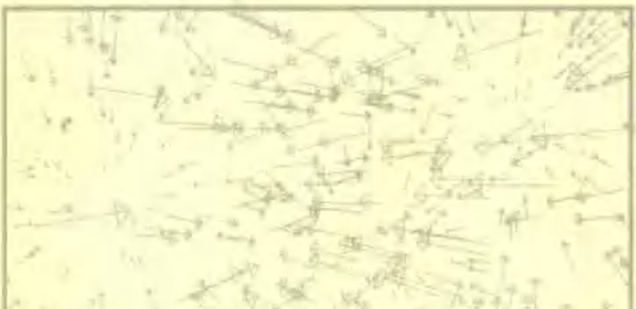
岩波講座 基礎工学 3

確率統計現象 I

瀧 保夫編



岩波書店

岩波講座　基礎工学 3

確率統計現象

I

瀧　　保夫
茅　　陽一
宮川　　洋
関根泰次

岩波書店

岩波講座　基礎工学　**3 確率統計現象 I**　（全 19 巻／第 1 回配本）

1967 年 11 月 25 日　第 1 刷発行 ©

東京都千代田区神田一ツ橋 2-3　株式会社 岩波書店／精興社印刷・松岳社製本

はじめに

この講義は，確率統計的な現象に対する基礎概念を与えるとともに，この様な現象に対処するための手法を解説することを目的としている．

工学において，確率統計的な考え方が重要であることは，今更いうまでもないことであろう．簡単な機械を設計する場合を考えて見ても，まずそれに必要なデータが完全にそろっているとは限らないし，また将来その機械が如何なる状態で使用されるかを完全に予想することもできない．工学では，過去の不完全なデータを基にし，将来の状態に対しある予測を立て，それ等から何等かの推論を経てある決断を行ない，それを実行に移すという過程が屢々起こる．ここでは未知のもの，不確定な要素が，対象とされなければならない．このような不確定的な現象を取り扱う手段を与えるのが確率統計論である．このように，確率統計的なものの見かたが，工学全般に対して不可欠のものであるにもかかわらず，従来の工学教育において，これが十分とり入れられているとはいい難い．

一方，今日においては，数学者によって書かれた確率統計に関するすぐれた著書が数多く出版されているのも事実である．しかし，数学としての確率論，統計論は，それ自身相当高度の数学であって，初学者特に応用を主とする工学部学生には，近寄り難い面も少なくない．基礎工学の立場から，工学部学生に要求されることは，数学としての確率統計論の体系を学ぶことではなく，現実の現象に直面してこれを確率的統計的に把握する能力と，これに対処する手段を身につけることである．本講義も，この趣旨に則って執筆した．従って，数学的な体系や厳密性にこだわらず，常に応用を頭に置いて，それに必要な基礎的事項の解説に重点をおいた．その意味から，本講義の名称も，確率統計現象とした．また，数学には全く素人の我々著者が，敢えて本講義の執筆を引き受けた所以でもある．しかし，上述のような意図の下に出発したにもかかわらず，出来上がった本講義が甚だ理想から遠いものになったのは，ひとえに著者の浅

学菲才によるものである.

第1章は，確率統計現象の概念を与え，又これらを取り扱う数学的手段としての確率の初等的解説を試みたもので，準備的な章である．第2章では，不規則変数とその確率分布関数について説明し，確率論の基礎的事項を取り扱った．第3章は，母集団からの標本という概念を解説したものであり，第4章は，これらの標本から母集団そのものの性質について推測を行なう統計的推測，或いは更にこれらのデータを基にして何等かの決定を行なう統計的決定の方法について解説したものである．これらは，問題を確率統計的に把握し処理するために，極めて重要な考え方である．第5章は，最小二乗法と回帰推定および分散分析の手法を述べ，実験計画法の基礎を与えたものである．第6章では，時間パラメーターを含む不規則過程の概念とその取扱い方の初歩的解説を試みた．不規則過程は，物理学，工学において極めて重要なものであるが，理論的取扱いがやや高級になるので，本講義では，読者が今後さらに深く学ぶためのいと口を与える意味で，最小限の記述にとどめざるを得なかった．本講義の内容は，本講座の全般にわたって関連を持つが，特に熱力学，情報論，制御工学，システム工学等を学ぶためには不可欠のものである.

