 + \alpha_2)\mathbf{A} = \alpha_1\mathbf{A} + \alpha_2\mathbf{A} \tag{13.14}$$

(4) 分配法則 2

$$\alpha(\mathbf{A} + \mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B} \tag{13.15}$$

練習問題 13.2

$\mathbf{A} = \begin{bmatrix} 2 & 1 & -3 \\ 4 & -6 & 5 \end{bmatrix}$,　$\mathbf{B} = \begin{bmatrix} 1 & 3 & 4 \\ 3 & 0 & -1 \end{bmatrix}$ について，次の式を計算せよ（矢野・田代 (1993: 5) より）．

(1) $\mathbf{A} + \mathbf{B}$　(2) $\mathbf{A} - \mathbf{B}$

(3) $3\mathbf{A}$　(4) $2\mathbf{A} + 3\mathbf{B}$

13.3　行列の積

行列の積は $m \times n$ 次元の行列 $\mathbf{A}$ と $n \times p$ 次元の行列 $\mathbf{B}$ の間で定義され,

$$\begin{aligned}\mathbf{AB} &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{k=1}^n a_{1k}b_{k1} & \sum_{k=1}^n a_{1k}b_{k2} & \cdots & \sum_{k=1}^n a_{1k}b_{kp} \\ \sum_{k=1}^n a_{2k}b_{k1} & \sum_{k=1}^n a_{2k}b_{k2} & \cdots & \sum_{k=1}^n a_{2k}b_{kp} \\ \vdots & \vdots & & \vdots \\ \sum_{k=1}^n a_{mk}b_{k1} & \sum_{k=1}^n a_{mk}b_{k2} & \cdots & \sum_{k=1}^n a_{mk}b_{kp} \end{bmatrix} \end{aligned} \tag{13.16}$$

となる．行列の積 $\mathbf{AB}$ は $m \times p$ 次元の行列になることに注意.

一般的に行列の積は，左の項の列数と右の項の行数が一致している場合にのみ定義される．ゆえに，$\mathbf{AB}$ が求められるときに，$\mathbf{BA}$ が求められるとは限らない．また，$\mathbf{BA}$ が求められるときにも，$\mathbf{AB}$ と一致するとは限らない.

行列の積を行列を構成する行ベクトル，列ベクトルとして理解しよう．行列 $\mathbf{A}$ を m 個の n 次元ベクトル $\mathbf{a}_i(i = 1, 2, \cdots, m)$ の転置（行ベクトル）を積み重ねた列ベクトルとみなす．同様に，行列 $\mathbf{B}$ を p 個の n 次元ベクトル $\mathbf{b}_j(j = 1, 2, \cdots, p)$ を横に並べた行ベクトルとみなす．すなわち,

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^{\mathrm{T}} \\ \mathbf{a}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{a}_m^{\mathrm{T}} \end{bmatrix}, \qquad \mathbf{B} = [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_p]. \tag{13.17}$$

このとき，行列の積 $\mathbf{AB}$ は，$\mathbf{a}_i$ と $\mathbf{b}_j$ の内積 $\mathbf{a}_i \cdot \mathbf{b}_j$ を ij 要素とする $m \times p$ 次元の行列として定義される.

$$\mathbf{AB} = \begin{bmatrix} \mathbf{a}_1^{\mathrm{T}} \\ \mathbf{a}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{a}_m^{\mathrm{T}} \end{bmatrix} [\mathbf{b}_1 \ \mathbf{b}_2 \ \cdots \ \mathbf{b}_p]$$

$$= \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{b}_1 & \mathbf{a}_1 \cdot \mathbf{b}_2 & \cdots & \mathbf{a}_1 \cdot \mathbf{b}_p \\ \mathbf{a}_2 \cdot \mathbf{b}_1 & \mathbf{a}_2 \cdot \mathbf{b}_2 & \cdots & \mathbf{a}_2 \cdot \mathbf{b}_p \\ \vdots & \vdots & & \vdots \\ \mathbf{a}_m \cdot \mathbf{b}_1 & \mathbf{a}_m \cdot \mathbf{b}_2 & \cdots & \mathbf{a}_m \cdot \mathbf{b}_p \end{bmatrix} \tag{13.18}$$

ここから，特殊例として $1 \times n$ 次元行列と $n \times 1$ 次元行列の行列の積と，それぞれを n 次元ベクトルとみた場合の内積を同一視してよいことが分かる．

行列の積について以下の法則が成立する．

(1) 結合法則

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) \tag{13.19}$$

(2) 分配法則

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC} \tag{13.20}$$

(3) 転置

$$(\mathbf{AB})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}} \tag{13.21}$$

(4) 単位行列： $m \times n$ 行列 $\mathbf{A}$ について

$$\mathbf{AI}_n = \mathbf{I}_m\mathbf{A} = \mathbf{A} \tag{13.22}$$

(5) 零行列： $m \times n$ 行列 $\mathbf{A}$ について

$$\mathbf{AO}_{n\times p} = \mathbf{O}_{m\times p}, \quad \mathbf{O}_{p\times m}\mathbf{A} = \mathbf{O}_{p\times n} \tag{13.23}$$

練習問題 13.3

次の積を求めよ（矢野・田代 (1993: 7) より）．

(1) $\begin{bmatrix} 2 & 1 & -3 \\ 4 & -6 & 5 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 3 & 0 \\ 2 & 4 \end{bmatrix}$　　(2) $\begin{bmatrix} -1 & 3 \\ 3 & 0 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 & -3 \\ 4 & -6 & 5 \end{bmatrix}$

(3) $\begin{bmatrix} 3 & 5 & -4 \\ 1 & 2 & 1 \\ -6 & 3 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$　　(4) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 4 & -1 \\ 5 & -2 \end{bmatrix}$

13.4　行列の積の幾何学的表現

行列の積には具体的にどのような意味があるだろうか．$n \times n$ 次元正方行列 $\mathbf{A}$ と n 次元列ベクトル（$n \times 1$ 次元行列）$\mathbf{x}$ との積は n 次元列ベクトル $\mathbf{y}$ を出力する．つまり，

$$\mathbf{y} = \mathbf{A}\mathbf{x}. \tag{13.24}$$

このことは，正方行列 $\mathbf{A}$ は，$\mathbf{x}$ を $\mathbf{y}$ に写す写像 $\mathbf{A} : \mathbf{x} \to \mathbf{y}$ の働きをしていることを意味する．さらに，この写像は線形性，つまり

$$\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y} \tag{13.25}$$

$$\mathbf{A}(\alpha\mathbf{x}) = \alpha\mathbf{A}\mathbf{x} \tag{13.26}$$

という特性を持つものであり，**線形写像** (linear mapping) といわれるものである．

たとえば

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \tag{13.27}$$

とし，$\mathbf{A}$ を $\mathbf{x} = [1\ 1]^{\mathrm{T}}$ に作用させると，

$$\mathbf{A}\mathbf{x} = \begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \mathbf{y} \tag{13.28}$$

というように $\mathbf{x} = [1\ 1]^{\mathrm{T}}$ を $\mathbf{y} = [3\ 2]^{\mathrm{T}}$ に写している（図 13.1）．

また，たとえば

$$[\mathbf{x}_1\ \mathbf{x}_2] = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \tag{13.29}$$

に $\mathbf{A}$ を作用させると

$$\begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & 3 \end{bmatrix} = [\mathbf{y}_1\ \mathbf{y}_2] \tag{13.30}$$

となるので，$\mathbf{x}_1 = [1\ 1]^{\mathrm{T}}$ を $\mathbf{y}_1 = [3\ 2]^{\mathrm{T}}$ に，$\mathbf{x}_2 = [0\ 1]^{\mathrm{T}}$ を $\mathbf{y}_1 = [1\ 3]^{\mathrm{T}}$ にそれぞれ写していることになる（図 13.2）．

