統計学を理解するための簡単な数学

三輪哲久

国立研究開発法人
農業・食品産業技術総合研究機構
フェロー

2017

目 次

1　統計学を理解するための簡単な数学

本節（第 1 節）で，統計学を学習するときに良く使われる簡単な数学をまとめておく。ほとんどが中学・高校の数学で学習する内容である。厳密な証明よりも，各数式の持つ意味を理解しておくことが重要である。(厳密な証明に興味のある読者は，高校数学の教科書・参考書を参照されたい)。

状況に応じて，**例・演習問題・注**の形で説明を加えている。内容は多岐にわたっているので，最初から全てを通して読む必要はない。統計手法を習得する過程で必要になったときに確認してもよい。ただし，第 1.2 節の和の記号 $\sum$ に関しては，この記号を駆使できるかどうかが，統計学を理解する上での鍵になる。決して難しい内容ではないので，紙と鉛筆を用意して数式を確認し，この記号の使い方を体で覚えていただきたい。

1.1　統計学と数学

統計学を学習するときに，数学の知識は必要だろうか。市販の入門的な統計書のなかには，数学を使うことを避け，和の記号 $\sum$ すら使用しないものもある。

統計学と数学との関連について，次の簡単な例で考えてみよう。表 1.1 は，同一条件で栽培した 5 つのポットにおける「水稲葉いもち病斑面積率」の観測値を示している。

表 1.1　水稲葉いもち病斑面積率 (%)

ポット番号	病斑面積率	偏差
1	21	−2.0
2	25	2.0
3	20	−3.0
4	21	−2.0
5	28	5.0
合計	115	0.0
平均	23.0	

この 5 つの観測値の平均は

$$\text{平均:}\quad \frac{21+25+20+21+28}{5}=23.0$$

である。また，個々の観測値と平均との差

$$-2.0,\ 2.0,\ -3.0,\ -2.0,\ 5.0$$

を偏差という（表 1.1 にも示されている）。この例で，これら偏差の値を合計すると

$$(-2.0)+2.0+(-3.0)+(-2.0)+5.0=0.0$$

のように，偏差の合計はゼロになる。

さて，我々の扱うデータは，毎回 $\{21, 25, 20, 21, 28\}$ と同じ値が得られるわけではない。そこで，5 つの観測値を

$$x_1, x_2, x_3, x_4, x_5$$

と記号を用いて表わし，平均も $\bar{x}$ という記号で表わすことにすれば，どのような観測値が得られても，平均は

$$\bar{x}=\frac{x_1+x_2+x_3+x_4+x_5}{5}=\frac{1}{5}\sum_{j=1}^{5}x_j$$

と表現することができる（$\sum$ の記号については，以下の第 1.2 節で説明する）。このような数学記号を使わなければ，次に，{30, 19, 25, 27, 29} という値が観測されたとき，いちいち具体的に

$$\text{平均：}\ \frac{31+19+25+27+29}{5}=26.2$$

と書き下さなければばならない。さらに，実験に用いるポットの数も 5 つとは限らず，n 個を用いることにすれば，平均は，もっと一般的に

$$\bar{x}=\frac{x_1+x_2+\cdots+x_n}{n}=\frac{1}{n}\sum_{j=1}^{n}x_j$$

と表わすことができる。これで，何個のポットを実験に用いても，そして，どのような値が観測されても平均を表現することができる。すなわち，数式を用いることにより，個別の具体的なデータではなく，一般的なデータを扱うことができるようになる。

さらに，表 1.1 の例で偏差の合計がゼロになったのは，特別なことなのであろうか。数学を用いなくても，たとえば 5 つのポットを用いた実験を 1000 回行なって，いつも偏差の合計がゼロになることを確認すれば，「偏差の合計はゼロになる」という事実が納得できるかもしれない。しかし，第 1.2 節で示すように，簡単な数学の知識を使えば，どのような n，どのような $x_1, x_2, \ldots, x_n$ に対しても，

$$\sum_{j=1}^{n}(x_j-\bar{x})=0$$

が成り立つことを示すことができる。つまり，数学の知識を用いることにより，本質的な事実を知ることができるのである。もっと役に立つ例として，n 個の独立な確率変数 $x_1, x_2, \ldots, x_n$ の平均 $\bar{x}$ の分散は，個々の確率変数の分散の $1/n$ になるという事実がある:

$$\mathrm{V}[\bar{x}]=\frac{\sigma^2}{n}$$

現場の技術者は，永年の経験から，平均を取ればバラツキが小さくなることを知っている。しかし，どれくらい小さくなるか（すなわち，$1/n$ になるということ）については，経験だけから導くことは困難である。実は簡単な数学を使うことによって，この事実を確かめることができる。

このように，数学の知識を使うことによって，

- 扱う対象を一般化する
- 本質的な事実を確かめる

ことが可能となる。（このことは，統計学に限らない。物理学，経済学，心理学，文学 $\cdots$ など，多くの分野において，数学を便利な道具として使うことができる。）

ただし，統計学を理解するためには高度な数学を必要とするわけではない。たとえば，統計学においては，和の記号 $\sum$（第 1.2 節）が頻出する。この和の記号に関しては，3 つの等式

$$(1)\quad \sum_{j=1}^{n}(x_j+y_j)=\Bigl(\sum_{j=1}^{n}x_j\Bigr)+\Bigl(\sum_{j=1}^{n}y_j\Bigr)$$

$$(2)\quad \sum_{j=1}^{n}(c\,x_j)=c\Bigl(\sum_{j=1}^{n}x_j\Bigr)$$

$$(3)\quad \sum_{j=1}^{n}c=n\cdot c$$

が成り立つことを確認するだけでよい（第 1.2 節）。あとは，この 3 つの等式を積み上げていくだけである。ときには，

$$式_1 = 式_2 = \cdots = \cdots = \cdots = \cdots = \cdots = \cdots = \cdots = 式_{10}$$

のように全体として複雑に見える数式が登場することがある。しかし，この中の 1 つの等号の部分 “$\cdots = \cdots$” は，単純な式変形である場合が多い。以下の第 1.2 節に示すように，上記の 3 つの等式を用いるだけで，「分散分析」までもが理解できるのである。

1.2　和の記号

n 個の数 $x_1, x_2, \ldots, x_n$ の和を，ギリシャ文字 $\sum$（シグマと読む）を用いて

$$\sum_{j=1}^{n} x_j = x_1 + x_2 + \cdots + x_n \tag{1.1}$$

と表わす。(1.1) 式の左辺は，x_j の添え字 j の値を 1 から n まで変化させて和を計算することを意味している。

例 1　n 個の数値 $x_1, x_2, \ldots, x_n$ の合計 T，平均 $\bar{x}$ は，次のように表わされる。

$$合計: T = x_1 + x_2 + \cdots + x_n = \sum_{j=1}^{n} x_j \tag{1.2}$$

$$平均: \bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{j=1}^{n} x_j = \frac{T}{n} \tag{1.3}$$

例 2　$x_j = j\ (j = 1, 2, \ldots, n)$，あるいは $x_j = j^2\ (j = 1, 2, \ldots, n)$ とすると，1 から n までの整数の和と 2 乗和は

$$\sum_{j=1}^{n} j = 1 + 2 + \cdots + n = \frac{n\,(n+1)}{2}$$

$$\sum_{j=1}^{n} j^2 = 1^2 + 2^2 + \cdots + n^2 = \frac{n\,(n+1)(2n+1)}{6}$$

と表わされる。証明は高校数学の教科書・参考書を参照せよ。（あるいは自力で考えよ。2 乗和については数学的帰納法を使うのが早道であろう。）　［例 2 終］

注　“x_j” の右下の小さい文字 “j” を添え字という。添え字には，適宜 i や j や k などを用いる。したがって，$\sum_{i=1}^{n} x_i$ と $\sum_{k=1}^{n} x_k$ とは同じものを表わす。一般に，i, j, k, $l, \ldots$ など，アルファベット列の中央あたりの文字が添え字として使用されることが多い。しかし数学書によっては，s や t を添え字として用いたり，なんとギリシャ文字を添え字として使用したりすることがあるので注意が必要である。

$$\sum_{i=1}^{n} x_i = \sum_{j=1}^{n} x_j = \sum_{k=1}^{n} x_k = \sum_{s=1}^{n} x_s = \sum_{\alpha=1}^{n} x_\alpha = x_1 + x_2 + \cdots + x_n$$

また，“j” は添え字としてだけでなく，実際に j という値を持つ数として現れる場合もある（例 2 参照）。

注　文章中では，行間のスペースを節約するため，$\sum_{j=1}^{n} x_j$ のように表現される場合もある。また，添え字 j の動く範囲が明らかな場合は，$\sum_j x_j$ や $\sum x_j$ のように簡略化して表わすことも多い。

注　スペースを節約するため，x_2 を省略して，“$x_1, \dots, x_n$” や “$x_1 + \cdots + x_n$” のように表わすこともある。

和に関する基本性質

c, d を添え字 j によらない定数，$y_1, \dots, y_n$ を別の n 個の数とすると次の性質が成り立つ。

(1) $$\sum_{j=1}^{n}(x_j + y_j) = \Big(\sum_{j=1}^{n} x_j\Big) + \Big(\sum_{j=1}^{n} y_j\Big) \tag{1.4}$$

(2) $$\sum_{j=1}^{n}(c\,x_j) = c\Big(\sum_{j=1}^{n} x_j\Big) \tag{1.5}$$

(3) $$\sum_{j=1}^{n} c = n \cdot c \tag{1.6}$$

(4) $$\sum_{j=1}^{n}(c\,x_j + d\,y_j) = c\Big(\sum_{j=1}^{n} x_j\Big) + d\Big(\sum_{j=1}^{n} y_j\Big) \tag{1.7}$$

この 4 つの性質は極めて重要であり，自由に扱うことができるようになっておく必要がある。以下，第 1.2 節における数式は，この 4 つの性質を使って導かれる。

演習問題 1　上の性質 (1), (2), (3) を確認せよ。

演習問題 2　(1) と (2) を使って，(4) を導け。逆に，(1) と (2) は (4) の特別な場合であることを確かめよ。

例 3　数値 $x_1, \dots, x_n$ が与えられたとき，各数値 x_j と平均 $\bar{x}$ との差 $x_j - \bar{x}$ $(j = 1, \dots, n)$ は偏差 (deviation) とよばれる。n 個の偏差の合計はゼロである。

$$\sum_{j=1}^{n}(x_j - \bar{x}) = 0 \tag{1.8}$$

演習問題 3　(1.8) 式を確かめよ。

［例 3 終］

例 4　偏差の 2 乗和を偏差平方和，あるいは単に平方和 (sum of squares) という。偏差平方和に関して，次の等式が成り立つ。

$$\sum_{j=1}^{n}(x_j - \bar{x})^2 = \sum_{j=1}^{n} x_j^2 - n\,\bar{x}^2 = \sum_{j=1}^{n} x_j^2 - \frac{T^2}{n} \tag{1.9}$$