本講義を草するに当たって，各方面から貴重な御意見を頂いた．特に本講座の編集委員をはじめ，椹木義一，高村仁一，得丸英勝，森美郎(以上京都大学)，伊沢計介，岡田利弘，進藤益男，高島洋一(以上東京工業大学)，中村伝(大阪大学)，並木美喜雄(早稲田大学)，大川善邦(山形大学)，横堀武夫(東北大学)の諸先生は，草稿を閲読し詳細な御意見御叱正を賜った．ここに深甚な謝意を表する．これ等の御意見は，いずれも誠に当を得たものであり，本講義の企画の核心をつくものであった．著者は極力これをとり入れることに努力したが，場合により相矛盾するものもあり，本講義の全体的バランスを考えて敢えて従わなかったものもあるが，著者の力及ばなかった点が少なくない.

目　　次

確立統計現象 II 目次

表紙カット製作：
渡辺茂・槌屋治紀・藤野孝爾・幸村真佐男
(IBM データセンター設置の IBM7090 およびプロッター使用)

第1章

確率統計現象と確率

本章は，確率統計現象の概念を説明し，かつこのような現象を理論的に取り扱うための数学的手段としての確率の基礎を与え，またその基本的な性質を導いて，後の各章の準備をするためのものである．

1.1 確率統計現象

いま重力の場の中の質点の運動を考えて見る．この質点の運動はニュートンの運動方程式によって規定されているから，適当な初期条件の下でこの方程式を解くことによって運動は完全に決定される．すなわち，初期時点における質点の位置とその速度(初期条件)が与えられれば，それ以後のすべての時刻における質点の位置と速度が一義的に決まる．このような現象を**確定現象**という．確定現象においては，原因と結果の間に，確定的な因果関係が成り立っている．力学や電磁気学が対象としているのは，主としてこのような現象である．

しかし現実の世界の現象は，必ずしもこのように簡単ではない．たとえば，ボールを投げる場合を考えて見ると，ボールも力学の法則に従って運動するのであるから，その初期条件が与えられれば，以後の運動は確定するはずである．しかし，もし人間がボールを投げるとすると，その初期条件を完全に一定に保つことは不可能である．またボールが飛行中に風が吹けばその運動は影響を受けるであろう．ボールを投げるという簡単な実験においても，その運動を規定するいろいろな要素をすべて知り，それを完全に取り入れて解析することは不可能に近い．その結果，“ボールを一定の力で投げる”という原因に対して，“ボールの飛距離”という結果が，必ずしも一義的には決まらない．これは，この例のような荒っぽい実験でなく，精密な装置を使い細心の注意を払った実験においても同様であり，一定条件の下で行なった測定に対しても，その測定値

が常に同じ値になるとは限らない．

さいころをふるような場合には，この傾向はさらに著しい．全く同じように投げても，出る目の数は全く不確定で予想がつかないのが普通である．これは，その運動が極めて複雑で，初期条件のわずかな違いによって，その結果がまるで違ったものになるからである．実際に我々がぶつかる現象には，その現象が非常に複雑で条件を一定に保つことが困難であるとか，現象に影響する要因が多数あってその影響の仕方が明確でないとか，あるいはさらに，現象を支配する法則そのものが不明であるとか，いろいろの原因によって，事実上一定の条件の下でも，結果が一つに決まらない現象が少なくない．むしろ，現実には真に確定的な現象は存在しないといってよいであろう．

このように，実際の現象には必ず不確定の要素が伴う．そこで，一定の原因に対して，結果が一つに決まらないような現象を抽象的に想定し，これを**不確定現象**と呼ぶ．実際の現象にぶつかったとき，それに伴う不確定さが，我々の要求する精度より十分小さいものであれば，これを確定現象として取り扱ってもよい．しかし，その不確定さが相対的に大きい場合には，これを不確定現象として取り扱う必要がある．機械の運動や，電気回路の動作は，確定現象として解析するのが普通である．しかし，これらの実験をする場合，実験精度を上げていくと，測定値のばらつきが問題になる．これらのばらつきは，不確定的な実験誤差として処理されなければならない．さいころをふるような現象は，もはや確定現象としての取扱いは不可能に近いであろう．

このように現実の現象には，確定現象と不確定現象の間に，はっきりした境界があるわけではなく，ある現象を確定現象と見るか，不確定現象と見るかは，その現象の性質と取扱いの目的によって決まるものである．しかし，工学においては，非常に複雑な問題や，未知の要素の多い問題に対処しなければならぬ場合が少なくないので，不確定現象としての取扱いが重要になる．次にこのような例を二三掲げておこう．