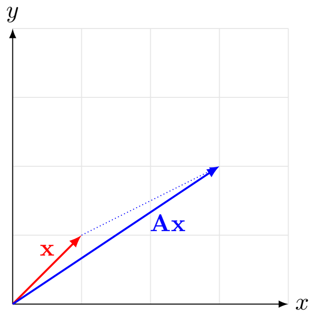

図 13.1　$\mathbf{A}$ による $\mathbf{x}$ の写像

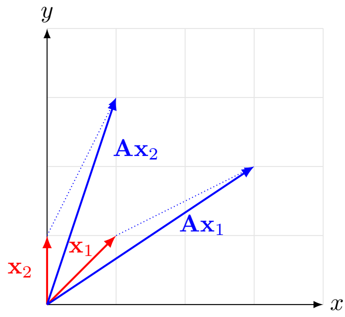

図 13.2　$\mathbf{A}$ による $[\mathbf{x}_1, \mathbf{x}_2]$ の写像

13.5　分散共分散行列の導出

$n \times p$ 次元行列 $\mathbf{X} = [\mathbf{x}_1 \ \cdots \ \mathbf{x}_p]$ を転置すると，$p \times n$ 次元行列 $\mathbf{X}^{\mathrm{T}}$ になる．これを $\mathbf{X}$ の左から掛けると，行列 $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ は $p \times p$ 次元の正方行列になる．具体的には，

$$\begin{aligned}
\mathbf{X}^{\mathrm{T}}\mathbf{X} &= \begin{bmatrix} \mathbf{x}_1^{\mathrm{T}} \\ \mathbf{x}_2^{\mathrm{T}} \\ \vdots \\ \mathbf{x}_p^{\mathrm{T}} \end{bmatrix} [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_p] \\
&= \begin{bmatrix} \mathbf{x}_1 \cdot \mathbf{x}_1 & \mathbf{x}_1 \cdot \mathbf{x}_2 & \cdots & \mathbf{x}_1 \cdot \mathbf{x}_p \\ \mathbf{x}_2 \cdot \mathbf{x}_1 & \mathbf{x}_2 \cdot \mathbf{x}_2 & \cdots & \mathbf{x}_2 \cdot \mathbf{x}_p \\ \vdots & \vdots & & \vdots \\ \mathbf{x}_p \cdot \mathbf{x}_1 & \mathbf{x}_p \cdot \mathbf{x}_2 & \cdots & \mathbf{x}_p \cdot \mathbf{x}_p \end{bmatrix} \\
&= \begin{bmatrix} \|\mathbf{x}_1\|^2 & \mathbf{x}_1 \cdot \mathbf{x}_2 & \cdots & \mathbf{x}_1 \cdot \mathbf{x}_p \\ \mathbf{x}_2 \cdot \mathbf{x}_1 & \|\mathbf{x}_2\|^2 & \cdots & \mathbf{x}_2 \cdot \mathbf{x}_p \\ \vdots & \vdots & & \vdots \\ \mathbf{x}_p \cdot \mathbf{x}_1 & \mathbf{x}_p \cdot \mathbf{x}_2 & \cdots & \|\mathbf{x}_p\|^2 \end{bmatrix} \qquad (13.31)
\end{aligned}$$

である．

ここで，各ベクトル $\mathbf{x}_k$ はそれぞれの変数を表すベクトルである．各ベクトルが各

要素と平均の差で構成されるベクトル（偏差ベクトル）であるとする．つまり，

$$\mathbf{x}_k = \begin{bmatrix} x_{1k} - \bar{x}_k \\ x_{2k} - \bar{x}_k \\ \vdots \\ x_{nk} - \bar{x}_k \end{bmatrix}. \tag{13.32}$$

このとき，行列の積 $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ を n で割った行列は，対角成分を分散，非対角成分を共分散とする行列になる．つまり，

$$\frac{1}{n}\mathbf{X}^{\mathrm{T}}\mathbf{X} = \begin{bmatrix} s_1^2 & s_{12} & \cdots & s_{1p} \\ s_{21} & s_2^2 & \cdots & s_{2p} \\ \vdots & \vdots & & \vdots \\ s_{p1} & s_{p2} & \cdots & s_p^2 \end{bmatrix}. \tag{13.33}$$

この行列を**分散共分散行列** (variance-covariance matrix) という．非対角要素の ij 成分と ji 成分は同じ共分散 $s_{ij} = s_{ji}$ であり等しい．このように対角項を対称に等しい正方行列を対称行列という．

分散共分散行列は，多変量解析において基本的な情報となる，各変数の変動の情報とそれぞれの変数間の共変動の情報がまとめられており，大変重要なものである．

さらに，各ベクトル $\mathbf{x}_k$ の要素は自らのノルムで基準化されているとする．

$$\mathbf{x}_k = \frac{1}{\|\mathbf{x}_k\|}\begin{bmatrix} x_{1k} - \bar{x}_k \\ x_{2k} - \bar{x}_k \\ \vdots \\ x_{nk} - \bar{x}_k \end{bmatrix} = \begin{bmatrix} \frac{x_{1k} - \bar{x}_k}{\|\mathbf{x}_k\|} \\ \frac{x_{2k} - \bar{x}_k}{\|\mathbf{x}_k\|} \\ \vdots \\ \frac{x_{nk} - \bar{x}_k}{\|\mathbf{x}_k\|} \end{bmatrix} \tag{13.34}$$

このとき，行列の積 $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ は，対角成分を 1，非対角成分を相関係数とする**相関行列** (correlation matrix) になる．つまり，

$$\mathbf{X}^{\mathrm{T}}\mathbf{X} = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & & \vdots \\ r_{p1} & r_{p2} & \cdots & 1 \end{bmatrix}. \tag{13.35}$$

付録 A

A.1　逆関数の存在

定理 A.1. 関数 $f : X \to Y$ について以下のことが成立する．

(1)　f が全単射 $\Longleftrightarrow (g \circ f = I_X) \wedge (f \circ g = I_Y)$ を満たす関数 $g : Y \to X$ が存在する．

(2)　そのような関数 g はただ 1 つ存在する．

証明．(1) 同値 $p \Longleftrightarrow q$ の証明なので，$p \Longrightarrow q$ の証明と $p \Longleftarrow q$ の証明の 2 つのパートに分ける．

($\Longrightarrow$)「f が全単射である」という命題から出発し，「$(g \circ f = I_X) \wedge (f \circ g = I_Y)$ を満たす関数 $g : Y \to X$ が存在する」という命題が論理的に導かれることを証明する．

f が全単射であることから，「$\forall y \in Y$ について，$y = f(x)$ となるただ 1 つの $x \in X$ が存在する」ことが導ける．このとき，関数 $g : Y \to X$ を $g(y) = x$ によって定義する．そうすると，

$$\forall y \in Y,\quad (f \circ g)(y) = f(g(y)) = f(x) = y \tag{A.1}$$

となる．これは関数 $f \circ g$ が Y 上の恒等関数であるということを示している．つまり，$f \circ g = I_Y$．一方，f が全単射であることから，「$\forall x \in X$ について，$y = f(x)$ となるただ 1 つの $y \in Y$ が存在する」ことが導ける．これより，

$$\forall x \in X,\quad (g \circ f)(x) = g(f(x)) = g(y) = x \tag{A.2}$$

となる．これで十分条件が証明された．

($\Longleftarrow$) $f \circ g = I_Y$ とする．このとき，

$$\forall y \in Y,\quad (f \circ g)(y) = f(g(y)) = I_Y(y) = y \tag{A.3}$$

である．ここで，関数 $g : Y \to X$ を $g(y) = x$ によって定義すると

$$\forall y \in Y, \quad \exists x \in X, \quad f(x) = y \tag{A.4}$$

が成り立つ．これは全射の定義に他ならない．一方，$g \circ f = I_X$ とする．$\forall x, x' \in X$, $f(x) = f(x')$ を仮定すると，

$$x = (g \circ f)(x) = g(f(x)) = g(f(x')) = (g \circ f)(x') = x' \tag{A.5}$$

となる．ゆえに，

$$\forall x, x' \in X, \quad f(x) = f(x') \Longrightarrow x = x' \tag{A.6}$$

が成立する．これは単射の定義そのものである．ゆえに f は全単射である．これで必要条件が証明された．
(2) 唯一性を証明する．2 つの関数 g_1, g_2 が (1) の性質を満たすと仮定する．このとき，$f \circ g_i = I_Y$ なので，

$$\forall y \in Y, \quad f(g_1(y)) = y = f(g_2(y)) \tag{A.7}$$

である．ところで，f は単射であるので，

$$\forall y \in Y, \quad g_1(y) = g_2(y) \tag{A.8}$$

である．ゆえに $g_1 = g_2$ が成り立つ． □

A.2　微分法を導入するための基礎

A.2.1　区間

実数 $a, b \in \mathbb{R}$ に対して，$a \leq x \leq b$ となるような x の集合を**閉区間** (closed interval) といい，

$$[a, b] \tag{A.9}$$

と表す．集合として定義すれば，

$$[a, b] = \{x \mid a \leq x \leq b\} \tag{A.10}$$

である．また，$a < x < b$ であるような x の集合を**開区間** (open interval) といい，

$$(a, b) \tag{A.11}$$

と表す．5 章で導入した順序対と同じ書き方をすることに注意せよ．普通は文脈から順序対なのか開区間なのかは明らかになるように書かれているので問題ない．先と同様に集合として定義すれば，

$$(a, b) = \{x \mid a < x < b\} \tag{A.12}$$

である．

実数を 1 つの直線で表すとき，これを数直線とよぶが，閉区間は数直線上で端点を含む線分であり，開区間は端点 a, b を含まない線分である．

また，片側の端点を含みもう片側の端点を含まないような区間を**半閉区間** (half-closed interval) または**半開区間** (half-open interval) と呼んで

$$[a, b) = \{x \mid a \leq x < b\} \tag{A.13}$$

$$(a, b] = \{x \mid a < x \leq b\} \tag{A.14}$$

と表す．また無限大 ∞ と無限小 $-\infty$ を用いて，$a \leq x$ となる x の集合や，$x < b$ となる x の集合もまた区間として表される．すなわち，

$$[a, \infty) = \{x \mid a \leq x\} \tag{A.15}$$

$$(-\infty, b] = \{x \mid x < b\} \tag{A.16}$$

である．また実数全体の集合 $\mathbb{R}$ も 1 つの区間と見なして

$$(-\infty, \infty) \tag{A.17}$$

で表す[*1]．

図 A.1 は，閉区間，開区間，半開区間の模式図である．

関数の独立変数 x が動く範囲をその関数の**定義域** (domain)D という．定義域は 1 つの区間である場合も，複数の区間の和集合である場合もある．普通，関数 $f(x)$ の