演習問題 4　(1.9) 式を確かめよ。

［例 4 終］

注　偏差平方和は n 個の偏差 $x_j - \bar{x}$ $(j = 1, \dots, n)$ の 2 乗和である。しかし，これら n 個の偏差の間には，(1.8) 式のように，合計するとゼロになるという 1 つの制約条件が成り立っている。$n-1$ 個の偏差 $x_1 - \bar{x}, \dots, x_{n-1} - \bar{x}$ の値が決まれば，$x_n - \bar{x}$ の値は自動的に決まってしまう。つまり，n 個の偏差のうち自由に値を取りうるのは $n-1$ 個である。この $n-1$ のことを，平方和の自由度 (degrees of freedom) という。

注　電卓を使って平方和を計算するとき（あるいは筆算で計算するとき）は，(1.9) 式右辺の $\sum_{j=1}^{n} x_j^2 - T^2/n$ が便利である。しかし，コンピュータ等でプログラムを書く場合は，左辺の $\sum_{j=1}^{n}(x_j - \bar{x})^2$ を使った方が一般に精度が高い。

例 5　定数 c, d を用いて，$y_j = c\,x_j + d$ $(j = 1, \ldots, n)$ と変換する。つまり，x_j の各値を c 倍し，そのあと d だけ平行移動する。この y_j に関して，

$$\bar{y} = c\,\bar{x} + d \tag{1.10}$$

$$\sum_{j=1}^{n}(y_j - \bar{y})^2 = c^2 \sum_{j=1}^{n}(x_j - \bar{x})^2 \tag{1.11}$$

が成り立つ。y_j に関する平方和は，平行移動 d の値に依存しないことに注意せよ。

演習問題 5　(1.10) 式，(1.11) 式を確かめよ。

［例 5 終］

例 6　2 組の数値 $\{p_1, \ldots, p_n\}$ と $\{x_1, \ldots, x_n\}$ を考える。ただし，p_j $(j = 1, \ldots, n)$ については，

$$p_j \geq 0, \quad \sum_{j=1}^{n} p_j = 1$$

の条件を満たすものとする。ここで，次の 2 つの数値を定義する。

表 1.2　$p_1, \ldots, p_n$ $(\sum p_j = 1)$

1	2	$\cdots$	j	$\cdots$	n
p_1	p_2	$\cdots$	p_j	$\cdots$	p_n
x_1	x_2	$\cdots$	x_j	$\cdots$	x_n
y_1	y_2	$\cdots$	y_j	$\cdots$	y_n

$y_j = c\,x_j + d$

$$\mu = \sum_{j=1}^{n} x_j \cdot p_j \quad (\mu \text{ はミューと読む}) \tag{1.12}$$

$$\sigma^2 = \sum_{j=1}^{n}(x_j - \mu)^2 \cdot p_j \quad (\sigma \text{ はシグマと読む}) \tag{1.13}$$

このとき，次の等式が成立する。

$$\sum_{j=1}^{n}(x_j - \mu) \cdot p_j = 0 \tag{1.14}$$

$$\sigma^2 = \sum_{j=1}^{n} x_j^2 \cdot p_j - \mu^2 \tag{1.15}$$

（例 6 は，離散確率変数の母平均と母分散の性質を説明している。ここでは，「母平均」，「母分散」の用語が理解できなくても構わない。数式が成り立つことが確認できればよい。）

演習問題 6　(1.14) 式，(1.15) 式を確かめよ。

［例 5 終］

例 7　（例 6 の続き）c, d を定数として，$y_j = c\,x_j + d$ $(j = 1, \ldots, n)$ と変換する。この y_j に関して，(1.12) 式，(1.13) 式と同様に

$$\mu_y = \sum_{j=1}^{n} y_j \cdot p_j$$

$$\sigma_y^2 = \sum_{j=1}^{n}(y_j - \mu_y)^2 \cdot p_j$$

を定義する。ここで，添え字の“ $_y$ ”は，y_j に関する μ，σ^2 の意味である。このとき，

$$\mu_y = c\,\mu + d \tag{1.16}$$
$$\sigma_y^2 = c^2\,\sigma^2 \tag{1.17}$$

が成り立つ。σ_y^2 は d の値（平行移動）に依存しない。（例 7 は，離散確率変数の線形変換の性質を説明している。）

演習問題 7　(1.16) 式，(1.17) 式を確かめよ。

［例 7 終］

例 8　（例 8 と例 9 は，分散分析を学習するまでは読み飛ばしてもよい。）

数値が，2 つの添え字で表現される場合も多い。たとえば m 種類の水稲品種に対して，それぞれ n 個の試験区で実験を行ない，各試験区で収量を観測する場合には，第 i 番めの品種の第 j 番めの試験区での収量は x_{ij} $(i = 1, \ldots, m;\ j = 1, \ldots, n)$ と表わされる。そのデータは m 行 $\times n$ 列の行列として与えられる（表 1.3）。x_{ij} の第 1 の添え字 i は第 i 行を表わし，縦方向に 1 から m まで動く。第 2 の添え字 j は第 j 列を表わし，横方向に 1 から n まで動く。

表 1.3　m 行 $\times n$ 列の数値 x_{ij}

	試験区 j							
品種 i	1	2	$\cdots$	第 j 列	$\cdots$	n	合計	平均
1	x_{11}	x_{12}	$\cdots$	x_{1j}	$\cdots$	x_{1n}	$T_{1\cdot}$	$\bar{x}_{1\cdot}$
2	x_{21}	x_{22}	$\cdots$	x_{2j}	$\cdots$	x_{2n}	$T_{2\cdot}$	$\bar{x}_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$
第 i 行	x_{i1}	x_{i2}	$\cdots$	x_{ij}	$\cdots$	x_{in}	$T_{i\cdot}$	$\bar{x}_{i\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\vdots$
m	x_{m1}	x_{m2}	$\cdots$	x_{mj}	$\cdots$	x_{mn}	$T_{m\cdot}$	$\bar{x}_{m\cdot}$
合計	$T_{\cdot 1}$	$T_{\cdot 2}$	$\cdots$	$T_{\cdot j}$	$\cdots$	$T_{\cdot n}$	$T_{\cdot\cdot}$	
平均	$\bar{x}_{\cdot 1}$	$\bar{x}_{\cdot 2}$	$\cdots$	$\bar{x}_{\cdot j}$	$\cdots$	$\bar{x}_{\cdot n}$		$\bar{x}_{\cdot\cdot}$

第 i 行の合計 $T_{i\cdot}$ と平均 $\bar{x}_{i\cdot}$ $(i = 1, \ldots, m)$ を

$$T_{i\cdot} = \sum_{j=1}^{n} x_{ij} = x_{i1} + x_{i2} + \cdots + x_{in}$$
$$\bar{x}_{i\cdot} = T_{i\cdot}/n$$

と表わす。ここで，添え字の“ . ”は，対応する部分の添え字を動かして合計・平均を計算することを意味している。すなわち $T_{i\cdot}$ については，x_{ij} の添え字 i を固定し，添え字 j を動かして合計を計算する。一方，第 j 列に着目すると，その合計 $T_{\cdot j}$ と平均 $\bar{x}_{\cdot j}$ $(j = 1, \ldots, n)$ は

$$T_{\cdot j} = \sum_{i=1}^{m} x_{ij} = x_{1j} + x_{2j} + \cdots + x_{mj}$$
$$\bar{x}_{\cdot j} = T_{\cdot j}/m$$

と表わされる。さらに，総合計を $T_{\cdot\cdot}$ とすると，

$$\begin{aligned} T_{\cdot\cdot} &= \sum_{i=1}^{m} T_{i\cdot} = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} x_{ij} \right) = \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} \\ &= \sum_{j=1}^{n} T_{\cdot j} = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} x_{ij} \right) = \sum_{j=1}^{n} \sum_{i=1}^{m} x_{ij} \end{aligned}$$

が成り立つ。この式は，i と j のどちらの添え字に関して先に和を計算しても結果が同じになることを示している。$\sum$ 記号の間の括弧は省略される場合が多い。

$m \cdot n$ 個の数値全体での総平均 $\bar{x}..$ は，

$$\bar{x}.. = \frac{T..}{m \cdot n}$$

で定義される。このとき，$m \cdot n$ 個の各数値 x_{ij} と総平均 $\bar{x}..$ との偏差 $x_{ij} - \bar{x}..$ に関して，総平方和は，

$$\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \bar{x}..)^2 = \sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \bar{x}_{i}.)^2 + n\sum_{i=1}^{m}(\bar{x}_{i}. - \bar{x}..)^2 \tag{1.18}$$

のように分解される。（例 8 は，繰返し数の等しい 1 元配置データの分散分析に対応している。）

演習問題 8　(1.18) 式を確認せよ。

［例 8 終］

例 9　数値 x_{ij} に関して，第 i 行ごとに添え字 j の取り得る範囲が異なる場合がある。たとえば，品種 i について n_i 個の試験区で実験を行ない，その試験区の数 $n_1, n_2, \ldots, n_m$ が一定ではない場合である。第 1 行には n_1 個の x_{1j} $(j = 1, \ldots, n_1)$，第 2 行には n_2 個の x_{2j} $(j = 1, \ldots, n_2)$，一般に第 i 行には n_i 個の x_{ij} $(j = 1, \ldots, n_i)$ が存在する場合である（表 1.4）。

表 1.4　異なる繰返し数のデータ x_{ij}

	行ごとに異なる数の x_{ij}					n_i	合計	平均
1	x_{11}	$\cdots$	$\cdots$	$\cdots$	x_{1n_1}	n_1	$T_1.$	$\bar{x}_1. = T_1./n_1$
2	x_{21}	$\cdots$	x_{2n_2}			n_2	$T_2.$	$\bar{x}_2. = T_2./n_2$
$\vdots$	$\vdots$	$\cdots$				$\vdots$	$\vdots$	$\vdots$
第 i 行	x_{i1}	$\cdots$	$\cdots$	x_{in_i}		n_i	$T_i.$	$\bar{x}_i. = T_i./n_i$
$\vdots$	$\vdots$	$\cdots$				$\vdots$	$\vdots$	$\vdots$
m	x_{m1}	$\cdots$	$\cdots$	$\cdots$	x_{mn_m}	n_m	$T_m.$	$\bar{x}_m. = T_m./n_m$
合計						N	$T..$	
平均								$\bar{x}.. = T../N$