(1)　電話の呼び　たとえば，ある電話局で毎日一定時間内にかけられる通話(呼び(call))の数を調べたとする．その結果は毎日の観測ごとに異なった値が得られるであろう．各加入者はそれぞれの個人的な事情により電話をかける(呼びを発生する)．しかし観測者は，すべての加入者の事情を知ることもでき

なければ，またそれをコントロールすることもできない．したがって，観測者として，できるだけ一定の条件の下で観測しても，毎日の観測値はばらついたものになる．電話局を設計するときには，このような不確定な呼びを対象にしなければならない．

すべての加入者が同時に電話をかけることも，可能性としては，あり得る．しかし，このようなことはめったに起こることではない．従ってこれに備えて大きな設備を用意することは不経済といわなければならない．一方，設備を節約して小さいものにしておくと，しばしば"お話し中"が起こってサービスが不十分になる．それでは，"実際上"満足できる程度のサービスを行なうにはどれほどの設備を必要とするであろうか．このような問題は，工学ではしばしばぶつかるものである．

(2) 品質検査 工場で物を生産している場合，その製品が所要の特性を満足しているかどうか検査する必要がある．しかしその検査が非常に複雑で時間がかかるとか，あるいは検査によって品物が破壊されるような時には，全製品に対して検査することは不可能であり，製品の中から適当数を無作為に抜き出して検査をする抜取り検査が行なわれる．その検査結果は，やはりばらついたものになるであろう．我々はこの少数の検査結果から，製品全体についての特性を推定し，もしそれが不満足なものであれば，生産過程になんらかの処置を施さなければならない．

(3) 気体 物理学の例をあげる．気体は数多くの気体分子の集りで，これらが互いに衝突を繰り返しつつ運動をしている．分子の運動自身は，力学の法則に従い確定的なものであるから，ある時刻における，すべての分子の位置と運動量は，初期条件により完全に確定するものである．しかし実際には分子の数が極めて多く，その運動を厳密に解析することは困難であり，またすべての分子の位置と運動量を知ることも不可能である．気体の圧力，温度，比熱等の物質としての巨視的な状態あるいは性質は，結局この多数の分子の運動の平均的な状態の結果として観測されるもので，これら不確定な分子運動と巨視的性質の間の関係を導くのが統計力学または熱力学といわれる分野である．

上に掲げた例は物理学，工学の分野で存在する不確定現象のほんの一例に過ぎない．しかし工学においては，このような不確定な現象を対象にし，しかも

限られた範囲の不確実なデータをもとにして，この現象になんらかの推論を下し，ある決断をしなければならない場合が非常に多い．だがこれらの問題から気のつくことは，観測の結果が確定しないといっても，それが全く無法則なものではなく，多数の観測をすれば，その結果にある種の規則性があることである．たとえば，電話の呼びも，その数は日により変動するものであっても，毎日観測していれば，ある平均値のまわりに集まっていることがわかる．したがってすべての加入者が同時に呼びを発することはほとんど起こらないこともわかる．また，抜取り検査にしても，検査見本の数をある程度増せば，見本の不良品率と製品全体の不良品率が大きく食い違うことはめったに起こらない．気体にしても，個々の分子の運動は不確定であっても，分子全体としての運動にはある法則性が存在する．この種の規則性は統計的規則性と呼ばれるもので，不確定現象が理論的考察の対象となり，物理学，工学の対象になるための根拠である．統計的規則性については改めて次節で述べるが，このような規則性が成立し，理論的取扱いのできる不確定現象を，**確率統計現象**と呼ぶことにする．本講義の目的は，このような確率統計現象の基本的な性質と，それを取り扱うための手法を解説しようとするものである．

1.2　確率の概念

再びさいころをふる例を考える．我々はまず“さいころをふってその目を読む”という操作を行なわなければならない．これは，いわば実験を行なうことである．前節の例で，ある時間の電話の呼びの数を数えること，製品の中から無作為に何個かの見本を抜き取って検査すること，ある時間の分子の位置と運動量を調べることは，いずれもこの実験である．このような実験，すなわちある確率統計現象についてある結果を得るのに必要な操作を**試行**(trial)，**実験**(experiment)，あるいは**観察**(observation)等と呼ぶ．これは必ずしも現実に行なえる実験である必要はないが，少なくとも思考の上では，理想化された一定の条件の下で何度でも繰り返すことができるものでなければならない．