[*1] 数直線の両端は無限に大きく（小さく）なるので，端点は含まれない（端点を 1 つに決められない）と考える．

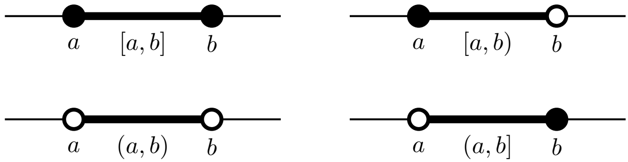

図 A.1　閉区間，開区間，半開区間

定義域としては，$f(x)$ が意味を持つもっとも広い実数の範囲を考える．例えば，一次関数 $f(x) = 2x + 3$ や二次関数 $f(x) = 2x^2 + 3x + 1$ の定義域は，実数全体の区間 $(-\infty, \infty)$ である．一方，分数関数

$$f(x) = \frac{1}{x-1} \tag{A.18}$$

は点 $x = 1$ において不定形となるので，定義域は $(-\infty, 1) \cup (1, \infty)$ である．

A.2.2　関数の極限

関数 $f(x)$ について，a を 1 つの定数としたとき，$x = a$ の場合を除いて a の十分近くの x において f が定義されており，$x = a$ のとき f は定義されていても定義されていなくてもよいとする．つまり，適当な正の実数 r をとったときの区間 $(a - r, a) \cup (a, a + r)$ が f の定義域に必ず含まれていると仮定する．このとき，x が a と異なる値をとりながら限りなく a に近づくとき，$f(x)$ が一定の値 α に近づくならば，「x が a に近づくとき $f(x)$ は α に**収束する** (converge)」といい，α を x が a に近づくときの関数 $f(x)$ の**極限値** (limit value) という．このことを

$$\lim_{x \to a} f(x) = \alpha \tag{A.19}$$

あるいは

$$f(x) \to \alpha \quad (x \to a) \tag{A.20}$$

と表す．lim は limit（極限）の意味である．x を a に近づけていくとき 2 つの方向からの接近が考えられる．1 つは数直線上の右から左方向への接近，つまり x が a よ

りも大きな値をとりながら a に近づく場合で，そのときの極限値を特に**右側極限値** (right-hand limit) といい，それが存在して α のとき

$$\lim_{x\to a+0} f(x) = \alpha \tag{A.21}$$

と表す．逆に左から右方向への接近，つまり x が a よりも小さな値をとりながら a に近づく場合，そのときの極限値を特に**左側極限値** (left-hand limit) といい，それが存在して β のとき

$$\lim_{x\to a-0} f(x) = \beta \tag{A.22}$$

と表す．$x \to a$ のとき，極限値が存在して α であるということは，

$$\lim_{x\to a-0} f(x) = \alpha = \lim_{x\to a+0} f(x) \tag{A.23}$$

であることに他ならない．反対に，

$$\lim_{x\to a-0} f(x) = \alpha \neq \beta = \lim_{x\to a+0} f(x) \tag{A.24}$$

であれば，$x \to a$ のときの極限値は存在しない．

極限値の計算として，自明なこととして，定数関数 $f(x) = c$（ただし c は定数）について，

$$\lim_{x\to a} c = c \tag{A.25}$$

が成り立ち，恒等関数 $f(x) = x$ について

$$\lim_{x\to a} x = a \tag{A.26}$$

が成り立つ．また，極限値と四則演算について次のような定理が成り立つ．これらの定理については厳密には証明を与える必要があるが，ここでは証明なしで真と認めることにする．

定理 A.2. $\lim_{x\to a} f(x) = \alpha, \lim_{x\to a} g(x) = \beta$ ならば，以下が成立する．

(1) $\lim_{x\to a} kf(x) = k\alpha$　　（ただし，k は定数）
(2) $\lim_{x\to a}\{f(x) + g(x)\} = \alpha + \beta$
(3) $\lim_{x\to a}\{f(x) - g(x)\} = \alpha - \beta$

(4) $\lim_{x \to a} f(x)g(x) = \alpha\beta$

(5) $\lim_{x \to a} \dfrac{f(x)}{g(x)} = \dfrac{\alpha}{\beta}$ （ただし，$\beta \neq 0$ のとき）

関数 $f(x)$ について，$x = a$ の場合を除いて a の十分近くの x において f が定義されており（$x = a$ のとき f は定義されていても定義されていなくてもよい），x が a に近づくとき，$f(x)$ が限りなく大きくなるならば，「$x \to a$ のとき $f(x)$ は**無限大に発散する** (diverge to positive infinity)」といい

$$\lim_{x \to a} f(x) = \infty \tag{A.27}$$

と表す[*2]．例えば，関数 $f(x) = 1/x^2$ について

$$\lim_{x \to 0} \frac{1}{x^2} = \infty \tag{A.28}$$

が成り立つ．

また，適当な実数 a をとれば区間 $[a, \infty)$ が定義域に含まれているような関数 $f(x)$ について，x が限りなく大きくなるにつれて $f(x)$ が限りなく α に近づくとき，そのことを

$$\lim_{x \to \infty} f(x) = \alpha \tag{A.29}$$

と表す[*3]．例えば，

$$\lim_{x \to \infty} x^2 = \infty \tag{A.30}$$

であるが，

$$\lim_{x \to \infty} \frac{1}{x^2} = 0 \tag{A.31}$$

である．

[*2] もし，$f(x)$ が限りなく小さくなるならば，「$x \to a$ のとき $f(x)$ は **負の無限大に発散する** (diverge to negative infinity)」といって $\lim_{x \to a} f(x) = -\infty$ と表す．

[*3] 同様にして，マイナス方向への無限大の時の極限も考えることができる．

A.2.3　関数の連続性

ところで，$x \to a$ のときの極限値が存在したとしても，$x = a$ としたときの関数の値 $f(a)$ と極限値が一致するとは限らない．以下，例を挙げよう．

(例 1) $f(x) = x^2 + 2x + 1$ とすると，

$$\lim_{x \to 1} f(x) = 4$$

であるので，$f(1)$ と一致する (図 A.2).

(例 2) 場合分け関数

$$f(x) = \begin{cases} x^2 + 2x + 1 & (x \neq 1) \\ 10 & (x = 1) \end{cases}$$

とする．この関数の $x \to 1$ の時の極限値は

$$\lim_{x \to 1} f(x) = 4$$

であるが，$f(1) = 10$ とは異なる (図 A.3).