このときは，各行ごとに合計 $T_i.$ と平均 $\bar{x}_i.$ が

$$T_i. = \sum_{j=1}^{n_i} x_{ij} = x_{i1} + x_{i2} + \cdots + x_{in_i}$$

$$\bar{x}_i. = \frac{T_i.}{n_i}$$

により計算される。総合計は

$$T.. = \sum_{i=1}^{m} T_i. = \sum_{i=1}^{m}\Big(\sum_{j=1}^{n_i} x_{ij}\Big)$$

である。この場合は，2 つの和の記号について，和を計算する順序を交換することはできない。まず各 i 行ごとに，添え字 j を動かして $T_i. = \displaystyle\sum_{j=1}^{n_i} x_{ij}$ を計算し，次に添え字 i に関して総合計 $T.. = \displaystyle\sum_{i=1}^{m} T_i. = \sum_{i=1}^{m}\big(\sum_{j=1}^{n_i} x_{ij}\big)$ を計算することになる。

データの個数の合計を

$$N = \sum_{i=1}^{m} n_i = n_1 + n_2 + \cdots + n_m$$

とする。総平均は

$$\bar{x}.. = \frac{T..}{N}$$

で与えられる。

この場合の偏差 $x_{ij} - \bar{x}..$ に関する総平方和は

$$\sum_{i=1}^{m}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}..)^2 = \sum_{i=1}^{m}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i.)^2 + \sum_{i=1}^{m} n_i\,(\bar{x}_i. - \bar{x}..)^2 \tag{1.19}$$

と分解される。（例 9 は，繰返し数の不揃いな 1 元配置データの分散分析に対応している。）

演習問題 9　(1.19) 式を確かめよ。

［例 9 終］

例 10　例 8 と同じように，添え字 i は 1 から m まで動き，添え字 j は 1 から n まで動く場合を考える。二つの添え字 i と j で表わされる数値 x_{ij} が，添え字 i のみに依存する数値 p_i と，添え字 j のみに依存する数値 q_j との積

$$x_{ij} = p_i \cdot q_j$$

として与えられる場合も多い。このとき，x_{ij} の総和は，それぞれの和の積になる。

$$\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij} = \sum_{i=1}^{m}\sum_{j=1}^{n} p_i \cdot q_j = \Bigl(\sum_{i=1}^{m} p_i\Bigr) \cdot \Bigl(\sum_{j=1}^{n} \cdot q_j\Bigr) \tag{1.20}$$

演習問題 10　(1.20) 式を確認せよ。

［例 10 終］

注　以上，第 1.2 節の和の記号 $\sum$ の使い方をマスターするだけで，確率変数の母平均・母分散の性質（例 6，例 7）や，1 元配置分散分析（例 8，例 9）を理解できることになる。

注　ここまでの議論では，$\displaystyle\sum_{j=1}^{n} x_j$ のように，添え字 j は 1 から n まで変化することになっていた。しかし，

$$\sum_{j=0}^{n-1} x_j = x_0 + x_1 + \cdots + x_{n-1}$$

のように，添え字 j の変化する範囲は，いろいろな場合がある。そのような場合でも，この節で説明した性質は全て成立する。さらに，

$$\sum_{j=0}^{\infty} p_j = p_0 + p_1 + \cdots + p_n + \cdots$$

のような無限個の数の和（無限級数）を考える場合もある。このときは，無限級数の収束に関する議論が必要になる。しかし初歩の統計学に登場する無限級数は，ほとんどの場合に収束し，本節で説明した性質が成立する。

1.3　積分

関数 $f(x)$ の a から b までの定積分は，図 1.1 のように，$f(x) \geq 0$ $(a \leq x \leq b)$ の場合には，斜線部分の面積 K の値に等しい。

$$\int_a^b f(x)\,dx = K$$

$f(x) \leq 0$ $(a \leq x \leq b)$ の場合（図 1.2）には，面積 K の値にマイナスの符号を付けたものになる。

$$\int_a^b f(x)\,dx = -K$$

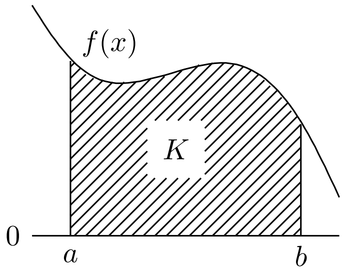

図 1.1　正の関数の積分

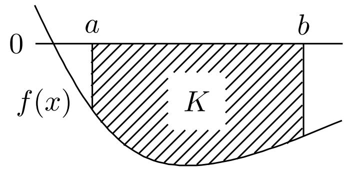

図 1.2　負の関数の積分

図 1.3 のように，$f(x)$ がプラスやマイナスの値を取る場合には，プラスの符号を持つ部分の面積から，マイナスの符号を持つ部分の面積を引いたものになる。

$$\int_a^b f(x)\,dx = K_1 - K_2 + K_3$$

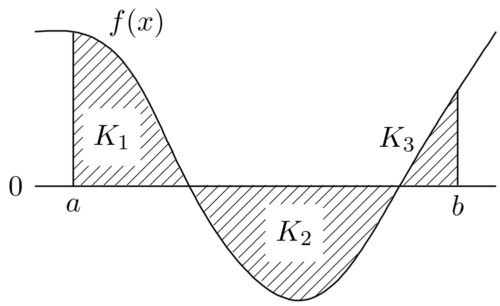

図 1.3　正負の関数の積分

注　定積分の値は，関数 $f(x)$ の形と上下限 a, b が与えられれば確定するので，積分に使われる変数は何でもよい。すなわち，

$$\int_a^b f(x)\,dx = \int_a^b f(t)\,dt = \int_a^b f(z)\,dz$$

である。このことは，和の記号 $\sum$ において，添え字として i や k を自由に使用したことに相当している。

注　定積分は，$f(x)$ の不定積分（微分すると $f(x)$ に等しくなる関数）を用いて定義される場合もある。すなわち，$F(x)$ を

$$\frac{d}{dx}F(x) = f(x)$$

となる関数とすると，関数 $f(x)$ の a から b までの定積分は

$$\int_a^b f(x)\,dx = F(b) - F(a) = \big[F(x)\big]_a^b \tag{1.21}$$

と定義される。ここで，$F(b) - F(a)$ を $\big[F(x)\big]_a^b$ と表わす。ただし，この定義では定積分を直感的に理解することが難しい。面積による定積分の定義と，不定積分による定義との関係は後で説明する。

積分に関する基本性質

c, d を x によらない定数，$g(x)$ を別の関数とすると次の性質が成り立つ。

$$(1)\quad \int_a^b \{f(x) + g(x)\}\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx \tag{1.22}$$

$$(2)\quad \int_a^b \{c\,f(x)\}\,dx = c\int_a^b f(x)\,dx \tag{1.23}$$

$$(3)\quad \int_a^b c\,dx = c\,(b-a) \tag{1.24}$$

$$(4)\quad \int_a^b \{c\,f(x) + d\,g(x)\}\,dx = c\int_a^b f(x)\,dx + d\int_a^b g(x)\,dx \tag{1.25}$$

ここでは証明は省略する。通常は，不定積分を使った定義式 (1.21) から (1) - (4) の性質を導く。しかし面積を使った定義からも，これらの性質が成り立つことは直感的に感じ取ることはできるであろう。

これらの性質は，第 1.2 節の和の演算に関する性質 (1) - (4) と全く同じ形をしている。そして定積分の性質においても，(4) は (1) と (2) から導かれる。逆に，(1) と (2) は，(4) の特別な場合であることも確認される。以下に述べる定積分に関する数式は，第 1.2 節と同様にして，上記の性質 (1) - (4) を使って導くことができる。

例 11　次の性質

$$f(x) \geq 0 \quad (a \leq x \leq b), \qquad \int_a^b f(x)\,dx = 1$$

を満たす関数 $f(x)$ を考える。ここで，次の 2 つの数値

$$\mu = \int_a^b x\,f(x)\,dx \tag{1.26}$$

$$\sigma^2 = \int_a^b (x-\mu)^2\,f(x)\,dx \tag{1.27}$$

を定義する。このとき，次の等式が成立する。

$$\int_a^b (x-\mu)\,f(x)\,dx = 0 \tag{1.28}$$

$$\sigma^2 = \int_a^b x^2\,f(x)\,dx - \mu^2 \tag{1.29}$$

(例 11 は，連続確率変数の母平均と母分散の性質を説明している。)

演習問題 11　(1.28) 式，(1.29) 式を確かめよ。

［例 11 終］

例 12　（例 11 の続き）c, d を定数として，$y = g(x) = c\,x + d$ と変換する。この $y = g(x)$ に関して，(1.26) 式，(1.27) 式と同様に

$$\mu_y = \int_a^b y\,f(x)\,dx = \int_a^b g(x)\,f(x)\,dx$$

$$\sigma_y^2 = \int_a^b (y-\mu_y)^2\,f(x)\,dx = \int_a^b (g(x)-\mu_y)^2\,f(x)\,dx$$

を定義する。このとき，

$$\mu_y = c\,\mu + d \tag{1.30}$$
$$\sigma_y^2 = c^2\,\sigma^2 \tag{1.31}$$

が成り立つ。σ_y^2 は d の値（平行移動）に依存しない。(例 12 は，連続確率変数の線形変換の性質を説明している。)

演習問題 12　(1.30) 式，(1.31) 式を確かめよ。

［例 12 終］

不定積分と定積分の関係

点 c を固定して点 x までの定積分

$$F(x) = \int_c^x f(t)\,dt \tag{1.32}$$

を考える。ここでは，元の関数を変数 t を用いて表わしていることに注意せよ。この値は上端の位置 x に依存するので，x の関数である。この $F(x)$ を x に関して微分すると $f(x)$ になる。すなわち，

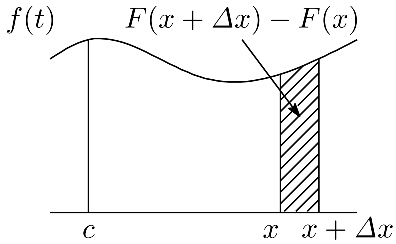

図 1.4　積分と微分の関係

$$\frac{d}{dx}\,F(x) = f(x) \tag{1.33}$$

が成り立つ。証明の概略は以下のとおりである。微分の定義によると

$$\frac{d}{dx}F(x) = \lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x}$$

である。ここで，右辺の分子 $F(x+\Delta x) - F(x)$ は，図 1.4 の斜線で示した柱の部分の面積に等しい。その柱の面積を Δx で割った値は，だいたい柱の高さ $f(x)$ に等しい。そして，$\Delta x \to 0$ の極限をとると $f(x)$ に一致する。

