

確率論 (2014 年度版)

稲垣敏之 : 3B413 (シス情研究科長室)
inagaki.toshiyuki.gb@u.tsukuba.ac.jp
<http://www.css.risk.tsukuba.ac.jp>

Chapter 1 確率

- 1.1 確率の定義について
- 1.2 標本空間と事象
- 1.3 確率
- 1.4 条件つき確率
- 1.5 独立事象
- 1.6 ベイズの定理

付録 1 確率空間

付録 2 確率の連続定理

Chapter 2 確率変数と分布関数

- 2.1 例題
- 2.2 確率変数
- 2.3 分布関数
- 2.4 離散型確率変数
- 2.5 連続型確率変数

Chapter 3 期待値

- 3.1 期待値
- 3.2 確率変数の関数の期待値

Chapter 4 確率変数の独立性

- 4.1 結合確率分布
- 4.2 確率変数の独立性
- 4.3 モーメント母関数
- 4.4 特性関数

Chapter 5 大数の法則と中心極限定理

- 5.1 大数の法則
- 5.2 中心極限定理

Chapter 1 確率

1.1 確率の定義について

どのようにして「確率」を定義するかについては、いくつかの流儀がある。本節では、それらのうち、「数学的確率」と「統計的確率」について考えてみよう。

(1) 数学的確率

ある試行において、同程度の確からしきで起こることが期待される場合の数を N ，そのうち、あることがら A が起こる場合の数を r とするとき、 r / N を「 A が起こる確率」という。

【前提】 1. N は有限確定

2. 試行の結果として起こり得る各場合は、「同程度の確からしき」を持つ

(例1) 「正しく作られたサイコロ」を振ったときに「1の目が出る確率は $1/6$ 」である。

(例2) 遠い宇宙から隕石が落ちてくるらしい。地球上のどこに落ちてても不思議ではなく、どこに落ちるかまったく見当がつかないという。この隕石が海に落ちる確率はどれくらいか、見積もることはできるだろうか？

(例3) 成田からフランクフルトへ向かう直行便に乗ることになった。航空機がハイジャックされることを恐れる人にとっては、自分が乗る便がハイジャック「される」、「されない」の2通りの場合がある。だからといって、「これから乗るフランクフルト行き直行便がハイジャックされる確率は $1/2$ である」といってよいだろうか？

(2) 統計的確率

n 回の試行に対して A の起こる相対度数 r/n を考える。 n を十分大きくしていくとき、 r/n がほぼ一定値 p に近づくなれば、 p を「 A が起こる確率」という。

【前提】 同一条件の下で、何度でも反復して「 A が起こるか否か」を試すことが許される

(問題) 1. どれくらいの n であれば「十分大きい」といえるのだろうか？

2. $n \rightarrow \infty$ のとき、 $r/n \rightarrow p$ となる保証はあるのだろうか？

(これについては、5.1「大数の法則」で詳述する)

(例4) 先行車が急減速したときにドライバーの対応が遅れると、衝突事故が発生する。そのような事故を防止するための切り札として、有能なエンジニアである第5代お茶の水博士は、ドライバーによる対応が検出されないときは、コンピュータが周辺の状況に応じて、ブレーキ操作とステアリング操作を巧みに組み合わせて先行車への衝突を回避する「先進衝突回避システム」を開発した。

博士は、システムが正しく稼動する確率（衝突回避確率とでもよんでおこう）を求めてこのシステムの有用性を示し、世の中に送り出したいと考えている。

さて、博士は、衝突回避確率を統計的確率として算出してよいだろうか？すなわち、衝突回避確率を求めようとすると、実環境のなかで実車を走らせ、先行車のドライバーが好きなきに好きなきように突然急ブレーキをかけ、その度に後続車に搭載された先進衝突回避システムが正しく稼動するかどうかを調べてみるといった実験を延々と繰り返す必要があるのだが、これは人道的に許されることだろうか？