(例 3) 分数関数

$$f(x) = \frac{x^2 - 1}{x - 1}$$

を考える．この関数は $x = 1$ のときには定義されない．一方，$x \neq 1$ のときには

$$f(x) = \frac{(x-1)(x+1)}{x-1} = x + 1$$

であるので，

$$\lim_{x \to 1} f(x) = 2$$

となる．

関数 $f(x)$ について，適当な正の実数 r をとったときの区間 $(a - r, a + r)$ が f の定義域に必ず含まれていると仮定する．このとき，もし

$$\lim_{x \to a} f(x) = f(a) \tag{A.32}$$

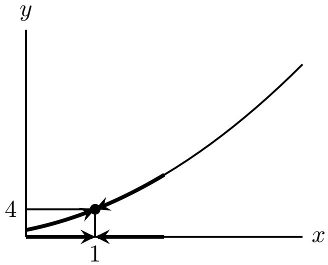

図 A.2　$x \to 1$ のときの $f(x)$(例 1)

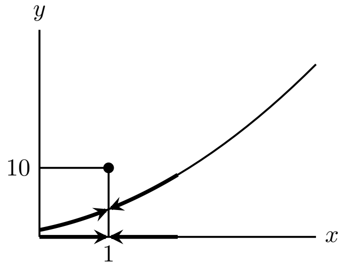

図 A.3　$x \to 1$ のときの場合分け関数 (例 2)

が成り立つならば，関数 $f(x)$ は **$\boldsymbol{x = a}$ において連続** (continuous at $x = a$) であるという．

例えば，関数 $f(x) = x^2$ は

$$\lim_{x \to 0} x^2 = 0 = f(0) \tag{A.33}$$

であるので，$x = 0$ において連続である（図 A.4）．また，x の**絶対値** (absolute value)$|x|$ を

$$|x| = \begin{cases} x & (x \geq 0) \\ -x & (x < 0) \end{cases} \tag{A.34}$$

と定義する．$f(x) = |x|$ とすると

$$\lim_{x \to 0} |x| = 0 = f(0) \tag{A.35}$$

が成立するので，この関数は $x = 0$ において連続である（図 A.5）．

$f(x)$ が区間 D の任意の点で連続であるとき，すなわち，

$$\forall a \in D, \lim_{x \to a} f(x) = f(a)$$

が成立するとき，$f(x)$ **は区間 $\boldsymbol{D}$ で連続** (continuous in interval D) であるという．特に閉区間 $[a, b]$ において f が連続であるという場合，端点 a においては**右側連続** (right continuous) であり，b においては**左側連続** (left continuous) である，つまり，

$$\lim_{x \to a+0} f(x) = f(a), \qquad \lim_{x \to b-0} f(x) = f(b) \tag{A.36}$$

と約束する．

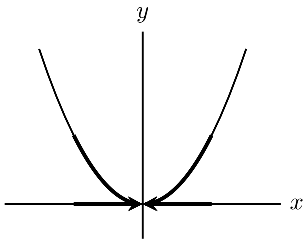

図 A.4　$\lim_{x\to 0} x^2$

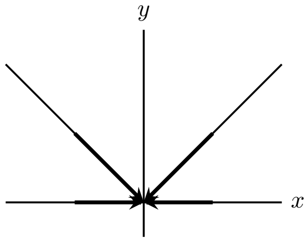

図 A.5　$\lim_{x\to 0} |x|$

A.3　微分可能と連続の関係

定理 A.3. 関数 $f(x)$ は $x=a$ において微分可能ならば，$f(x)$ は $x=a$ において連続である．したがって，$f(x)$ が区間 D で微分可能ならば，$f(x)$ は D で連続である．

証明. 関数 $f(x)$ が $x=a$ において微分可能であるとする．$x \neq a$ のとき $f(x)$ を変形すると，

$$f(x) = f(a) + f(x) - f(a) \tag{A.37}$$

$$= f(a) + \frac{f(x)-f(a)}{x-a}(x-a) \tag{A.38}$$

である．ところで，

$$\lim_{x\to a}\left\{\frac{f(x)-f(a)}{x-a}(x-a)\right\} = f'(a)\cdot 0 = 0 \tag{A.39}$$

であるから，恒等式

$$f(x) = f(a) + \frac{f(x)-f(a)}{x-a}(x-a) \tag{A.40}$$

の両辺について，$x \to a$ のときの極限をとると

$$\lim_{x\to a} f(x) = f(a) \tag{A.41}$$

となる．これは $f(x)$ が $x=a$ において連続であるという定義に他ならない．同様の推論を $\forall a \in D$ について行うことで，定理の後段も導かれる．　□

この定理を単純化して書くと,「微分可能 $\Longrightarrow$ 連続」であるので, 逆命題である「連続 $\Longrightarrow$ 微分可能」は常に成立するとは限らない. このことを示すには 1 つの反例を挙げれば十分である. 先に関数の連続性のところで検討した関数 $f(x) = |x|$ は, $x = 0$ において連続であった (図 A.5). しかし, 平均変化率について $a = 0$ のとき

$$\frac{f(b) - f(0)}{b - 0} = \frac{|b|}{b} \tag{A.42}$$

であるので, $b > 0$ のときは 1, $b < 0$ のときは -1 となる. ゆえに,

$$\lim_{b \to 0+0} \frac{f(b) - f(0)}{b - 0} = 1 \neq -1 = \lim_{b \to 0-0} \frac{f(b) - f(0)}{b - 0} \tag{A.43}$$

となり, 右側極限値と左側極限値は一致しないので, 極限値

$$\lim_{b \to 0} \frac{f(b) - f(0)}{b - 0} \tag{A.44}$$

は存在しない. ゆえに, 関数 $f(x) = |x|$ は $x = 0$ において微分可能ではない.

A.4 合成関数の微分

定理 A.4 (**合成関数の微分**). 関数 $y = f(x)$ が区間 I において微分可能で, その値域が区間 J に含まれているとする. このとき, $z = g(y)$ が区間 J において微分可能であれば, 合成関数

$$z = (g \circ f)(x) = g(f(x)) \tag{A.45}$$

は区間 I において微分可能であり,

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} = g'(f(x))f'(x) \tag{A.46}$$

である.

証明. まず, 関数 f について平均変化率を考えて,

$$\frac{f(x+h) - f(x)}{h} = f'(x) + \varepsilon_1(h) \tag{A.47}$$

とおく. ここで, $\varepsilon_1(h)$ は h を独立変数とする何らかの関数である.

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = f'(x) \tag{A.48}$$

であるので，$\lim_{h\to 0}\varepsilon_1(h)=0$ でなければならない．式 (A.47) の両辺に h をかけて，

$$f(x+h)-f(x)=\{f'(x)+\varepsilon_1(h)\}h \tag{A.49}$$

となる．次に，g の平均変化率を考える．先と同様に，

$$\frac{g(y+k)-g(y)}{k}=g'(y)+\varepsilon_2(k) \tag{A.50}$$

とおく．やはりここでも $\lim_{k\to 0}\varepsilon_2(k)=0$ が成り立つ．ただし，$y=f(x), y+k=f(x+h)$ となるように k を設定する．すると, $k=f(x+h)-f(x)$ であり, f の連続性（これは微分可能性から導かれる）により，$h\to 0$ のとき，$\lim_{x+h\to x}f(x+h)=f(x)$ なので，$k\to 0$ となる．式 (A.50) を変形すると,

$$\begin{aligned} g(f(x+h))-g(f(x)) &= \{g'(f(x))+\varepsilon_2(k)\}k & \text{(A.51)}\\ &= \{g'(f(x))+\varepsilon_2(k)\}\{f'(x)+\varepsilon_1(h)\}h & \text{(A.52)}\end{aligned}$$

結局,

$$\frac{g(f(x+h))-g(f(x))}{h}=\{g'(f(x))+\varepsilon_2(k)\}\{f'(x)+\varepsilon_1(h)\} \tag{A.53}$$

であるので，$h\to 0$ の極限値は

$$\lim_{h\to 0}\frac{g(f(x+h))-g(f(x))}{h}=g'(f(x))f'(x) \tag{A.54}$$

となる．　□

A.5　関数の挙動

A.5.1　接線

関数 $y=f(x)$ のグラフ上に点 $P(a,f(a))$ と $b\neq a$ となる点 $Q(b,f(b))$ をとると，平均変化率

$$\frac{\Delta y}{\Delta x}=\frac{f(b)-f(a)}{b-a} \tag{A.55}$$

は直線 PQ の傾きを示している．そこで，$b \to a$ とすれば，点 Q は $f(x)$ 上を動いて点 P に限りなく近づく．このとき，微分係数

$$f'(a) = \lim_{b \to a} \frac{f(b) - f(a)}{b - a} \tag{A.56}$$

が存在するならば，直線 PQ は $b \to a$ のとき，点 P を通り傾き $f'(a)$ をもつ直線に限りなく近づく．この直線を $y = f(x)$ のグラフ上の点 P におけるグラフの**接線** (tangent) といい，点 P をグラフと接線の**接点** (contact) という（図 A.6）．

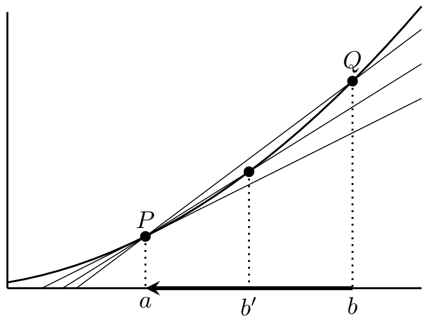

図 A.6　$f(x)$ の接線

具体的には，関数 $y = f(x)$ が $x = a$ において微分可能であれば，この関数のグラフ上の点 $P(a, f(a))$ における**接線の方程式** (tangent equation) は

$$y - f(a) = f'(a)(x - a) \tag{A.57}$$

で与えられる．

A.5.2　関数の増減

ここで，関数の増減についての定義を導入する．定義域に区間 D を含む $y = f(x)$ について，

$$\forall x_1, x_2 \in D, \qquad x_1 < x_2 \Longrightarrow f(x_1) < f(x_2)$$

が成り立つならば，$f(x)$ は D において**単調に増加する**，あるいは D における**単調増加関数**という．一方，

$$\forall x_1, x_2 \in D, \qquad x_1 < x_2 \Longrightarrow f(x_1) > f(x_2)$$

が成り立つならば，$f(x)$ は D において**単調に減少する**，あるいは D における**単調減少関数**という[*4].