(1.32) 式の関数 $F(x)$ を使えば，a から b までの定積分は，

$$\int_a^b f(x)\,dx = F(b) - F(a) = \big[F(x)\big]_a^b$$

で与えられる (図 1.5)。この $F(x)$ とは別に，導関数が $f(x)$ となる関数 $G(x)$

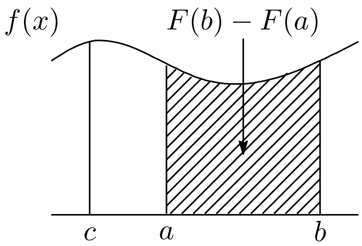

図 1.5　不定積分による定積分

$$\frac{d}{dx}G(x) = f(x)$$

があったとしよう。

$$\frac{d}{dx}\{G(x) - F(x)\} = f(x) - f(x) = 0$$

であるから，

$$G(x) - F(x) = C \quad （C \text{ は定数}）$$

でなければならない。したがって，$G(x) = F(x) + C$ より

$$\big[G(x)\big]_a^b = G(b) - G(b) = F(b) - F(a) = \int_a^b f(x)\,dx$$

が成り立つ。つまり，$\dfrac{d}{dx}F(x) = f(x)$ となる関数が一つ与えられれば，定積分の値

$$\int_a^b f(x)\,dx = F(b) - F(a) = \bigl[F(x)\bigr]_a^b$$

が確定する。

注　ここまでは，$\displaystyle\int_a^b f(x)\,dx$ のように，有限の区間 $a \leq x \leq b$ での定積分を考えてきた。しかし，確率や統計の議論では，$\displaystyle\int_0^\infty f(x)\,dx$ や $\displaystyle\int_{-\infty}^\infty f(x)\,dx$ のような無限区間での定積分を考えることもある。ここで，無限区間での定積分は，

$$\int_0^\infty f(x)\,dx = \lim_{b\to\infty}\int_0^b f(x)\,dx$$
$$\int_{-\infty}^\infty f(x)\,dx = \lim_{\substack{a\to-\infty\\ b\to\infty}}\int_a^b f(x)\,dx$$

と定義される。このときは極限の収束に関する議論が必要になる。しかし和の記号の場合と同様に，初歩の統計学に登場する無限区間での定積分は，ほとんどの場合に収束し，本節で説明した事項が成立する。

注　定積分は，関数が不連続であっても実行することができる。

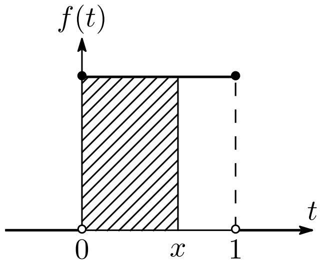

図 1.6　不連続な関数
（一様分布の密度関数）

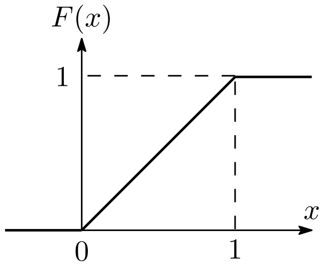

図 1.7　$F(x) = \int_{-\infty}^{x} f(t)\,dt$
（一様分布の分布関数）

例として，図 1.6 に示す次の関数を考える。

$$f(t) = \begin{cases} 0 & -\infty < t < 0 \\ 1 & 0 \leq t \leq 1 \\ 0 & 1 < t < \infty \end{cases}$$

$f(t)$ は，$t = 0$ および $t = 1$ において不連続な関数である（この関数は，一様分布の密度関数を表わしている）。このような不連続な関数であっても，関数 $f(t)$ と x 軸（この場合は t 軸）とで挟まれた部分の面積として定積分を考えることができる。$-\infty$ から x までの定積分は次式で与えられる（図 1.7）。

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \begin{cases} 0 & -\infty < x \leq 0 \\ x & 0 \leq x \leq 1 \\ 1 & 1 \leq x < \infty \end{cases}$$

1.4　微分と最大値・最小値

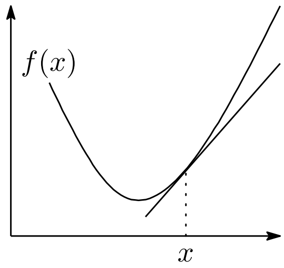

図 1.8　関数の微分

関数 $y = f(x)$ に対して，点 x における微分係数は

$$\frac{d}{dx}f(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \tag{1.34}$$

と定義され，点 $(x, f(x))$ における接線の傾きを表わす（図 1.8）。微分係数は

$$\frac{dy}{dx}, \quad f'(x), \quad \dot{f}(x)$$

などと表わされることもある（$f'(x)$ はエフプライムオブエックスと読む）。

微分係数の値は点 x によって異なる。すなわち，微分係数は x の関数となる。これを関数 $f(x)$ の導関数とよぶ。関数 $f(x)$ から導関数 $\dfrac{d}{dx}f(x)$ を求めることを「関数 $f(x)$ を微分する」という。

注　微分係数の定義 (1.34) 式において，関数 $f(x)$ が点 x において不連続な場合は，その点において微分係数は定義できない。

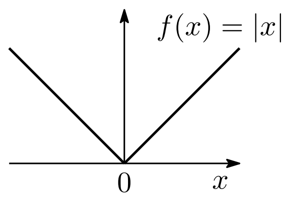

図 1.9　微分不可能

また，関数 $f(x)$ が点 x において連続であっても，$\Delta x > 0$ で $\Delta x \to 0$ の場合と，$\Delta x < 0$ で $\Delta x \to 0$ の場合とで極限の値が異なる場合は，微分係数は定義できない。極限が一致する場合，関数 $f(x)$ は点 x において微分可能であるという。

図 1.9 の関数 $f(x) = |x|$ は，$x = 0$ において連続（$f(0) = 0$）である。しかし，

$$\lim_{\substack{\Delta x > 0 \\ \Delta x \to 0}} \frac{f(0+\Delta x) - f(0)}{\Delta x} = \lim_{\substack{\Delta x > 0 \\ \Delta x \to 0}} \frac{|\Delta x|}{\Delta x} = \lim_{\substack{\Delta x > 0 \\ \Delta x \to 0}} \frac{\Delta x}{\Delta x} = 1$$

$$\lim_{\substack{\Delta x < 0 \\ \Delta x \to 0}} \frac{f(0+\Delta x) - f(0)}{\Delta x} = \lim_{\substack{\Delta x < 0 \\ \Delta x \to 0}} \frac{|\Delta x|}{\Delta x} = \lim_{\substack{\Delta x < 0 \\ \Delta x \to 0}} \frac{-\Delta x}{\Delta x} = -1$$

であるから，関数 $f(x) = |x|$ は，$x = 0$ において微分係数が定まらない（すなわち，微分可能ではない）。

微分に関する基本性質

c, d を x によらない定数，$g(x)$ を別の関数とすると次の性質が成り立つ（証明は省略）。

(1)　$\{f(x) + g(x)\}' = f'(x) + g'(x)$　(1.35)

(2)　$\{c\,f(x)\}' = c\,f'(x)$　(1.36)

(3)　$\dfrac{d}{dx}c = 0$　(1.37)

(4)　$\{c\,f(x) + d\,g(x)\}' = c\,f'(x) + d\,g'(x)$　(1.38)

注　ここまで，「和に関する基本性質」（第 1.2 節），「積分に関する基本性質」（第 1.3 節），「微分に関する基本性質」（第 1.4 節）を説明した。賢明な読者は，そこでの性質 (1) と (2) が，3 つの場合（和，積分，微分）において同じ形をしていることに気づいているであろう。すなわち，次のような対象物と演算

対象物:	数値ベクトル $(x_1, \ldots, x_n)$，関数 $f(x)$，...
演算:	和を取る，積分する，微分する，...

を考えたとき

(1) 演算 (対象物 1 + 対象物 2) = 演算 (対象物 1) + 演算 (対象物 2)

(2) 演算 (スカラー倍 × 対象物) = スカラー倍 × 演算 (対象物)

の関係が，いずれの場合にも成り立っていることが分かる。この性質 (1) と (2) が成り立つとき，その演算は線形 (linear) であるという。性質 (1) と (2) から性質 (4)

(4) 演算 (スカラー倍 1 × 対象物 1 + スカラー倍 2 × 対象物 2)
= スカラー倍 1 × 演算 (対象物 1) + スカラー倍 2 × 演算 (対象物 2)

導くことができ，逆に性質 (1) と (2) は，性質 (4) の特別な場合であることも 3 つの場合（和，積分，微分）に共通している。

その他の導関数の性質

(a) $\{f(x)\,g(x)\}' = f'(x)\,g(x) + f(x)\,g'(x)$ （積の微分公式） (1.39)

(b) $\left\{\dfrac{1}{g(x)}\right\}' = -\dfrac{g'(x)}{g(x)^2}$ （逆数の微分公式） (1.40)

(c) $\left\{\dfrac{f(x)}{g(x)}\right\}' = \dfrac{f'(x)\,g(x) - f(x)\,g'(x)}{g(x)^2}$ （商の微分公式） (1.41)

(d) $y = f(z),\ z = g(x)$（すなわち $y = f(g(x))$）とすると
$\dfrac{dy}{dx} = \dfrac{dy}{dz} \cdot \dfrac{dz}{dx}$ （合成関数の微分公式） (1.42)

(e) $\dfrac{d}{dx} f^{-1}(x) = \dfrac{1}{f'(f^{-1}(x))}$ （逆関数の微分公式） (1.43)

演習問題 13 気にかかる読者は，定義にしたがって (1.39) - (1.43) 式を確かめよ。

例 13 (1.39) 式より，不定積分と定積分の関係式 (1.21) 式を用いれば，部分積分の公式

$$\begin{aligned}
\int_a^b f'(x)\,g(x)\,dx &= \int_a^b \Big(\{f(x)\,g(x)\}' - f(x)\,g'(x)\Big)\,dx \\
&= \int_a^b \{f(x)\,g(x)\}'\,dx - \int_a^b f(x)\,g'(x)\,dx \\
&= \big[f(x)\,g(x)\big]_a^b - \int_a^b f(x)\,g'(x)\,dx \quad \text{（部分積分の公式）} \qquad (1.44)
\end{aligned}$$

が得られる。統計学では，この部分積分公式が頻繁に使われる。 [例 13 終]

例 14 (1.42) 式の合成関数の微分公式から，次の置換積分の公式

$$\int_a^b f(z)\,dz = \int_c^d f(g(x)) \cdot g'(x)\,dx \quad \text{（置換積分の公式）} \qquad (1.45)$$