上に述べたことから分かるように、数学的確率、統計的確率のいずれにも、何らかの問題点がある。これらの問題点を回避するひとつの方法は、「確率」が持つべき基本的性質を抽象化して公理を定め、その公理を満たすものを「確率」として定義する方法である。これが公理的確率論 (axiomatic probability theory) とよばれるものである。これはまた、測度論 (measure theory) に立脚することから、測度論的確率論ともよばれる。本講義では、公理的確率論を学んでいくことにしよう。なお、本講義では測度論 (ルベーグ積分論) の知識は前提にしないが、測度論の考え方は時折述べる。

(注1) 測度論を仮定した確率論には良書が多いが、測度論の知識を仮定せず高度な内容を分かりやすく説明している本は多くない。そのなかで、Sheldon M. Ross の著作は、教科書として、あるいは自習用参考書として秀逸である。版を重ねていることから分かるように、定評がある。例題もユニークで面白いものが多い。図書館で手にとってみることを強く勧めたい。

Sheldon M. Ross: A First Course in Probability, 9th revised Ed., Pearson, 2013.

Sheldon M. Ross: Introduction to Probability Models, 11th Ed., Academic Press, 2014.

(注2) 公理的確率論の体系は、コルモゴロフによる「確率論の基礎概念」(1933年刊)によって明らかにされた。ちくま学芸文庫の1冊として入手可能である。ただし、「基礎」といっても「易しい」という意味ではない。

(注3) 余裕のある人は、いつかルベーグ積分を学ぶことを勧める。リーマン積分(1年次の解析学で習ったもの)よりはるかにエレガントである。志賀浩二著「ルベーグ積分30講」(朝倉書店)は、説明が親切であり分かりやすい。古い本であるが、お薦めの1冊である。

1.2 標本空間と事象

試行 明確に定義された実験や観察の単一の実行

- (1) 起こりうる結果がどのようなものでありうるかは既知
- (2) 実際にどの結果が起こるのかは事前には未知

(例1) 人類ではじめて木星に到達したコロンブス 15 世は、地球風にいえば「洞窟」のように見える地形を見つけた。彼は、この中に入ってみようか、止めたほうがよいか、考えあぐねている。人が意思決定に迷ったときに、確率論は大きな手助けをしてあげることがあるが、事前の知識がまったくない状況では、確率論は無力である。実際、「洞窟の中で未知の生命体に襲われる確率」や、「洞窟の中で財宝を発見する確率」などを考えることはできない。

標本空間 Ω ある試行で起こりうる結果(ω)からなる集合 ($\omega \in \Omega$)

(例2) コイン投げ $\Omega = \{H, T\}$

2 枚のコイン $\Omega = \{HH, HT, TH, TT\}$

車の寿命 $\Omega = [0, \infty)$

「標本空間 Ω を何にするか」は自分で決めるものであり、基本となる物理実験と一意的に対応しているとは限らない。そのことは、つぎの例からも知ることができる。

(例3) サイコロ投げのとき、つぎのような Ω が対応しうる。

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

$\Omega = \{\text{偶}, \text{奇}\}$

$\Omega = \{(i, x, y) : i = 1, \dots, 6, a \leq x \leq b, c \leq y \leq d\}$

(x, y) はサイコロの落ちた位置を表す。

事象 Ω の部分集合 (厳密な事象の定義は後述する)

事象 E が起こる \Leftrightarrow 試行の結果 $\omega (\in \Omega)$ が $\omega \in E$ を満たす

(例4) サイコロ投げを考えよう. ただし, $\Omega = \{1, 2, 3, 4, 5, 6\}$ とする. このとき, サイコロを投げたときに偶数の目が出る (E) とは, サイコロ投げによって得られた結果 (サイコロの目) $\omega (\in \Omega)$ が, $\omega \in E = \{2, 4, 6\}$ を満たすことをいう.

問 1. H が出るまでコインを投げ続け, H が出た時点でやめる. このとき, 標本空間 Ω はどのように表すことができるか.

問 2. $n (< \infty)$ ケの要素から成る Ω があるとき, Ω の部分集合の総数はいくつか.

(補足) Ω の部分集合全体を Ω のベキ集合とよび, 2^Ω と書く. なお, 空集合は, あらゆる集合の部分集合である. また, 任意の集合は, それ自身の部分集合であるとみなす.