A.5.3　極値と最大・最小値

関数 $f(x)$ の定義域内の 1 点 c とその十分近くの任意の x について，$f(c) \geq f(x)$ が成立するとき，もう少し厳密に言うと，$a < c < b$ を満たす a と b を c の十分近くにとるとき

$$\forall x \in (a, b), \qquad f(c) \geq f(x) \tag{A.58}$$

が成立するならば，c は f の**極大点** (local maximum point)，$f(c)$ は f の**極大値** (local maximum value) という．一方，

$$\forall x \in (a, b), \qquad f(c) \leq f(x) \tag{A.59}$$

が成立するならば，c は f の**極小点** (local minimum point)，$f(c)$ は f の**極小値** (local minimum value) という．極大点と極小点を合わせて**極値点**，極大値と極小値を合わせて**極値**という．

さらに，

$$\forall x \in (a, c) \cup (c, b), \qquad f(c) > f(x) \tag{A.60}$$

が成立するならば，c を**狭義の極大点**，$f(c)$ を**狭義の極大値**ともいう．また，

$$\forall x \in (a, c) \cup (c, b), \qquad f(c) < f(x) \tag{A.61}$$

[*4] これらの単調関数を狭義単調増加・減少関数と呼んで，

$$\forall x_1, x_2 \in D, \qquad x_1 < x_2 \Longrightarrow f(x_1) \leq f(x_2)$$

のときの**広義単調増加関数（単調非減少関数）**，そして，

$$\forall x_1, x_2 \in D, \qquad x_1 < x_2 \Longrightarrow f(x_1) \geq f(x_2)$$

のときの**広義単調減少関数（単調非増加関数）**と区別することがある．

が成立するならば，c を**狭義の極小点**，$f(c)$ を**狭義の極小値**ともいう．

関数 $f(x)$ の定義域 D 内の 1 点 c について，

$$\forall x \in D, \qquad f(c) \geq f(x) \tag{A.62}$$

が成立するとき，c を D における f の**最大点** (global maximum point)，$f(c)$ を D における**最大値** (global maximum value) という．同様に，

$$\forall x \in D, \qquad f(c) \leq f(x) \tag{A.63}$$

が成立するとき，c を D における f の**最小点** (global minimum point)，$f(c)$ を D における**最小値** (global minimum value) という．

最大・最小値については次の定理が大変重要である．

定理 A.5 (最大・最小値の定理). 関数 $f(x)$ が閉区間で連続ならば，この区間において $f(x)$ の最大値および最小値が存在する．

この定理は直感的にはほぼ明らかであるが，きちんと証明するためには，実数の連続性についての厳密な議論が必要となるため，ここでは証明なしでこの定理を認めることにする．また，定理中の閉区間という限定が重要であり，開区間であれば最大値もしくは最小値が存在しない場合がある．

一般的に，極値点のうちの 1 つが必ず最大あるいは最小値であるとは限らない．閉区間である定義域の端点となることもある．

A.5.4　平均値の定理

次に，微分法を用いて関数の増減や極値を調べるために，いくつかの基本的な定理を導入する．

定理 A.6. $f(x)$ が区間 D で定義された関数，c を D の *1* つの端点ではない点とする．このとき，f は c において微分可能であり，かつ c が f の極値点であれば，

$$f'(c) = 0 \tag{A.64}$$

である．

証明. c が極大点であるとして証明する．c は端点ではなく微分可能であるので，左側極限値と右側極限値の両方が存在して，それらは微分係数 $f'(c)$ に等しい．

一方，c は f の極大点なので，h を絶対値が十分に小さい正または負の数とすれば，$c+h \in D$ かつ

$$f(c) \geq f(c+h) \tag{A.65}$$

である．したがって，$h > 0$ ならば

$$\frac{f(c+h)-f(c)}{h} \leq 0 \tag{A.66}$$

であり，$h < 0$ ならば

$$\frac{f(c+h)-f(c)}{h} \geq 0 \tag{A.67}$$

である．ゆえに，$h \to 0$ の右側極限値は

$$\lim_{h \to 0+} \frac{f(c+h)-f(c)}{h} \leq 0 \tag{A.68}$$

となり，左側極限値は

$$\lim_{h \to 0-} \frac{f(c+h)-f(c)}{h} \geq 0 \tag{A.69}$$

となる．これらは微分係数 $f'(c)$ に等しいのであるから，$f'(c) \leq 0$ かつ $f'(c) \geq 0$ が同時に成立しなければならない．ゆえに，$f'(c) = 0$ である．極小点である場合も同様である． □

定理 A.7 (ロルの定理). 関数 $f(x)$ が閉区間 $[a,b]$ で連続であり，開区間 (a,b) で微分可能であり，$f(a) = f(b)$ であるとする．このとき

$$f'(c) = 0 \tag{A.70}$$

であるような $c \in (a,b)$ が少なくとも 1 つ存在する．

証明. $f(x)$ が $[a,b]$ で定数ならば，$\forall c \in (a,b)$，$f'(c) = 0$ であるので定理が成立する．f は $[a,b]$ で定数ではないとする．このとき，(a,b) に $f(s) \neq f(a)$ となる s が存在する．

$f(s) > f(a)$ の場合を考察する．定理 A.5 より，$[a, b]$ において f の最大点が存在する．これを c とすれば，最大値の定義より，

$$f(c) \geq f(s) > f(a) \tag{A.71}$$

であるので，c は端点 a, b と等しくない．つまり，$c \in (a, b)$ である．c は f の最大点であるので，同時に極大点でもある．また，f は c において微分可能であるので，定理**??**より，$f'(c) = 0$ である．

$f(s) < f(a)$ の場合は最小点を考えれば同様に証明できる．□

ロルの定理をさらに一般化したものが，平均値の定理である．

定理 A.8 (**平均値の定理**)**.** 関数 $f(x)$ が閉区間 $[a, b]$ で連続で，開区間 (a, b) で微分可能ならば,

$$\frac{f(b) - f(a)}{b - a} = f'(c) \tag{A.72}$$

であるような $c \in (a, b)$ が少なくとも 1 つ存在する．

証明. $f(x)$ のグラフ上の点 $P(a, f(a))$ と $Q(b, f(b))$ を結ぶ直線は，1 次関数

$$g(x) = \frac{f(b) - f(a)}{b - a}(x - a) + f(a) \tag{A.73}$$

で表すことができる．ここで，$\phi(x) = f(x) - g(x)$ とおく．これを展開すると,

$$\phi(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a) - f(a) \tag{A.74}$$

である．このとき，ϕ もまた閉区間 $[a, b]$ で連続で開区間 (a, b) で微分可能であり,かつ,

$$\phi(a) = \phi(b) = 0 \tag{A.75}$$

が成り立つ．ゆえに, 定理 A.7「ロルの定理」が適用できて, $\phi'(c) = 0$ となる $c \in (a, b)$ が存在する．ところで,

$$\phi'(x) = f'(x) - \frac{f(b) - f(a)}{b - a} \tag{A.76}$$

であるので，

$$\phi'(c) = 0 \Longleftrightarrow \frac{f(b) - f(a)}{b - a} = f'(c) \tag{A.77}$$

となる． □

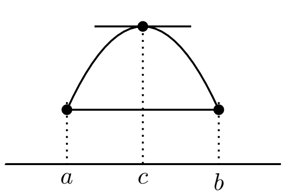

図 A.7　ロルの定理

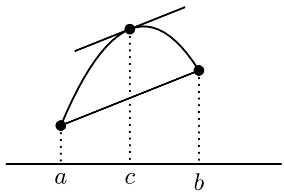

図 A.8　平均値の定理

A.5.5　関数の増減と極値

以下の定理は，関数の挙動を分析する上で必要になってくる．

定理 A.9. 関数 $f(x)$ が区間 D で連続で，かつ D の端点を取り除いた区間（これを D の内部という）で微分可能であるとする．そのとき，

(1) D の内部でつねに $f'(x) > 0$ ならば，f は D において単調に増加する．
(2) D の内部でつねに $f'(x) < 0$ ならば，f は D において単調に減少する．
(3) D の内部でつねに $f'(x) = 0$ ならば，f は D において定数である．

証明. x_1, x_2 を $x_1 < x_2$ を満たす D の任意の 2 数とする．そのとき，定理 A.8「平均値の定理」より，区間 (x_1, x_2) に