が得られる。ただし，$z = g(x)$, $a = g(c)$, $b = g(d)$ である。まず $\dfrac{d}{dz}F(z) = f(z)$ となる関数 $F(z)$ を使って，

$$\int_a^b f(z)\,dz = F(b) - F(a)$$

と表わす。ここで,

$$z = g(x), \quad a = g(c), \quad b = g(d)$$

と変換する。合成関数の微分公式から

$$\frac{d}{dx}F(g(x)) = \frac{dF}{dz} \cdot \frac{dz}{dx} = f(z) \cdot g'(x) = f(g(x)) \cdot g'(x)$$

が得られる。したがって，不定積分と定積分の関係式から

$$\int_c^d f(g(x)) \cdot g'(x)\, dx = F(g(d)) - F(g(c)) = F(b) - F(a) = \int_a^b f(z)\, dz$$

が成り立つ。 ［例 14 終］

よく使われる関数の導関数

- $\dfrac{d}{dx}x^n = n\, x^{n-1}$ （任意の自然数 n）
- 実は任意の実数 a に対して $\dfrac{d}{dx}x^a = a\, x^{a-1}$ $\left(\dfrac{d}{dx}\sqrt{x} = \dfrac{1}{2\sqrt{x}}, \quad a = 1/2\right)$
- $\dfrac{d}{dx}e^x = e^x, \quad \dfrac{d}{dx}a^x = (\ln a)\, a^x$
- $\dfrac{d}{dx}\ln x = \dfrac{1}{x}$
- $\dfrac{d}{dx}\sin x = \cos x, \quad \dfrac{d}{dx}\cos x = -\sin x, \quad \dfrac{d}{dx}\tan x = \dfrac{1}{\cos^2 x}$
- $\dfrac{d}{dx}\sin^{-1}(x) = \dfrac{1}{\sqrt{1-x^2}} \quad (-1 < x < 1)$

最大値・最小値

微分可能な関数 $f(x)$ の導関数 $f'(x)$ は点 x における接線の傾きを表わしているので，導関数の値によって，その点での関数の増減の具合を調べることができる。

$f'(x) < 0$	$f(x)$ は減少している。
$f'(x) = 0$	$f(x)$ は停留している（極小，極大，停留）。
$f'(x) > 0$	$f(x)$ は増加している。

関数 $f(x)$ が区間 $a \le x \le b$ において微分可能なとき，その最大値を求めるには， 次のようにする。

1) まず，導関数の値がゼロになる点を探す。その点で関数 $f(x)$ は，極大，または極小，または単なる停留点（x^3 の $x = 0$ のような点）となる。極大値であっても，必ずしも最大値を与えるとは限らない。

2) 極大値と端の値の中から最大値を求める（図 1.10）。

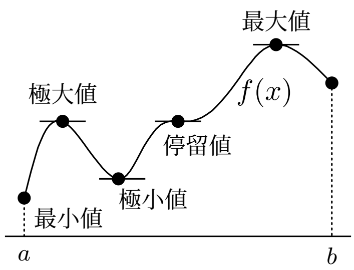

図 1.10 最大値・最小値

最小値に関しても同様である。

例 15 n 個の数値 $x_1, \dots, x_n$ が与えられている。ここで，a を変数とする関数

$$f(a) = \sum_{j=1}^{n}(x_j - a)^2 \tag{1.46}$$

を考える（ここでは，$x_1, \ldots, x_n$ は既知の定数であり，a を変数と考えている）。$f(a)$ を最小にする a の値を求めよう。まず，$f(a)$ を変数 a に関して微分する手順は次のとおりである。

$$
\begin{aligned}
\frac{d}{da}f(a) &= \frac{d}{da}\sum_{j=1}^{n}(x_j-a)^2 \quad \text{（性質 (1) により，微分演算 } \frac{d}{da} \text{ を } \sum \text{ の中に入れる。）} \\
&= \sum_{j=1}^{n}\frac{d}{da}(x_j-a)^2 \quad \text{（} y=z^2,\ z=(x_j-a)^2 \text{ とおいて，まず } z \text{ で微分する。）} \\
&= \sum_{j=1}^{n}2\,(x_j-a)\frac{d}{da}(x_j-a) \quad \text{（次に，} z \text{ を } a \text{ で微分する。「} x_j \text{ は定数」に注意。）} \\
&= \sum_{j=1}^{n}(-2)\,(x_j-a) = -2\,\Bigl(\sum_{j=1}^{n}x_j - \sum_{j=1}^{n}a\Bigr) = -2\,(n\,\bar{x}-n\,a) = 2\,n\,(a-\bar{x})
\end{aligned}
$$

したがって，$a=\bar{x}$ において $f'(a)=0$ となる。この点において実際に $f(a)$ が最小となっていることは，次のような増減表

増減表

a	$a<\bar{x}$	$a=\bar{x}$	$a>\bar{x}$
$f'(a)$	$-$	0	$+$
$f(a)$	$\searrow$	$\sum(x_j-\bar{x})^2$	$\nearrow$

により確認できる。

あるいは，$\sum_{j=1}^{n}(x_j-\bar{x})=0$ に注意すると，$f(a)$ は

$$
\begin{aligned}
f(a) &= \sum_{j=1}^{n}(x_j-a)^2 = \sum_{j=1}^{n}(x_j-\bar{x}+\bar{x}-a)^2 \\
&= \sum_{j=1}^{n}(x_j-\bar{x})^2 + \sum_{j=1}^{n}(\bar{x}-a)^2 + 2\,(\bar{x}-a)\sum_{j=1}^{n}(x_j-\bar{x}) \\
&= \sum_{j=1}^{n}(x_j-\bar{x})^2 + \sum_{j=1}^{n}(\bar{x}-a)^2 = \sum_{j=1}^{n}(x_j-\bar{x})^2 + n\,(\bar{x}-a)^2
\end{aligned}
$$

のように a の 2 次関数として書き直すことができる。第 1 項は変数 a に依存しない値である。この式より $a=\bar{x}$ のときに $f(a)$ が最小になることが分かる。このように $f(a)$ を a の 2 次関数として表わす方法では，$f(a)$ を a で微分するという操作を実行しなくても，$a=\bar{x}$ において $f(a)$ が最小になることが分かる。

平均の特徴づけ

n 個の数値 $x_1, \ldots, x_n$ が与えられたときの平均 $\bar{x}$ は，各 x_j と a との差の 2 乗和

$$
f(a) = \sum_{j=1}^{n}(x_j-a)^2
$$

を最小にする a の値として特徴づけることができる。 ［例 15 終］

例 16　指数関数 $\exp(x)=e^x$ に関して，

$$
e^x > 1+x \quad (x>0) \tag{1.47}
$$

$$
e^x > 1+x+\frac{x^2}{2} \quad (x>0) \tag{1.48}
$$

を証明する。ただし，$e^0=1$, $e^x>1$ $(x>0)$, $\dfrac{d}{dx}e^x=e^x$ の関係は既知とする（第 1.6 節）。

$$f_1(x)=e^x-(1+x)$$

とおくと，

$$\begin{aligned} f_1(0) &= e^0-1=0 \\ f_1'(x) &= e^x-1>0 \quad (x>0) \end{aligned}$$

である。$x>0$ に対して常に $f_1'(x)>0$ であるから $f_1(x)>0$ $(x>0)$，すなわち (1.47) 式が成り立つ。次に，

$$f_2(x)=e^x-(1+x+\frac{x^2}{2})$$

とおくと，

$$\begin{aligned} f_2(0) &= e^0-1=0 \\ f_2'(x) &= e^x-(1+x)=f_1(x)>0 \quad (x>0) \end{aligned}$$

となる。したがって $f_2(x)>0$ $(x>0)$ （すなわち，(1.48) 式）が成り立つ。この関係式を使えば，次の極限が成り立つことが分かる。

$$x\,e^{-x}=\frac{x}{e^x}<\frac{x}{1+x+\dfrac{x^2}{2}}\longrightarrow 0 \quad (x\to\infty) \tag{1.49}$$

さらに，数学的帰納法を使えば，任意の自然数 n に対して

$$\begin{aligned} &e^x>1+x+\frac{x^2}{2}+\cdots+\frac{x^n}{n!} \quad (x>0) \\ &x^n\,e^{-x}\longrightarrow 0 \quad (x\to\infty) \end{aligned}$$

が成り立つことが分かる（余力のある読者は証明を試みよ）。負の引数をもつ指数関数 e^{-x} は $x\to\infty$ のとき急速にゼロに近づくのである（x^n に打ち勝ってゼロに近づく）。 ［例 16 終］

例 17　置換積分・部分積分の公式と例 16 の極限を使って，

$$\int_0^1 \ln z\,dz=-1$$

を証明する。ここで，$\ln z$ は $z=0$ においては定義できないので，上記の積分は，

$$\lim_{a\to 0}\int_a^1 \ln z\,dz=-1$$

を意味している。$x=\ln z$ $(z=e^x)$ と変換する。$a=e^c\to 0$ $(c\to-\infty)$ である。$\dfrac{dz}{dx}=e^x$ であるから，置換積分と部分積分の公式を使えば

$$\begin{aligned} \int_a^1 \ln z\,dz &= \int_c^0 x\,\frac{dz}{dx}\,dx=\int_c^0 x\,e^x\,dx \\ &\qquad (f'(x)=e^x,\ g(x)=x \text{ として部分積分の公式を使う}) \\ &= \big[x\,e^x\big]_c^0-\int_c^0 e^x\,dx=\big[x\,e^x\big]_c^0-\big[e^x\big]_c^0 \\ &= -c\,e^c-1+e^c\to -1 \quad (c\to-\infty) \end{aligned}$$

が成り立つ。 ［例 17 終］

1.5 偏微分

2 つの変数 x, y の関数 $z = f(x,y)$ を考える。一方の変数 y に関して，$y = b$ に固定すると，$f(x,b)$ は x の関数とみなすことができる。これを x に関して微分したものを，

$$\frac{\partial}{\partial x} f(x,b) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, b) - f(x,b)}{\Delta x} \tag{1.50}$$

と表わし，変数 x についての偏微分係数とよぶ。この値は，点 x に依存するとともに，変数 y の固定点 b にも依存する（図 1.11）。すなわち，偏微分係数は x と b の関数となる。これを偏導関数という。通常は，b の代わりに元の y を用いる。

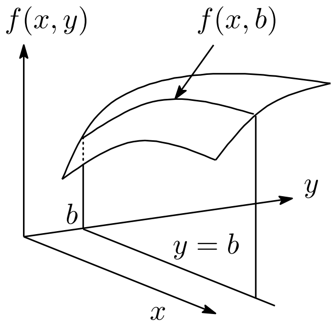

図 1.11　偏微分

$$\frac{\partial}{\partial x} f(x,y) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x, y) - f(x,y)}{\Delta x} \tag{1.51}$$