事象の演算 $E \subset \Omega, F \subset \Omega$

- | | | |
|---------|-------------------------|-------------------------|
| (1) 和事象 | $E \cup F$ | E, F のうち少なくとも一方が起こる |
| (2) 積事象 | $E \cap F$ | E, F がともに起こる |
| (3) 差事象 | $E - F$ | E が起こり, かつ F が起こらない |
| (4) 全事象 | Ω | 確実に起こる |
| (5) 空事象 | ϕ | 起こりえない |
| (6) 余事象 | $E^c \equiv \Omega - E$ | E が起こらない |

(例5) 「確率論」でどのような成績 (評点) が得られるか, その確率的な情報も含めて議論をしたい. 情報科学類では「履修放棄 (X) は評点には用いてはならない」とされていることを反映させると, $\Omega = \{A^+, A, B, C, D\}$ となる. このとき, 「確率論の単位が取れる」という事象 E は, $E = \{A^+, A, B, C\}$ と表すことができる. 「評点 B 以下は良い成績とはみなさない」という人であれば, 「良い成績で単位が取れる」という事象 F は, $F = \{A^+, A\}$ と表すことになる.

問3. 「確率論」と「力学」の2科目を受講した人が、それらの科目でどのような成績（評点）が得られるかに関心を持っているとする.

(1) 上記の議論の際に必要な標本空間 Ω はどのようなものか.

(2) 「確率論の単位を取得する」ことを E , 「力学の単位を取得する」ことを F と書くとき, 「確率論と力学の少なくとも一方の単位は取れる」という事象はどのように表現できるか.

(3) 「確率論は単位が取れるが, 力学の単位取得には失敗する」という事象はどのように表現すればよいか.

Def. $E \cap F = \phi$ のとき, E と F はたがいに排反であるという.

分かりやすくいえば, E と F はたがいに排反であるとは, E と F は同時には起こり得ないことである. 1.5 節で述べる「 E と F が独立である」ことと, ここで述べた「 E と F がたがいに排反である」ことを混同してはならない. 「 E と F が独立である」ときは, 「 E が起こるか否か」は, 「 F が起こるか否か」にまったく影響を及ぼさない. つまり, 「 F は起こる」かもしれないが, 「 F は起こらない」かもしれない. しかし, 「 E と F がたがいに排反である」ときは, 「 E が起こった」とすると, もはや「 F は起こりえない」.

(例6) サイコロの1から6までの目は, たがいに排反である.

(例7) E さんと F さんのうち, くじ引きでどちらか一人だけが100万円を獲得することができるとき, 「 E さんが100万円を獲得する」と「 F さんが100万円を獲得する」ことは, たがいに排反である.

さて、p. 5 で述べた事象演算には、 $\cup, \cap, ^c, -$ の 4 種類があったが、実は、 \cup と c 、または \cap と c があれば、他の 2 つの演算は表現することができる。すなわち、4 種類の演算記号のうち 2 つをうまく選んでおけば、残りの 2 つの記号は不要である。この事実は、後に付録 1 で重要な役割を果たす。

問 4. \cup と c を用いて、 $E \cap F$ および $E - F$ を表せ。

問 5. \cap と c を用いて、 $E \cup F$ および $E - F$ を表せ。

問 6. 1 枚のコインと 1 枚のサイコロを投げる。

(1) 標本空間 Ω を表せ。

(2) つぎの各事象を表せ。

$$A \triangleq \{\text{コインは } H \text{ が出、サイコロは偶数の目が出る}\}$$

$$B \triangleq \{\text{サイコロの目は素数}\} \quad (\text{注}) 1 \text{ は素数ではない.}$$

$$C \triangleq \{\text{コインは } T \text{ が出、サイコロは奇数の目が出る}\}$$

(3) $A \cup B, B \cap C$ を求めよ。

(4) A, B, C のうち B のみが起こることは、どのような事象として表現できるか。

(5) A, B, C のうち、たがいに排反なものはどれか。

1.3 確率

実数値集合関数 P がつぎの $1^\circ \sim 3^\circ$ を満たすとき, $P(E)$ を事象 E の確率とよぶ.