このような試行によって我々はある結果を得る．試行によって得られる結果を**事象**(event)と呼ぶ．“1の目が出る”こと，“6の目が出る”こと等はそれぞれ一つの事象であるが，“偶数の目が出る”というのも事象である．後者の事

象は，“2の目が出る”か，または“4の目が出る”か，または“6の目が出る”という三つの事象に分解できる．このようにいくつかの事象を“または”でつなぐ形に分解できる事象を**複合事象**(compound event)といい，それに対し，もはやこれ以上分解できない事象を**単純事象**(simple event)という．さいころの場合には“1の目が出る”,“2の目が出る”,…,“6の目が出る”の六つの事象が単純事象である．

そこで我々は上述の試行を同一条件の下でN回行なったものとし，その結果Aなる事象が$n(A)$回起こったとしよう．たとえば，N回さいころをふる実験を行ない，“1の目の出た”回数を$n(1)$回とする．この時$n(A)/N$を，Aの事象の起こる**相対度数**(relative frequency)という．さいころの実験では試行回数Nを大きくしていくと，1の目の出る相対度数$n(1)/N$は1/6という一定値に近づくことが，経験的に知られている．一般に相対度数が，試行回数が増すと共に一定値に収束すること，すなわち

$$\frac{n(A)}{N} \to \mathbf{P}(A), \quad N \to \infty \tag{1.1}$$

の成り立つことを狭い意味で**統計的規則性**(statistical regularity)があるという．統計的規則性がある場合には我々は素朴に“事象Aの起こる**確率**は$\mathbf{P}(A)$である”といっている．これが我々が**直観的**に持っている確率の概念である．

次に“偶数の目が出る”という事象をAとし，その相対度数を調べて見る．N回の試行においてAが実現した回数$n(A)$は明らかに

$$n(A) = n(2) + n(4) + n(6) \tag{1.2}$$

であるから，相対度数は

$$\frac{n(A)}{N} = \frac{n(2)}{N} + \frac{n(4)}{N} + \frac{n(6)}{N} \tag{1.3}$$

となる．ここで統計的規則性が成り立つとすれば，

$$\frac{n(2)}{N} \to \mathbf{P}(2), \quad \frac{n(4)}{N} \to \mathbf{P}(4), \quad \frac{n(6)}{N} \to \mathbf{P}(6)$$

のように，相対度数はそれぞれある一定値$\mathbf{P}(2)$, $\mathbf{P}(4)$, $\mathbf{P}(6)$に収束するから，

$$\frac{n(A)}{N} \to \mathbf{P}(2) + \mathbf{P}(4) + \mathbf{P}(6) \tag{1.4}$$

すなわち，$n(A)/N$ の収束値としての確率 $\mathbf{P}(A)$ は

$$\mathbf{P}(A)=\mathbf{P}(2)+\mathbf{P}(4)+\mathbf{P}(6) \tag{1.5}$$

となる．

上に述べたような統計的規則性は，現実の世界に存在する多くの現象に成立しており，あるいは少なくともこれを仮定した上で，はじめて現象が理論的な解析の対象になるものであることは前節で述べた通りである．そしてこの際の理論的な取扱いの手段を与えるものが**確率論**(probability theory)といわれる数学である．そこで次に数学としての確率論の考え方に簡単に触れておこう．

我々は現実の世界の観察から経験的に統計的規則性というものを知っており，相対度数の収束値として，素朴な確率の概念を持っている．したがってこれから出発して確率論を組み立てることは極めて自然であり，事実，確率論の起源はこのような自然の観察から生まれたものである．しかし一方，厳密な数学的立場からいえば，純粋な形式論理であるべき数学の出発点に，統計的規則性というわば経験的なものを取り入れるのは，いろいろな点で困難が生ずる．数学としての確率論は，一口にいえば，比較的簡単な事象の確率から，複雑な事象の確率を計算するとか，あるいは現象に関連した統計的パラメーターの値を計算するための手段を与えればよいのである．したがって確率というものの定義とか意味とかを論議することは止めて，すべての起こり得る事象に確率という数値が割り当てられているものとし，それが従うべき規則を約束しておくに止め，これを土台として確率論という数学が組み立てられる．しかし，このような立場から作られた確率論が，現実の世界の現象の取扱いに応用できるためには，その確率が，我々の直観的に持っている確率の概念と矛盾しないものでなければならない．したがって確率論で約束する確率の規則(公理)はそのように定めなければならない．