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1) \tag{A.78}$$

となる c が存在する．いま，$(x_2 - x_1) > 0$ なので，$f(x_2) - f(x_1)$ と $f'(c)$ の正負は一致する． □

ある区間で定義された関数 $y = f(x)$ の導関数 $y' = f'(x)$ が，その区間においてさらに微分可能であれば，その導関数を**二次導関数** (second derivative) といって，

$$f''(x), \quad y'', \quad \frac{d^2y}{dx^2}, \quad \frac{d^2}{dx^2}f(x)$$

などと表す．

定理 A.6 で $f'(a) = 0$ となるとき a で極値をとることが分かっているが，この二次導関数を用いて，さらにその点が極大点か極小点かを判定をすることができる．

定理 A.10. 関数 $f(x)$ が区間 D で第 *2* 次導関数 $f''(x)$ をもち，f'' が連続であるとする．このとき，D の内部の点である a について，

$$f'(a) = 0, f''(a) > 0 \text{ ならば, } a \text{ は } f \text{ の狭義の極小点であり,}$$
$$f'(a) = 0, f''(a) < 0 \text{ ならば, } a \text{ は } f \text{ の狭義の極大点である.}$$

証明. $f'(a) = 0, f''(a) > 0$ とする．f'' が連続なので，十分小さな $\epsilon > 0$ について a の近くの区間 $(a - \epsilon, a + \epsilon)$ で $f''(x) > 0$ である．したがって，$f'(x)$ は区間 $(a - \epsilon, a + \epsilon)$ において単調増加している．$f'(a) = 0$ なので，区間 $(a - \epsilon, a)$ で $f'(x) < 0$，区間 $(a, a + \epsilon)$ で $f'(x) > 0$ となる．ゆえに，点 a の左側で $f(x)$ は単調減少し，点 a の右側で単調増加する．よって，

$$\forall x \in (a - \epsilon, a) \cup (a, a + \epsilon), \quad f(a) < f(x) \tag{A.79}$$

が成立する．これは a が狭義極小点であることに他ならない．$f'(a) = 0, f''(a) < 0$ についても同様に証明できる． □

注意点として，$f'(a) = 0, f''(a) = 0$ となる場合は，a がつねに極値であるとは限らない．この場合，さらなる極値判定条件が必要となる．ちなみに，$f''(a) = 0$ となる点 a を f の**変曲点** (inflection point) という．

A.5.6 増減表

ここまでの定理を用いて，具体的に関数の増減ならびに極値の分析を行う．例として，

$$f(x) = x^3 - 3x + 1 \tag{A.80}$$

を検討する．

$$f'(x) = 3x^2 - 3 = 3(x+1)(x-1), \qquad f''(x) = 6x \tag{A.81}$$

である．この結果からまず，$f'(1) = f'(-1) = 0$ であること，さらに

$$f'(x) \begin{cases} > 0 & (x < -1) \\ < 0 & (-1 < x < 1) \\ > 0 & (1 < x) \end{cases} \tag{A.82}$$

である．また，二次導関数に関しては，

$$f''(x) \begin{cases} > 0 & (x < 0) \\ = 0 & (x = 0) \\ < 0 & (x > 0) \end{cases} \tag{A.83}$$

である．この結果から，-1 が極大点で $f(-1) = 3$，1 が極小点で $f(1) = -1$ であることが分かる．また，関数の増減を表の形でまとめると表 A.1 のようになる．このような表を**増減表**という．

表 A.1　$f(x) = x^3 - 3x + 1$ の増減表

x	$\cdots$	-1	$\cdots$	0	$\cdots$	1	$\cdots$
$f''(x)$	$-$	$-$	$-$	0	$+$	$+$	$+$
$f'(x)$	$+$	0	$-$	$-$	$-$	0	$+$
$f(x)$	$\nearrow$	3	$\searrow$	$\searrow$	$\searrow$	-1	$\nearrow$

　もう一つの例として，

$$f(x) = x^4 + 2x^3 \tag{A.84}$$

を分析する．

$$f'(x) = 2x^2(2x+3), \qquad f''(x) = 12x(x+1) \tag{A.85}$$

である．ここから増減表を作成すると表 A.2 のようになる．ここで，$-3/2$ は極小点であるが，0 は極値点ではない．

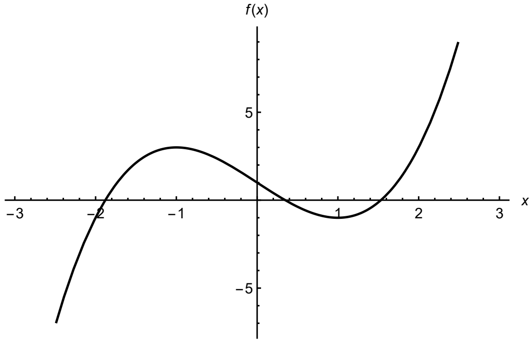

図 A.9　$f(x) = x^3 - 3x + 1$ のグラフ

表 A.2　$f(x) = x^4 + 2x^3$ の増減表

x	$\cdots$	$-\frac{3}{2}$	$\cdots$	-1	$\cdots$	0	$\cdots$
$f''(x)$	$+$	$+$	$+$	0	$-$	0	$+$
$f'(x)$	$-$	0	$+$	$+$	$+$	0	$+$
$f(x)$	$\searrow$	$-\frac{27}{16}$	$\nearrow$	$\nearrow$	$\nearrow$	0	$\nearrow$

ここでは，極値のみを検討しているが，ある区間内における最大・最小値を見つけるという問題の場合，区間内のすべての極値に加えて，端点，さらに場合によって微分不可能な点を加えて比較する必要がある．

A.6　定積分の基本的性質

定理 A.11. 区間 $[a, b]$ で連続な関数 f について

$$m(b-a) \leq \int_a^b f \leq M(b-a). \tag{A.86}$$

証明.

$$m(b-a) \leq L(P, f) \leq R(P, f) \leq U(P, f) \leq M(b-a) \tag{A.87}$$

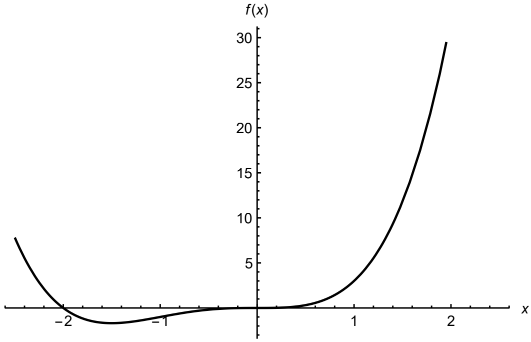

図 A.10　$f(x) = x^4 + 2x^3$ のグラフ

という関係性より明らかである．□

定理 A.12. 関数 f が積分可能である区間 D において，任意の $a, b, c \in D$ について，

$$\int_a^b f = \int_a^c f + \int_c^b f \tag{A.88}$$

が成り立つ．

証明．$a < c < b$ の場合，区間 $[a, b]$ の分割 $P = (x_0, x_1, \cdots, x_n)$ の点のなかの 1 つを c となるようにとる．そうすると，区間 $[a, c]$ についての分割 $P_1 = (x_0, \cdots, c)$ と区間 $[c, b]$ についての分割 $P_2 = (c, \cdots, x_n)$ のそれぞれが得られることになる．リーマン和の定義から

$$R(P, f) = R(P_1, f) + R(P_2, f) \tag{A.89}$$

である．$|P| = \max\{|P_1|, |P_2|\}$ なので，$|P| \to 0$ のとき，$|P_1| \to 0$ かつ $|P_2| \to 0$ である．ゆえに

$$\lim_{|P|\to 0} R(P, f) = \lim_{|P_1|\to 0} R(P_1, f) + \lim_{|P_2|\to 0} R(P_2, f) \tag{A.90}$$