$$1^\circ \quad P(E) \geq 0 \quad (E \subset \Omega)$$

$$2^\circ \quad P(\Omega) = 1$$

$$3^\circ \quad E_i \cap E_j = \phi \quad (i \neq j) \text{ なる事象列 } E_1, E_2, \dots \text{ に対し,}$$

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i) \quad (\text{完全加法性という})$$

「確率」が満たすべき必要最小限の性質を書き上げてその全体 (ここでは $1^\circ \sim 3^\circ$) を公理系とし, それらだけを用いて定理を証明していくアプローチを公理的確率論とよぶ. 以下に示すように, わずか3つの式 ($1^\circ \sim 3^\circ$) だけで強力な理論体系を構築していくことができる点は, 驚嘆すべきことではないだろうか. ぜひ味わってほしいところである.

公理 $1^\circ \sim 3^\circ$ から導かれる性質

$$(P1) \quad P(\phi) = 0$$

$$(P2) \quad E_i \cap E_j = \phi \quad (i \neq j) \text{ なる事象列 } E_1, E_2, \dots, E_n \text{ に対し,}$$

$$P\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n P(E_i) \quad (\text{有限加法性という})$$

$$(P3) \quad P(E^c) = 1 - P(E)$$

$$(P4) \quad 0 \leq P(E) \leq 1$$

$$(P5) \quad E \subset F \text{ ならば, } P(E) \leq P(F) \quad (\text{単調性という})$$

$$(P6) \quad P(E - F) = P(E) - P(E \cap F)$$

$$(P7) \quad P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

$$(P8) \quad P(E \cup F) \leq P(E) + P(F)$$

あるいは一般化して,
$$P\left(\bigcup_{i=1}^n E_i\right) \leq \sum_{i=1}^n P(E_i) \quad (\text{劣加法性という})$$

$$(P9) \quad P\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} P(E_i) \quad (\text{劣加法性の一般形})$$

(注) 「(P8) で $n \rightarrow \infty$ とすればよいではないか」と考えるのはダメ (その方法は大丈夫そうに見えるかもしれないが, 極限操作の対象がまったく異なるという誤りを犯している)

$$(P10) \quad P\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n P(E_i) - \sum_{i < j} P(E_i \cap E_j) + \sum_{i < j < k} P(E_i \cap E_j \cap E_k) + \dots + (-1)^{n-1} P(E_1 \cap \dots \cap E_n) \quad (\text{複合確率の公式という})$$

問1. 正しく作られたサイコロ (fair die) は, i の目 ($i = 1, \dots, 6$) が出る事象を E_i とするとき, $P(E_i) = 1/6$ でモデル化できる. では, unfair die の例はどのようなものか, 作ってみよ.

完全加法性と有限加法性 (この議論はやや高度であるので, はじめは読み飛ばしてもよい)

(1) 完全加法的ならば有限加法的.

(公理 3° において, $E_{n+1} = E_{n+2} = \dots = \phi$ とせよ)

(2) 有限加法性 (P2) で $n \rightarrow \infty$ としても完全加法性 (公理 3°) は得られない.

(理由) 任意の n について $\bigcup_{i=1}^n E_i \subset \Omega$ ゆえ,

$$P\left(\bigcup_{i=1}^n E_i\right) \leq P(\Omega) = 1$$

が成立する. ここで, $E_i \cap E_j = \phi$ ($i \neq j$) を仮定しているから, 任意の n について

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

が成立するが, これは数列 $\left\{\sum_{i=1}^n P(E_i)\right\}$ が有界な単調数列であることを示している.