さいころをふった時に，1の目の出る確率を1/6とすることが，はたして妥当なことかどうかは，数学としての確率論の考察の対象からは除外されている．これは確率論をさいころに応用する時に，我々が経験的あるいはその他の考慮から決定することである．そして各々の目の出る確率がすべて1/6と定められた上で，これから偶数の目の出る確率，あるいは2個のさいころをふった時，(1, 1)の組合せの出る確率等の計算の手段を与えるものが確率論である．ただ

し，現実の問題に適用できるためには，それぞれの目に割り当てた確率の和が1であるとか，その他必要最小限の規則を，直観と矛盾した計算結果が出ないように定めておく必要がある．次節以下にこのような確率論の基本的な概念を述べよう．

1.3 標本空間と事象

ある試行において，その結果として起こりうる単純事象のすべての集合を考え，これを**標本空間**(sample space)といい，それぞれの単純事象のことを**標本点**(sample point)と呼ぶ．さいころの例において，1, 2, … の目の出る事象を，それぞれ (1), (2), … と書けば，これらはそれぞれ標本点であり，(1), (2), …, (6) の全体 {(1), (2), …, (6)} が標本空間である．任意の事象は，これら標本点のいくつかの集り(部分集合)で表わせる．たとえば“偶数の目の出る”事象は {(2), (4), (6)} の三つの標本点の集りで代表される．

次に事象について二三の言葉を定義しておく．

定義 標本点を一つも含まない事象 A を**空事象**といい，

$$A=\phi \tag{1.6}$$

と表わす．

空事象とは絶対起こり得ない事象のことである．

定義 事象 A に含まれる標本点以外の，すべての標本点から成る事象を事象 A の**補事象**(complementary event)といい，A' で表わす．

A' は“A 以外のことが起こる”あるいは“A は起こらない”ことを意味し，さいころの例で $(1)'$ とは 2, 3, …, 6 のいずれかの目が出ることである．標本空間全体を S で表わすと，

$$S'=\phi \tag{1.7}$$

すなわち 1, 2, …, 6 のどの目も出ないことは起こり得ないということを表わす．

定義 事象 A および事象 B に対し，“A または B，あるいはその両方が起こる”という事象を，事象 A および B の**和事象**(union)といい，$A\cup B$ と書く．

これを拡張して事象 $A, B, C, \cdots$ に対して，“$A, B, C, \cdots$ のうち少なくともどれか一つは起こる”という事象を $A\cup B\cup C\cup\cdots$ と書く．

前の例で，偶数の目の出る事象は，事象 (2), (4), (6) の和事象で，(2)∪(4)∪(6) と表わせる．

定義　“事象 A と事象 B が同時に起こる”事象を，事象 A と事象 B の**共通事象**(intersection)または**積事象**(product)と呼び $A\cap B$ で表わす．

これを拡張して，事象 $A, B, C, \cdots$ の共通事象を $A\cap B\cap C\cap\cdots$ と表わす．

たとえば“1, 2, 3 のいずれかの目が出る”ことを事象 A とし，“偶数の目が出る”ことを事象 B とすれば，

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

であるから，

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

である．

事象 A と事象 B が

$$A\cap B = \phi \tag{1.8}$$

の関係にあるとき，A と B は互いに**排反**(exclusive)な事象であるといい，A と B は同時には起こり得ないことを示す．たとえば，1 の目が出ることと，偶数の目が出ることは同時に起こり得ないから

$$\{(1)\}\cap\{(2),\ (4),\ (6)\} = \phi$$

である．

定義　事象 A に属する標本点がすべて事象 B に含まれている時，これを $A\subset B$ あるいは $B\supset A$ と書き，“A が起こることは必ず B が起こることである”を意味する．

$A\subset B$ の場合，B には属するが A には属さない事象，すなわち $A'\cap B$ を $B-A$ と書き，B と A の**差事象**(difference)という．