である．定積分の定義より定理が成立する．その他の場合は，まずは上述のように定

積分の分解をした上で式を変形すれば定理の通りになる．例えば $a < b < c$ の場合，

$$\int_a^c f = \int_a^b f + \int_b^c f \tag{A.91}$$

であるので，式を変形すると

$$\int_a^b f = \int_a^c f - \int_b^c f = \int_a^c f + \int_c^b f \tag{A.92}$$

を得る． □

関数 f をある区間 D において連続な関数とする．関数 f は D において積分可能であるので，ある $a \in D$ を下端として，任意の $x \in D$ を上端とする定積分が存在する．これは変数 x の関数と見なすことができるので，この関数を $G(x)$ とおくと

$$G(x) = \int_a^x f(t)dt \tag{A.93}$$

と定義される．このとき，以下の定理が成立する．

定理 A.13. 上記の仮定の下で，関数 $G(x)$ は区間 D で微分可能で，

$$G'(x) = f(x) \tag{A.94}$$

が成立する．

証明. x は D の端点ではないとする[*5]．定理の証明のためには，$x \in D$ について，$x + h \in D$ となるような，絶対値の十分小さな h について，

$$\lim_{h \to 0} \frac{G(x+h) - G(x)}{h} = f(x) \tag{A.95}$$

が成立することを示せばよい．ここでは $h > 0$ のケースを示す．関数 G の定義より，

$$G(x+h) - G(x) = \int_a^{x+h} f - \int_a^x f \tag{A.96}$$

である．ここで，定理 A.12 より

$$\int_a^{x+h} = \int_a^x f + \int_x^{x+h} f \tag{A.97}$$

[*5] 端点を考慮する場合は，片側連続や片側微分可能の議論を導入する必要がある．

なので，

$$G(x+h) - G(x) = \int_x^{x+h} f \tag{A.98}$$

である．区間 $[x, x+h]$ における f の最大点を c，最小点を d とする．このとき，

$$\forall t \in [x, x+h], \quad f(d) \leq f(t) \leq f(c) \tag{A.99}$$

が成立する．定理 A.11 を用いると

$$f(d)h \leq \int_x^{x+h} f \leq f(c)h \tag{A.100}$$

となる．$h > 0$ と仮定しているので，各辺を h で割ると

$$f(d) \leq \frac{G(x+h) - G(x)}{h} \leq f(c) \tag{A.101}$$

が成り立つ．ところで，$h \to 0$ のとき区間 $[x, x+h]$ は徐々に小さくなっていく．よって区間内の点 c, d も x に近づく．ゆえに，各辺について $h \to 0$ の時の極限をとると

$$\lim_{d \to x} f(d) \leq \lim_{h \to 0} \frac{G(x+h) - G(x)}{h} \leq \lim_{c \to x} f(c) \tag{A.102}$$

となる[*6]．関数 f の連続性より，

$$\lim_{d \to x} f(d) = \lim_{c \to x} f(c) = f(x) \tag{A.103}$$

なので，

$$\lim_{h \to 0} \frac{G(x+h) - G(x)}{h} = f(x) \tag{A.104}$$

でなければならない[*7]．$h < 0$ のときは，符号の向きと不等式の扱いに気をつけながら，同様にすれば証明できる． □

[*6] 関数の極限の特性，$f(x) \leq g(x) \Longrightarrow \lim_{x \to a} f(x) \leq \lim_{x \to a} g(x)$ を用いた．

[*7] このように，不等式 $f(x) \leq g(x) \leq h(x)$ の各辺について，$x \to a$ のときの極限をとって，$\lim_{x \to a} f(x) = \lim_{x \to a} h(x) = \alpha$ から，$\lim_{x \to a} g(x) = \alpha$ を導き出すやり方は「はさみうち法」などと呼ばれる．

A.7　置換積分とその他の積分テクニック

A.7.1　不定積分の計算

以下に述べるのは，合成関数の微分法から導出される積分法で，置換積分法と呼ばれるものである．

定理 A.14 (置換積分法). x を t を変数とする微分可能な関数 $\varphi(t)$ を用いて，$x = \varphi(t)$ とおいたとき

$$\int f(x)dx = \int f(\varphi(t))\varphi'(t)dt. \tag{A.105}$$

証明. $F(x)$ を $f(x)$ の 1 つの不定積分であるとする．すなわち，

$$F(x) = \int f(x)dx \tag{A.106}$$

とする，このとき，$x = \varphi(t)$ とおいて，$F(x) = F(\varphi(t))$ を t で微分すると，合成関数の微分法により，

$$\frac{d}{dt}F(\varphi(t)) = \frac{d}{dx}F(x)\frac{dx}{dt} = f(x)\varphi'(t) = f(\varphi(t))\varphi'(t) \tag{A.107}$$

である．左辺と右辺を t について積分すると，定理の式を得る □

置換積分法を用いて，x についての積分を $x = \varphi(t)$ とおいて，t についての積分に帰着させることで，積分を解くのが簡単になることがある．例えば，不定積分

$$\int (2x-3)^5 dx \tag{A.108}$$

を求める．ここで $t = 2x - 3$ とおくと，

$$x = \varphi(t) = \frac{t+3}{2}, \qquad \varphi'(t) = \frac{1}{2} \tag{A.109}$$

である．置換積分法の公式に，それぞれを当てはめると

$$\int (2x-3)^5 dx = \int t^5 \cdot \frac{1}{2}dt = \frac{1}{12}t^6 + C \tag{A.110}$$

となる．この結果を x についてまとめると，

$$\int (2x-3)^5 dx = \frac{1}{12}(2x-3)^6 + C. \tag{A.111}$$

そのほか，複雑な関数について積分を解くときには，いくつかのテクニックを知っておく必要がある．ここでは，さらに部分積分法と未定係数法に簡単に触れるだけにとどめる．

定理 A.15 (部分積分法)**.**

$$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx \tag{A.112}$$

証明. 定理**??**(4) より，

$$\{f(x)g(x)\}' = f'(x)g(x) + f(x)g'(x) \tag{A.113}$$

であり，両辺を x について積分すると，

$$f(x)g(x) = \int f'(x)g(x)dx + \int f(x)g'(x)dx \tag{A.114}$$

となる．これをまとめると公式を得る． □

部分積分法によって

$$\int \log x dx \tag{A.115}$$

を求める．$\log x = (\log x) \cdot 1$ と考え，$f(x) = \log x$，$g'(x) = 1$ とおく．このとき，

$$f'(x) = \frac{1}{x}, \qquad g(x) = x \tag{A.116}$$

であるので，公式に当てはめると，

$$\int \log x dx = x \log x - \int \frac{1}{x} \cdot x dx = x \log x - x + C \tag{A.117}$$

である．

最後に，分数関数を**未定係数法** (method of undetermined coefficients) によって解く方法を例示しておく．

$$\int \frac{x-7}{(x+3)(x-2)} dx \tag{A.118}$$

を解く．このとき A, B を定数として，

$$\frac{x-7}{(x+3)(x-2)} = \frac{A}{x+3} + \frac{B}{x-2} \tag{A.119}$$

とおく，この A, B が求まれば，簡単な分数関数に分解することができる．両辺に $(x+3)(x-2)$ をかけると，

$$x - 7 = A(x-2) + B(x+3) = (A+B)x - (2A - 3B) \tag{A.120}$$

となる．これは x についての恒等式であるので，連立方程式

$$1 = A + B, \qquad 7 = 2A - 3B \tag{A.121}$$

が成立しなければならない．これを A, B について解けば，$A = 2, B = -1$ である．したがって，

$$\int \frac{x-7}{(x+3)(x-2)}dx = \int \frac{2}{x+3}dx - \int \frac{dx}{x-2} \tag{A.122}$$

$$= 2\log|x+3| - \log|x-2| + C \tag{A.123}$$

$$= \log\frac{(x+3)^2}{|x-2|} + C \tag{A.124}$$

A.7.2 定積分の計算

積分可能な関数 $f(x)$ について，a から b の範囲で定積分を計算するためには，前節の知識を用いて関数 $f(x)$ の不定積分 $F(x)$ を計算し，微分積分学の基本定理より，

$$\int_a^b f(x)dx = [F(x)]_a^b = F(b) - F(a) \tag{A.125}$$

を計算すればよい．ここでは，定積分における置換積分法と部分積分法について確認しておく．

定理 A.16 (**定積分の置換積分法**)**.** 関数 $f(x)$ が連続であり，$x = \varphi(t)$ が連続な導関数 $\varphi'(t)$ をもち，t の値 α, β に対して

$$\varphi(\alpha) = a, \qquad \varphi(\beta) = b \tag{A.126}$$

ならば,

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)dt. \tag{A.127}$$

証明. 合成関数の微分

$$\frac{d}{dt}F(\varphi(t)) = \frac{d}{dx}F(x)\frac{dx}{dt} = f(x)\varphi'(t) = f(\varphi(t))\varphi'(t) \tag{A.128}$$

より,

$$F(\varphi(t)) = \int f(\varphi(t))\varphi'(t)dt. \tag{A.129}$$

よって,

$$\int_\alpha^\beta f(\varphi(t))\varphi'(t)dt = [F(\varphi(t))]_\alpha^\beta = F(\varphi(\beta)) - F(\varphi(\alpha)) \tag{A.130}$$

$$= F(b) - F(a) = [F(x)]_a^b = \int_a^b f(x)dx \tag{A.131}$$

□

例えば,

$$\int_1^2 (2x-1)^3 dx \tag{A.132}$$

を求める．$2x-1=t$ であり，$x=\varphi(t)=(t+1)/2$，また $\varphi'(t)=1/2$ である．さらに，$x=1$ のとき $t=1$，$x=2$ のとき $t=3$ なので，公式より，

$$\int_1^3 t^3 \cdot \frac{1}{2}dt = \frac{1}{2}\left[\frac{t^4}{4}\right]_1^3 = 10. \tag{A.133}$$

定積分の部分積分法は，部分積分法の自然の拡張として得られる．

定理 A.17 (定積分の部分積分法)**.**

$$\int_a^b f(x)g'(x)dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x)dx \tag{A.134}$$

A.8　相関係数の特性の証明

定理 A.18.

$$-1 \leq r_{xy} \leq 1$$

証明. 常に非負となる式 $(1/n)\sum_{i=1}^{n}(z_i \pm w_i)^2$ を展開する.