偏微分の場合は，(1.50) 式や (1.51) 式のように，d の代わりに ∂ の記号が用いられる。また，x の値を固定して，y に関して偏微分を行なう場合も同様に定義される。

例 18　$f(x,y) = x^2 + 3x\,y + 2y^2$ とすると，y を定数と考えて，

$$\frac{\partial}{\partial x} f(x,y) = 2x + 3y$$

が得られる。y についての偏微分も同様に，x を定数と考えて，

$$\frac{\partial}{\partial y} f(x,y) = 3x + 4y$$

となる。　［例 18 終］

ある領域 $(x,y) \in D$ で偏微分可能な関数 $f(x,y)$ について，最大値・最小値を求める方法は，ひとつの変数の場合と同様である。$f(x,y)$ の極値（極大値，極小値）を与える点では，偏導関数の値はゼロになる（図 1.11 参照）。そこで，まず x, y について偏導関数がゼロになる点を求める。

$$\frac{\partial}{\partial x} f(x,y) = 0, \quad \frac{\partial}{\partial y} f(x,y) = 0$$

次に，領域 D の境界を含めて最大値，または最小値を探す。

例 19　（例 19 - 21 は，回帰分析を学習するまでは読み飛ばしてもよい。）
n 組の数値 $(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)$ が与えられている。ここで，a と b を変数とする関数

$$f(a,b) = \sum_{j=1}^{n} (y_j - a - b\,x_j)^2 \tag{1.52}$$

を考える（ここでは，x_j, y_j $(j = 1,\ldots,n)$ は定数であり，a と b を変数と考えている）。

演習問題 14　この $f(a,b)$ に対し，a と b に関する偏微分係数がゼロになるときの a と b は

$$\hat{b} = \frac{\displaystyle\sum_{j=1}^{n} (x_j - \bar{x})\,(y_j - \bar{y})}{\displaystyle\sum_{j=1}^{n} (x_j - \bar{x})^2}, \quad \hat{a} = \bar{y} - \hat{b}\,\bar{x}$$

で与えられることを確かめよ。$\hat{b}$ の分母は，例 4 で説明したように，$x_1, \ldots, x_n$ の偏差平方和（平方和）である。$\hat{b}$ の分子は，$(x_1, y_1), \ldots, (x_n, y_n)$ の偏差積和とよばれる（単に積和ということもある）。

［例 19 終］

例 20　（例 19 の続き）任意の $a,\ b$ に対して

$$
\begin{aligned}
f(a,b) &= \sum_{j=1}^{n}(y_j - a - b\,x_j)^2 = \sum_{j=1}^{n}(y_j - \hat{a} - \hat{b}\,x_j + \hat{a} - a + \hat{b}\,x_j - b\,x_j)^2 \\
&= \sum_{j=1}^{n}(y_j - \hat{a} - \hat{b}\,x_j)^2 + \sum_{j=1}^{n}\left\{(\hat{a} - a) + (\hat{b} - b)\,x_j\right\}^2 \geq \sum_{j=1}^{n}(y_j - \hat{a} - \hat{b}\,x_j)^2 = f(\hat{a}, \hat{b})
\end{aligned}
$$

より，$a = \hat{a},\ b = \hat{b}$ において $f(a,b)$ が最小となることが分かる。ここで，

$$
\begin{aligned}
& y_j - \hat{a} - \hat{b}\,x_j = y_j - \bar{y} - \hat{b}\,(x_j - \bar{x}) \\
& \sum_{j=1}^{n}\{y_j - \bar{y} - \hat{b}\,(x_j - \bar{x})\} = 0 \\
& \sum_{j=1}^{n}\{y_j - \bar{y} - \hat{b}\,(x_j - \bar{x})\}\,x_j = \sum_{j=1}^{n}\{y_j - \bar{y} - \hat{b}\,(x_j - \bar{x})\}\,(x_j - \bar{x}) = 0
\end{aligned}
$$

の関係を利用している。　［例 20 終］

例 21　（例 19 の続き）(1.52) 式の $f(a,b)$ を最小にする a と b を，偏微分を使わないで求めてみよう。b が与えられたとき，$y_j - b\,x_j = z_j$ とおけば，$f(a,b)$ は

$$
f(a,b) = \sum_{j=1}^{n}(y_j - a - b\,x_j)^2 = \sum_{j=1}^{n}(z_j - a)^2
$$

の形をしている。したがって，例 15 の議論により，与えられた b に対して

$$
\hat{a} = \bar{z} = \bar{y} - b\,\bar{x}
$$

のときに $f(a,b)$ は最小となり，その最小値は

$$
g(b) = \sum_{j=1}^{n}(z_j - \bar{z})^2 = \sum_{j=1}^{n}\{y_j - \bar{y} - b\,(x_j - \bar{x})\}^2
$$

である。これは変数 b のみの関数であるから，変数 b で微分して増減表により調べる方法，あるいは変数 b の 2 次関数に展開する方法により，

$$
\hat{b} = \frac{\displaystyle\sum_{j=1}^{n}(x_j - \bar{x})(y_j - \bar{y})}{\displaystyle\sum_{j=1}^{n}(x_j - \bar{x})^2}
$$

のときに $g(b)$ が最小になることが分かる（余力のある読者は確認せよ）。　［例 21 終］

1.6　指数関数と対数関数

指数関数

正の実数 a $(a > 0,\ a \neq 1)$ に対して，関数

$$f(x) = a^x > 0 \quad (-\infty < x < \infty) \tag{1.53}$$

を，a を底（てい）とする指数関数という（$a = 1$ の場合は，常に $a^x = 1^x = 1$ となる)。引数 x は，任意の実数の値を取り得る $(-\infty < x < \infty)$。しかし，関数 $f(x)$ の値は常に正である。

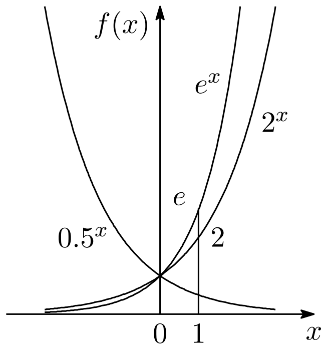

図 1.12　指数関数

指数関数は次の性質を持つ。

(1)　$f(-x) = a^{-x} = \dfrac{1}{a^x}$

(2)　$f(0) = a^0 = 1$

(3)　$a > 1$ のとき $f(x) = a^x$ は単調増加関数で

$$\lim_{x \to -\infty} a^x = 0$$

$$\lim_{x \to \infty} a^x = \infty$$

(4)　$0 < a < 1$ のとき $f(x) = a^x$ は単調減少関数で

$$\lim_{x \to -\infty} a^x = \infty$$

$$\lim_{x \to \infty} a^x = 0$$

(5)　任意の実数 x, y に対して

$$f(x+y) = a^{x+y} = a^x \times a^y = f(x) \times f(y)$$

a として特に自然対数の底 $e = 2.71828\cdots$ を用いた

$$f(x) = e^x = \exp(x) \quad (-\infty < x < \infty) \tag{1.54}$$

を単に指数関数とよぶこともある。e^x は $\exp(x)$ と表わされることも多い。

指数関数 $e^x = \exp(x)$ の特徴のひとつとして，導関数が元の関数に等しいことが挙げられる。

$$\frac{d}{dx}e^x = e^x \tag{1.55}$$

証明は難しい。逆に，一般の指数関数 a^x において，(1.55) 式の関係が成り立つ場合の底 a を $e = 2.71828\cdots$ として定義する方式もある。

例 22　標準正規分布の密度関数は

$$\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2} = \frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{x^2}{2}\Big) \tag{1.56}$$

で与えられる（図 1.13)。

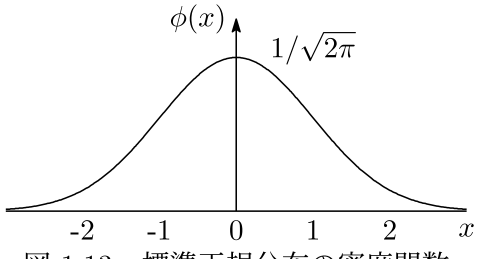

図 1.13　標準正規分布の密度関数

関数 $\exp(-x^2/2)$ は，$x=0$ を中心として左右対称であり，

$$\lim_{x\to-\infty}\exp(-x^2/2)=0$$
$$\lim_{x\to\infty}\exp(-x^2/2)=0$$

となることに注意せよ。また，指数関数の導関数の公式 (1.55) 式と，合成関数の微分公式 (1.42) 式を使えば $\exp(-x^2/2)$ の導関数は

$$\frac{d}{dx}\exp\left(-\frac{x^2}{2}\right)=\exp\left(-\frac{x^2}{2}\right)\cdot\frac{d}{dx}\left(-\frac{x^2}{2}\right)=-x\cdot\exp\left(-\frac{x^2}{2}\right)$$

となる。この式は，正規分布を学習するときに良く登場する。 ［例 22 終］

対数関数

指数関数の逆関数を対数関数という。すなわち，正の定数 a $(a>0,\ a\neq 1)$ を定めたとき，正の実数 x に対して $a^y=x$ となる y を

$$y=f(x)=\log_a x \tag{1.57}$$

と表わし，a を底（てい）とする対数関数とよぶ。

対数関数は $x>0$ に対して定義される。$f(x)=\log_a x$ の値は任意の実数値 y $(-\infty<y<\infty)$ を取り得る。

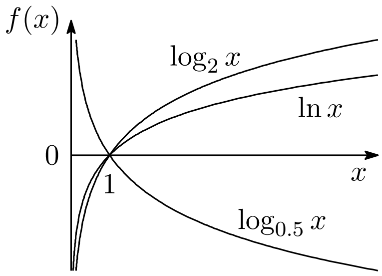

図 1.14　対数関数

対数関数に関しては次の性質が成り立つ。これらの性質は，指数関数の逆関数としての定義 $(y=\log_a x,\ x=a^y)$ から導かれる。

(1)′　$f(1/x)=\log_a(1/x)=-\log_a x$

(2)′　$f(1)=\log_a 1=0$

(3)′　$a>1$ のとき，$f(x)=\log_a x$ は単調増加関数で

$$\lim_{x\to 0}\log_a x=-\infty$$
$$\lim_{x\to\infty}\log_a x=\infty$$

(4)′　$0<a<1$ のとき，$f(x)=\log_a x$ は単調減少関数で

$$\lim_{x\to 0}\log_a x=\infty$$
$$\lim_{x\to\infty}\log_a x=-\infty$$

(5)′　任意の正の実数 $x>0, y>0$ に対して

$$f(x\,y)=\log_a(x\,y)=\log_a x+\log_a y=f(x)+f(y)$$

底として $a=10$ を用いた対数 $\log_{10}x$ を常用対数とよぶ。文献によっては常用対数 $\log_{10}x$ を単に $\log x$ と表わすことがある。自然対数の底 $e=2.71828\cdots$ を底とする対数を自然対数とよび，$\log_e x=\ln x$ と表わす（定義が循環しており，不自然な日本語になっている）。別の文献によっては，自然対数 $\log_e x=\ln x$ を単に $\log x$ と表わしている場合があり，注意が必要である。常に，