たとえば，

$$A = \{(2),\ (4)\}$$
$$B = \{(2),\ (4),\ (6)\}$$

とすれば，$A\subset B$ であり

$$B-A=\{(6)\}$$

である．この記号を使えば A の補事象は

$$A'=S-A \tag{1.9}$$

と書ける．これらの関係を図 1.1 に示す．

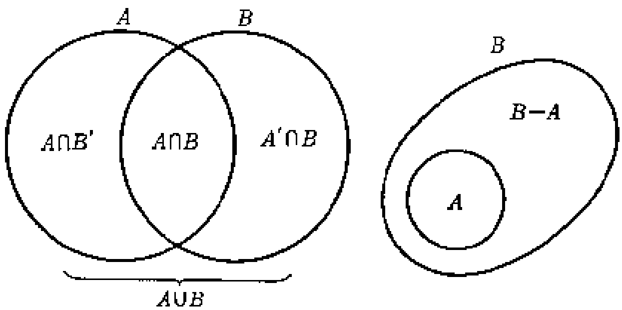


図 1.1 補事象，和事象，積事象，差事象の関係

以上述べたように，ある実験(試行)の結果生ずる事象を表わすのに，標本点および標本空間の概念を導入すると，任意の標本点の組はある一つの事象を表わす．標本点のすべての組合せを考えれば，起こりうるすべての事象を表わすことができる．またその組合せ方を拡張し，一つも標本点を含まない空事象(絶対起こり得ない事象)，および全部の標本点を含む事象(必ず起こる事象)も事象の中に含めて考えることにする．

なお，標本空間は無限の標本点を含む場合にも拡張できる．とくに試行の結果が，連続的な値をとり得る実数値で与えられる場合には，標本空間は連続した直線で代表される．たとえば，ラジオ受信機から発生する雑音の観測を考える．スイッチを入れて後ある時間たった瞬間の電圧を測定するものとすれば，雑音は確率統計的な現象であるから，その測定値はやはり観測ごとに異なった値になる．さいころの実験の場合には，その結果は 1 から 6 までの 6 個の数値のいずれかで与えられるのに対し，この場合の電圧値は，連続したすべての実数値をとり得る．いま，原点Oを通る 1 本の直線を考え，観測値 x に対して，直線上に原点Oから距離 x (x が正ならばOから右方向へ，負ならば左方向へ距離 $|x|$ を測ることにする)の点 a を決めると，すべての観測値と直線の上の 1 点とが 1 対 1 に対応する．したがって，この直線上の 1 点は，電圧のとり得る一つの値を表わすことができる．すなわちこの直線上の任意の点は標本点であ

り，直線全体が標本空間であると見ることができる．この直線上の任意の2点a, bで区切られる線分は，電圧値が$\overrightarrow{\mathrm{Oa}}$と$\overrightarrow{\mathrm{Ob}}$（Oからaまたはbに向かう方向によって符号を決める）の間であるという事象を表わす．

1.4 確率の公理

1.2節で述べたように確率論では，確率を，すべての事象にあらかじめ割り当てられた単なる数値と考えてしまうが，その規則すなわち公理を約束する際に，これが我々が直観的に持っている概念と矛盾しないように決めなければならない．ここですべての事象というのは，前節で述べた通り，あらゆる標本点の組合せを意味し，事象Aがあれば事象A'が存在し，事象Aと事象Bがあれば$A\cup B$, $A\cap B$の事象もこの中に含まれている．また空事象と標本空間全体も一つの事象と考えておく．

公理1　ある試行の結果として起こり得るすべての事象を考えた時，いかなる事象Aに対しても，必ず**確率**(probability)と呼ぶ一つの負でない実数値$\mathbf{P}(A)$が割り当てられている．

公理2　標本空間全体すなわち必ず起こる事象の確率を1とする．

必ず起こる事象の相対度数は常に1であるから，公理2は我々の直観と合致している．

公理3　事象Aと事象Bが互いに排反な事象であれば，

$$\mathbf{P}(A\cup B)=\mathbf{P}(A)+\mathbf{P}(B) \tag{1.10}$$

1.2節において，相対度数にこれと同じ形の(*1.3*)が成立し，相対度数の収束値にも同様(*1.5*)が成立すべきことを知った．したがって公理3はこの性質を確率に要求したものである．

公理3を繰り返し適用すれば次の系が得られる．

系　$A_1, A_2, \cdots, A_k$をすべて互いに排反な事象であるとすれば

$$\mathbf{P}(A_1\cup A_2\cup\cdots\cup A_k)=\mathbf{P}(A_1)+\mathbf{P}(A_2)+\cdots+\mathbf{P}(A_k) \tag{1.11}$$

ある事象Aとその補事象A'は明らかに排反である（“事象Aが起こる”ことと，“事象Aが起こらない”ことが同時に実現することはない）．また事象Aとその補事象A'の和は標本空間全体であり，したがって必ず起こる事象であるから，公理2と公理3から