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}(z_i \pm w_i)^2 &= \frac{1}{n}\sum_{i=1}^{n}(z_i^2 \pm 2z_i w_i + w_i^2) \\
&= \frac{1}{n}\sum_{i=1}^{n} z_i^2 \pm \frac{2}{n}\sum_{i=1}^{n} z_i w_i + \frac{1}{n}\sum_{i=1}^{n} w_i^2 \\
&= 1 \pm 2r_{xy} + 1 = 2(1 \pm r_{xy}) \geq 0
\end{aligned}$$

$\pm r_{xy} \geq -1$ より，$-1 \leq r_{xy} \leq 1$ である. □

定理 A.19.

$$\forall i,\ y_i = ax_i + b \Longleftrightarrow |r_{xy}| = 1$$

証明. $(\Longrightarrow)$ $\forall i,\ y_i = ax_i + b$ とすると，$s_y = |a|s_x$．また

$$\begin{aligned}
s_{xy} &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \\
&= \frac{a}{n}\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x}) = as_x^2
\end{aligned}$$

である．ゆえに,

$$r_{xy} = \frac{as_x^2}{s_x|a|s_x} = \frac{a}{|a|}.$$

つまり，$a > 0$ のとき相関係数は 1 で，$a < 0$ のとき -1 となる.

$(\Longleftarrow)$ $r_{xy} = \pm 1$ とする．このとき,

$$s_{xy} = \pm s_x s_y = \pm \frac{s_y}{s_x} s_x^2$$

が成り立つ．$a = \pm s_y/s_x$ とすると，上式より

$$\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = a\sum_{i=1}^{n}(x_i - \bar{x})(x_i - \bar{x})$$

が成り立つ．ゆえに,

$$\forall i, (y_i - \bar{y}) = a(x_i - \bar{x})$$

でなければならない．$b = \bar{y} - a\bar{x}$ とすると $\forall i,\ y_i = ax_i + b$ を得る． $\square$

付録 B

B.1　練習問題の解答

練習問題 2.3

p と q の真理値の 4 つのパターン TT,TF,FT,FF に対応する合成命題の真理値のみを記す．p のみの合成命題の場合は 2 つのパターン T,F に対応する真理値.

(1)　FTTT　(2)　FFFT　(3)　FFFT

(4)　TT　(5)　FF　(6)　FFFF

(7)　FFFT　(8)　FTTF

練習問題 2.4

(1)　TTTT　(2)　TTTT　(3)　TT

(4)　TTTT　(5)　TFTT　(6)　TTTT

練習問題 2.5

省略.

練習問題 3.1

省略.

練習問題 3.2

推論の妥当性

(1) 妥当 (2) 妥当 (3) 妥当でない

間接証明

(1) p が偽，すなわち $\neg p$ が真であるとする．ここから，$\neg p \wedge \neg q \Longleftrightarrow \neg(p \vee q)$ であるが，前提の $p \vee q$ と矛盾するため，p は真でなければならない.
(2) $\neg p$ は偽，すなわち p が真であるとする．すると，$p \wedge (p \leftrightarrow q) \Longrightarrow q$ であり，さらに $q \wedge (q \rightarrow \neg r) \Longrightarrow \neg r$ となる．しかしながら，$\neg r$ は前提である r が真ということと矛盾する．よって，$\neg p$ は真である．

練習問題 4.1

$$\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\}$$

部分集合関係は，すべての要素について自分自身との間で成り立つ（たとえば，$\{2,3\} \subset \{2,3\}$）．それ以外では，$\emptyset \subset \{1\}, \emptyset \subset \{2\}, \emptyset \subset \{3\}, \emptyset \subset \{1,2\}, \emptyset \subset \{1,3\}, \emptyset \subset \{2,3\}, \emptyset \subset \{1,2,3\}, \{1\} \subset \{1,2\}, \{1\} \subset \{1,3\}, \{1\} \subset \{1,2,3\}, \{2\} \subset \{1,2\}, \{2\} \subset \{2,3\}, \{2\} \subset \{1,2,3\}, \{3\} \subset \{1,3\}, \{3\} \subset \{2,3\}, \{3\} \subset \{1,2,3\}, \{1,2\} \subset \{1,2,3\}, \{2,3\} \subset \{1,2,3\}, \{1,3\} \subset \{1,2,3\}$.

練習問題 4.2

省略.

練習問題 5.1.1

省略.

練習問題 5.1.2

省略.

練習問題 6.3

(1)　1　　(2)　0　　(3)　4/6　　(4)　5/6

練習問題 6.4

(1)　1/3　　(2)　1/6　　(3)　1/216　　(4)　1/72

練習問題 7.5

省略.

練習問題 8.2

次の式を展開せよ.

(1)　30

(2)　$\displaystyle\sum_{i=1}^{n} a_i - \sum_{i=1}^{n} b_i$

(3)　$\displaystyle 3\sum_{i=1}^{n} a_i + 2b$

(4)　$\displaystyle\sum_{i=1}^{n} a_i^2 + 2\sum_{i=1}^{n} a_i b_i + \sum_{i=1}^{n} b_i^2$

練習問題 8.5

省略.

練習問題 9.1

(1)　1　　(2)　$2a$　　(3)　$3a^2$　　(4)　0　　(5)　$-\dfrac{1}{a^2}$

練習問題 9.3

(1) $2x-4$ (2) $3x^2+3$ (3) $6x+13$

(4) $-\dfrac{1}{x^2}+2x$ (5) $4x^3+6x^2+2x$ (6) $\dfrac{17}{(4x+1)^2}$

(7) $-\dfrac{2}{x^3}$ (8) $-\dfrac{3}{x^4}$ (9) $2x-\dfrac{2}{x^3}$

練習問題 9.4

(1) $6x^2(x^3+2)$ (2) $4(x^2+1)(3x-1)^3(6x^2-x+3)$

(3) $\dfrac{-2}{(x+3)^3}$

練習問題 10.5

積分定数は省略する．

(1) x^3-2x^2-2x (2) $x^2+\dfrac{1}{4}x^4$

(3) $\dfrac{2}{3}x^3-\dfrac{5}{2}x^2+3x$ (4) $-\dfrac{1}{x}$

(5) $\dfrac{1}{5x^5}$

練習問題 11.3

(1) $\dfrac{3}{x}$ (2) $\log x+1$ (3) $\dfrac{2x}{x^2+1}$

(4) $-2e^{-2x}$ (5) xe^x (6) $\dfrac{e^x}{1+e^x}$

練習問題 11.4

(1) $2x\sqrt{x}$ (2) $x-4\sqrt{x}+\log x$

(3) $x+2\log|x|-\dfrac{1}{x}$ (4) $4e^x+5x$

練習問題 12.2

(1) $\begin{bmatrix} 17 \\ 13 \\ -11 \end{bmatrix}$ (2) $\begin{bmatrix} 3 \\ 1 \\ 11 \end{bmatrix}$ (3) $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

練習問題 12.4

(1) 21 (2) -6 (3) 18

練習問題 13.2

(1) $\begin{bmatrix} 3 & 4 & 1 \\ 7 & -6 & 4 \end{bmatrix}$ (2) $\begin{bmatrix} 1 & -2 & -7 \\ 1 & -6 & 6 \end{bmatrix}$

(3) $\begin{bmatrix} 6 & 3 & -9 \\ 12 & -18 & 15 \end{bmatrix}$ (4) $\begin{bmatrix} 7 & 11 & 6 \\ 17 & -12 & 7 \end{bmatrix}$

練習問題 13.3

(1) $\begin{bmatrix} -5 & -6 \\ -12 & 32 \end{bmatrix}$ (2) $\begin{bmatrix} 10 & -19 & 18 \\ 6 & 3 & -9 \\ 20 & -22 & 14 \end{bmatrix}$

(3) $\begin{bmatrix} 1 \\ 8 \\ 6 \end{bmatrix}$ (4) $\begin{bmatrix} 2 & 3 \\ 4 & -1 \\ 5 & -2 \end{bmatrix}$