自然対数を $\ln x$

常用対数を $\log_{10}x$

と表わすことにすれば混乱を防ぐことができる。

自然対数 $f(x) = \ln x$ の導関数は

$$\frac{d}{dx}\ln x = \frac{1}{x}, \quad (x \neq 0) \tag{1.58}$$

（より正確には，$\dfrac{d}{dx}\ln|x| = \dfrac{1}{x}$）

で与えられる。この関係式は，$\dfrac{d}{dx}e^x = e^x$ を使えば証明できる。

1.7　順列と組合せ

順列

n 個の異なる物から r 個を選んで並べたものを順列という。可能な順列の総数を ${}_nP_r$ と表わすと，${}_nP_r$ は

$${}_nP_r = \underbrace{n\,(n-1)\,(n-2)\cdots(n-r+1)}_{r\text{ 項の積}} \tag{1.59}$$

で与えられる。

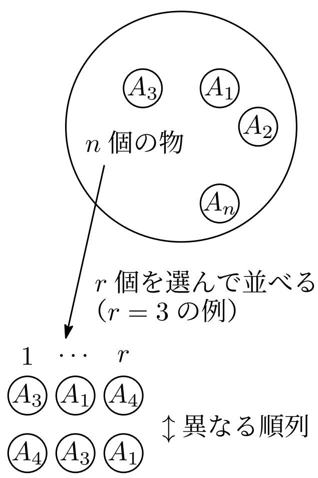

図 1.15　順列

注　証明の概略。n 個の異なる物を $A_1, A_2, \ldots, A_n$ とする。そこから r 個を選んで並べるための r 個の場所を用意する。

場所			
${}^1\square$	${}^2\square$	$\cdots$	${}^r\square$
$\uparrow$	$\uparrow$	$\cdots$	$\uparrow$
n とおり	$n-1$ とおり	$\cdots$	$n-r+1$ とおり

第 1 番めの場所に置く物としては，n とおりの可能性がある。第 1 番めが決まると，第 2 番めの場所に置く物としては，$n-1$ とおりの可能性が残されている。以下，順番に場所を埋めていって，第 r 番めの場所については，$r-1$ 個の物がすでに使われているので，$n-(r-1)$ とおりの可能性が残されている。

とくに $n = r$ のときは,

$${}_nP_n = n\,(n-1)\,(n-2)\cdots 3\cdot 2\cdot 1 = n\,! \tag{1.60}$$

となる。$n\,!$ は 1 から n までの整数を乗じたもので，n の階乗という。階乗の記号を使えば，順列の総数 ${}_nP_r$ は

$${}_nP_r = n\,(n-1)\,(n-2)\cdots(n-r+1) = \frac{n\,!}{(n-r)\,!} \tag{1.61}$$

と表わすことができる。ただし，$0\,! = 1$ と約束する。

演習問題 15　(1.61) 式を確認せよ。

組合せ

n 個の異なる物から，r 個を選んだものを組合せという。つまり，組合せにおいては，順列の場合とは違って選ばれる順序は問題にしない。同じ r 個の物が選ばれていれば，同じ組合せと考える。組合せの総数を ${}_nC_r$ と表わすと，${}_nC_r$ は

$$
{}_nC_r = \binom{n}{r} = \frac{{}_nP_r}{r\,!} = \frac{n\,!}{r\,! \times (n-r)\,!} \tag{1.62}
$$

で与えられる。

注　証明の概略。まず，順列の総数は ${}_nP_r$ である。その中の 1 つの順列に注目する。そこに含まれる r 個の物の並べ方の総数は $r\,!$ とおりある。これらは同じ組合せと考えるので，組合せの総数は ${}_nC_r = {}_nP_r/r\,!$ となる。図 1.15 の例では，$r = 3$ 個の物として，A_1, A_3, A_4 が選ばれている。そして，$3\,! = 6$ とおりの順列

$$
\begin{aligned}
&\{A_1, A_3, A_4\},\ \{A_1, A_4, A_3\},\ \{A_3, A_1, A_4\}, \\
&\{A_3, A_4, A_1\},\ \{A_4, A_1, A_3\},\ \{A_4, A_3, A_1\}
\end{aligned}
$$

は同じ組合せと考える。

注　日本では組合せの総数として記号 ${}_nC_r$ がよく使われる。外国の文献では $\displaystyle\binom{n}{r}$ の記号が使われる場合が多い。しかし，文章中で記号 $\displaystyle\binom{n}{r}$ が使われると，この記号のためだけに行間を広くしなければならず，ひどく迷惑する。そのため文章中ではフォントを小さくし，$\binom{n}{r}$ とすることもある。しかし，この小さいフォントの $\binom{n}{r}$ は，あまり美しいとはいえない。

注　$r = 0$ の場合も，定義に従って

$$
{}_nC_0 = \frac{n\,!}{0\,! \times (n-0)\,!} = \frac{n\,!}{0\,! \times n\,!} = 1
$$

と約束する。

(1.62) 式から明らかなように，組合せの重要な性質として

$$
{}_nC_r = {}_nC_{n-r}, \quad \binom{n}{r} = \binom{n}{n-r} \tag{1.63}
$$

の関係が成り立つ。とくに r が n に近いときは，この式が有効である。たとえば ${}_{10}C_8$ は

$$
{}_{10}C_8 = {}_{10}C_2 = \frac{{}_{10}P_2}{2\,!} = \frac{10 \cdot 9}{2 \cdot 1} = 45
$$

のように計算できる。

2 項定理

$$
(p+q)^n = \overbrace{(p+q)\,(p+q)\cdots(p+q)}^{n\text{ 項の積}} = \sum_{r=0}^{n} {}_nC_r\, p^r\, q^{n-r} \tag{1.64}
$$

注　証明の概略。$(p+q)^n$ を展開すると，$n+1$ とおりの $p^r\, q^{n-r}$ $(r = 0, 1, \ldots, n)$ の形の項が現れる。特定の $p^r\, q^{n-r}$ に着目すると，n 項の $(p+q)$ のうちの r 項から p を取りだして掛け合わせることになる。その選び方は ${}_nC_r$ とおりあるので，$p^r\, q^{n-r}$ の係数が ${}_nC_r$ となる。

1.8 ギリシャ文字

小文字	大文字	英語綴り	[英国式発音] / [アメリカ式発音]	通常の日本語発音
α	A	alpha	[ǽlfə]	アルファ
β	B	beta	[bíːtə] / [béitə, bíːtə]	ベータ
γ	Γ	gamma	[gǽmə]	ガンマ
δ	Δ	delta	[déltə]	デルタ
ε, ϵ	E	epsilon	[epsáilən, épsilən, -lɔ̀n,] / [épsəlɔ̀n, -lən]	イプシロン
ζ	Z	zeta	[zíːtə] / [zéitə, zíːtə]	ジータ
η	H	eta	[íːtə] / [éitə, íːtə]	イータ（エータ）
θ	Θ	theta	[θíːtə] / [θéitə, θíːtə]	シータ
ι	I	iota	[aióutə]	イオタ
κ	K	kappa	[kǽpə]	カッパ
λ	Λ	lambda	[lǽmdə]	ラムダ
μ	M	mu	[mjuː] / [mjuː, muː]	ミュー
ν	N	nu	[njuː] / [nuː, njuː]	ニュー
ξ	Ξ	xi	[sai, ksai, gzai, zai] / [zai, sai]	グザイ（クシー）
o	O	omicron	[oumáikrɔ̀n, əm-, -krən, ɔ́mik-] / [ɔ́mikrɔ̀n, óumikrɔ̀n]	オミクロン
π	Π	pi	[pai]	パイ
ρ	P	rho	[rou]	ロー
σ	Σ	sigma	[sígmə]	シグマ
τ	T	tau	[tau, tɔː]	タウ
υ	Υ	upsilon	[júːpsilən, -lɔ̀n, juːpsáilən, ʌps-] / [ʌ́psəlɔ̀n, júːpsəlɔ̀n]	ユプシロン
ϕ, φ	Φ	phi	[fai]	ファイ
χ	X	chi	[kai]	カイ
ψ	Ψ	psi	[psai] / [sai, psai]	プサイ
ω	Ω	omega	[óumigə, oumíːgə] / [ouméigə, oumégə, oumíːgə]	オメガ

注 英語発音に関しては，手元にある複数の辞書で調べた結果を載せた。アメリカ式発音は，英国式と異なる場合のみ載せている。辞書によっては，ここに載せたもの以外の発音を示すこともある。

注 小文字を一列に並べると次のようになる。

$$\alpha\ \beta\ \gamma\ \delta\ \varepsilon\ \zeta\ \eta\ \theta\ \iota\ \kappa\ \lambda\ \mu\ \nu\ \xi\ o\ \pi\ \rho\ \sigma\ \tau\ \upsilon\ \phi\ \chi\ \psi\ \omega$$

γ（ガンマ）や χ（カイ）は，基線から下にはみ出しており，r（アール）や x（エックス）とは完全に異なる文字である。

ガンマ　カイ
γ r χ x X
アール　エックス（小，大）

注 プログラムなどを作成するときに，$\lambda = 1$ のつもりで “`RAMUDA = 1.0`” などと書かないように注意せよ。

注 24 個のギリシャ文字を全て手書きで書けるとともに，正しく発音できるようにしておくと尊敬される（誰に?）。

1.9　付録（演習問題略解）

演習問題 1

$$(1)\quad \sum_{j=1}^{n}(x_j+y_j) = (x_1+y_1)+\cdots+(x_n+y_n)$$
$$= (x_1+\cdots+x_n)+(y_1+\cdots+y_n) = \Big(\sum_{j=1}^{n}x_j\Big)+\Big(\sum_{j=1}^{n}y_j\Big)$$

$$(2)\quad \sum_{j=1}^{n}(c\,x_j) = c\,x_1+\cdots+c\,x_n = c\,(x_1+\cdots+x_n) = c\Big(\sum_{j=1}^{n}x_j\Big)$$

$$(3)\quad \sum_{j=1}^{n}c = \underbrace{c+\cdots+c}_{n\text{ 個の和}} = n\cdot c$$

演習問題 2

$$(4)\quad \sum_{j=1}^{n}(c\,x_j+d\,y_j) = \Big(\sum_{j=1}^{n}c\,x_j\Big)+\Big(\sum_{j=1}^{n}d\,y_j\Big)\quad \text{（性質 (1) より）}$$
$$= c\Big(\sum_{j=1}^{n}x_j\Big)+d\Big(\sum_{j=1}^{n}y_j\Big)\quad \text{（性質 (2) より）}$$

逆に，(4) において $c=d=1$ とすれば (1) が得られる。また，$d=0$ とすれば (2) が得られる。

演習問題 3

$$\sum_{j=1}^{n}(x_j-\bar{x}) = \sum_{j=1}^{n}x_j-\sum_{j=1}^{n}\bar{x} = T-n\,\bar{x} = 0$$

第 1 の等式は性質 (4)，第 2 の等式は性質 (3)，第 3 の等式は平均の定義による。

演習問題 4　$\bar{x}$ が添え字 j に依存しない定数であることと，$\sum x_j = n\,\bar{x}$ に注意して，

$$\sum_{j=1}^{n}(x_j-\bar{x})^2 = \sum(x_j^2-2\,\bar{x}\cdot x_j+\bar{x}^2) = \sum x_j^2-2\,\bar{x}\sum x_j+\sum\bar{x}^2$$
$$= \sum x_j^2-2\,\bar{x}\cdot(n\,\bar{x})+n\,\bar{x}^2 = \sum x_j^2-n\,\bar{x}^2 = \sum x_j^2-T^2/n$$

演習問題 5

$$\bar{y} = \frac{1}{n}\sum_{j=1}^{n}y_j = \frac{1}{n}\sum(c\,x_j+d) = \frac{c}{n}\sum x_j+\frac{1}{n}\sum d = c\,\bar{x}+d$$

$$\sum_{j=1}^{n}(y_j-\bar{y})^2 = \sum\{(c\,x_j+d)-(c\,\bar{x}+d)\}^2 = \sum\{c\,(x_j-\bar{x})\}^2 = c^2\sum(x_j-\bar{x})^2$$

演習問題 6　$\sum p_j = 1$ に注意して，

$$\sum_{j=1}^{n}(x_j - \mu)\cdot p_j = \sum x_j \cdot p_j - \sum \mu \cdot p_j = \mu - \mu \sum p_j = \mu - \mu = 0$$

$$\begin{aligned}\sigma^2 &= \sum_{j=1}^{n}(x_j - \mu)^2 \cdot p_j = \sum (x_j^2 - 2\mu\, x_j + \mu^2)\cdot p_j = \sum x_j^2 \cdot p_j - 2\mu \sum x_j \cdot p_j + \mu^2 \sum p_j \\ &= \sum x_j^2 \cdot p_j - 2\mu \cdot \mu + \mu^2 = \sum x_j^2 \cdot p_j - \mu^2\end{aligned}$$

演習問題 7

$$\begin{aligned}\mu_y &= \sum_{j=1}^{n} y_j \cdot p_j = \sum (c\, x_j + d)\cdot p_j = c \sum x_j \cdot p_j + d \sum p_j = c\,\mu + d \\ \sigma_y^2 &= \sum_{j=1}^{n}(y_j - \mu_y)^2 \cdot p_j = \sum \{(c\, x_j + d) - (c\,\mu + d)\}^2 \cdot p_j = \sum \{c\,(x_j - \mu)\}^2 \cdot p_j \\ &= c^2 \sum (x_j - \mu)^2 \cdot p_j = c^2\,\sigma^2\end{aligned}$$

演習問題 8

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \bar{x}..)^2 &= \sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \bar{x}_i. + \bar{x}_i. - \bar{x}..)^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \bar{x}_i.)^2 + 2\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \bar{x}_i.)(\bar{x}_i. - \bar{x}..) + \sum_{i=1}^{m}\sum_{j=1}^{n}(\bar{x}_i. - \bar{x}..)^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \bar{x}_i.)^2 + n\sum_{i=1}^{m}(\bar{x}_i. - \bar{x}..)^2\end{aligned}$$

ここで，$\bar{x}_i. - \bar{x}..$ は添え字 j に関しては定数なので，

$$\begin{aligned}&\sum_{j=1}^{n}(x_{ij} - \bar{x}_i.)(\bar{x}_i. - \bar{x}..) = (\bar{x}_i. - \bar{x}..)\sum_{j=1}^{n}(x_{ij} - \bar{x}_i.) = 0 \\ &\sum_{j=1}^{n}(\bar{x}_i. - \bar{x}..)^2 = n\,(\bar{x}_i. - \bar{x}..)^2\end{aligned}$$

の関係を利用している。

演習問題 9

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}..)^2 &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i. + \bar{x}_i. - \bar{x}..)^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i.)^2 + 2\sum_{i=1}^{m}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i.)(\bar{x}_i. - \bar{x}..) + \sum_{i=1}^{m}\sum_{j=1}^{n_i}(\bar{x}_i. - \bar{x}..)^2 \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i.)^2 + \sum_{i=1}^{m} n_i\,(\bar{x}_i. - \bar{x}..)^2\end{aligned}$$

演習問題 8 と同様に

$$\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_{i\cdot})(\bar{x}_{i\cdot}-\bar{x}_{\cdot\cdot})=0$$
$$\sum_{j=1}^{n_i}(\bar{x}_{i\cdot}-\bar{x}_{\cdot\cdot})^2=n_i\,(\bar{x}_{i\cdot}-\bar{x}_{\cdot\cdot})^2$$

の関係を利用している。

演習問題 10

$$\begin{aligned}\sum_{i=1}^{m}\Big(\sum_{j=1}^{n}x_{ij}\Big) &= \sum_{i=1}^{m}\Big(\sum_{j=1}^{n}p_i\cdot q_j\Big) \quad \text{（}p_i\text{ が }j\text{ に依存しないので}\sum_{j=1}^{n}\text{の外に出る）}\\ &= \sum_{i=1}^{m}\Big(p_i\sum_{j=1}^{n}q_j\Big) \quad \text{（}\sum_{j=1}^{n}q_j\text{ が }i\text{ に依存しない定数になり}\sum_{i=1}^{m}\text{の外に出る）}\\ &= \Big(\sum_{i=1}^{m}p_i\Big)\cdot\Big(\sum_{j=1}^{n}\cdot q_j\Big)\end{aligned}$$

演習問題 11　省略。演習問題 6 と同様。

演習問題 12　省略。演習問題 7 と同様。

演習問題 13　積の微分の (1.39) 式，合成関数の微分の (1.42) 式を証明する。

$$\begin{aligned}\frac{d}{dx}\{f(x)\,g(x))\} &= \lim_{\Delta x\to 0}\frac{f(x+\Delta x)\,g(x+\Delta x)-f(x)\,g(x)}{\Delta x}\\ &= \lim_{\Delta x\to 0}\frac{f(x+\Delta x)\,g(x+\Delta x)-f(x)\,g(x+\Delta x)+f(x)\,g(x+\Delta x)-f(x)\,g(x)}{\Delta x}\\ &= \lim_{\Delta x\to 0}\frac{\{f(x+\Delta x)-f(x)\}\,g(x+\Delta x)}{\Delta x}+\lim_{\Delta x\to 0}\frac{f(x)\{g(x+\Delta x)-g(x)\}}{\Delta x}\\ &= f'(x)\,g(x)+f(x)\,g'(x)\end{aligned}$$

逆数の微分の (1.40) 式，商の微分の (1.41) 式は，積の微分の (1.39) 式の応用として導くことができる（自ら試みよ）。

合成関数の微分公式 (1.42) 式については，$y=f(z)$, $z=g(x)$ において，$g(x+\Delta x)=z+\Delta z$ とおくと，

$$\begin{aligned}\frac{f(g(x+\Delta x))-f(g(x))}{\Delta x} &= \frac{f(g(x+\Delta x))-f(g(x))}{g(x+\Delta x)-g(x)}\cdot\frac{g(x+\Delta x)-g(x)}{\Delta x}\\ &= \frac{f(z+\Delta z)-f(z)}{\Delta z}\cdot\frac{g(x+\Delta x)-g(x)}{\Delta x}\to\frac{dy}{dz}\cdot\frac{dz}{dx}\end{aligned}$$

が得られる。逆関数の微分については，合成関数の微分公式を利用して，

$$f(f^{-1}(x))=x$$

の両辺を x で微分すればよい。

演習問題 14　$f(a,b)=\sum_{j=1}^{n}(y_j-a-b\,x_j)^2$ を a について偏微分すると，b を定数と考えて，

$$\frac{\partial}{\partial a}f(a,b)=-2\sum(y_j-a-b\,x_j)$$

となる。同様に b について偏微分すると，

$$\frac{\partial}{\partial b} f(a,b) = -2 \sum x_j \left(y_j - a - b\,x_j\right)$$

が得られる。この二つの式をゼロにする a, b を $\hat{a}, \hat{b}$ とおくと，

$$\begin{aligned}
&(1) \quad \sum (y_j - \hat{a} - \hat{b}\,x_j) = n\,\bar{y} - n\,\hat{a} - n\,\hat{b}\,\bar{x} = 0 \\
&(2) \quad \sum x_j\,(y_j - \hat{a} - \hat{b}\,x_j) = \sum x_j\,y_j - n\,\hat{a}\,\bar{x} - \hat{b} \sum x_j^2 = 0
\end{aligned}$$

となる。まず (1) 式より

$$(3) \quad \hat{a} = \bar{y} - \hat{b}\,\bar{x}$$

である。(3) 式を (2) 式に代入すると，

$$(4) \quad \sum x_j\,y_j - n\,\bar{x}\,\bar{y} - \hat{b}\,(\sum x_j^2 - n\,\bar{x}^2) = 0$$

が得られる。ここで，例 4 （演習問題 4）より，

$$\sum x_j^2 - n\,\bar{x}^2 = \sum (x_j - \bar{x})^2$$

が成り立つ。同様な計算により，

$$\sum (x_j - \bar{x})(y_j - \bar{y}) = \sum x_j\,y_j - \bar{x} \sum y_j - \bar{y} \sum x_j + n\,\bar{x}\,\bar{y} = \sum x_j\,y_j - n\,\bar{x}\,\bar{y}$$

が成り立つ。したがって，(4) 式より

$$(5) \quad \hat{b} = \frac{\displaystyle\sum_{j=1}^{n} (x_j - \bar{x})(y_j - \bar{y})}{\displaystyle\sum_{j=1}^{n} (x_j - \bar{x})^2}$$

が得られる。

演習問題 15

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