さて, 有界な単調数列は収束するから, $\lim_{n \rightarrow \infty} \sum_{i=1}^n P(E_i)$ が存在する. この極限値を

$\sum_{i=1}^{\infty} P(E_i)$ と書くことにすると, (P2) から主張できるのは次式である.

$$(a) \lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n E_i\right) = \lim_{n \rightarrow \infty} \sum_{i=1}^n P(E_i) = \sum_{i=1}^{\infty} P(E_i)$$

一方, 公理 3° が主張していることは, つぎのとおりである.

$$(b) P\left(\lim_{n \rightarrow \infty} \bigcup_{i=1}^n E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

(a) と (b) を見比べると, (b) が成立するためには.

$$(c) \lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n E_i\right) = P\left(\lim_{n \rightarrow \infty} \bigcup_{i=1}^n E_i\right) \quad (\text{連続性という})$$

の成立が求められることがわかる.

しかし、一般には、作用素の順序交換は無条件には許されないことから、(c)の成立は保証されない。すなわち、有限加法性において「 $n \rightarrow \infty$ 」としても、完全加法性を得ることはできない。有限加法性と完全加法性の間にある「ギャップ」を埋めるものが (c) である。

ちなみに、(c) は $P(\bullet)$ の「連続性」とよばれている。これを簡潔に表すと、つぎのようになる。

$$(\text{完全加法性}) = (\text{有限加法性}) + (\text{連続性})$$

なぜ (c) が「連続性」という名前と呼ばれるかについては、1年次の解析学で学んだ「連続関数」の定義を思い出してみれば理解できるのではないだろうか。

(補足) 本節冒頭に示した公理的確率の定義をフォーマルにしたものを「付録1」に示す。また、確率の連続性に関連した性質を「付録2」に示す。いずれもやや理屈っぽく見える内容であるが、いくつかの概念について、「なぜ、そのように定義される必要があったのか」を理解しておくことは重要である。余力のある人は、一度挑戦してみることを勧める。

1.4 条件つき確率

赤のサイコロと黒のサイコロ（いずれも正しく作られたサイコロであるとする）を投げたとき、それらの目の和が6になれば勝つゲームがある。赤と黒のサイコロの目をそれぞれ d_1, d_2 とするとき、

$P\{d_1 + d_2 = 6\} = \frac{5}{36}$ を得る。すなわち、このゲームに勝つ確率は $5/36$ である。

さて、このゲームに参加している人が赤と黒のサイコロを投げたとき、2つともテーブルから落ちて、カーテンの陰に隠れてしまった。そのとき、係りの人が、「赤のサイコロの目は4です」というのが聞こえた。「 $d_1 = 4$ 」という情報が得られたわけである。

では、この「赤のサイコロの目は4です」という情報が得られたことによって、ゲームに勝つ確率は大きくなったのだろうか、小さくなったのだろうか、それとも変わらないのだろうか。

(解)「 $d_1 = 4$ 」という情報のもとで、 $d_1 + d_2 = 6$ となる確率はどのように変わるのかを調べてみよう。
 $d_1 = 4$ のとき、可能な目の組合せ (d_1, d_2) は、つぎの6通り。

$(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)$

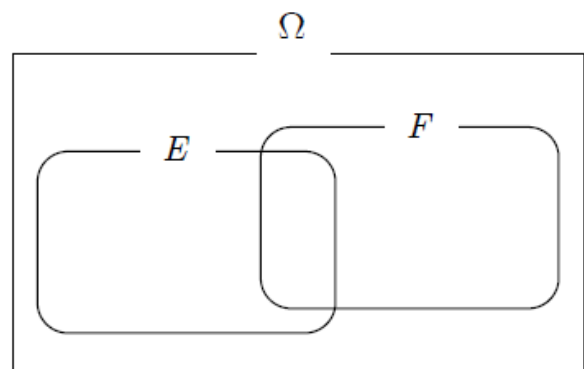
これらは、それぞれ等確率 $1/6$ で起こる。このうち、 $d_1 + d_2 = 6$ を満たすものは、 $(4,2)$ のみである。よって、求める確率は $1/6$ 。すなわち、「 $d_1 = 4$ 」という情報が与えられたとき、その情報が与えられる前と比べると、 $d_1 + d_2 = 6$ となる確率は大きくなったといえることができる。

上の問題で求めた確率は、

$$E = \{(d_1, d_2) : d_1 + d_2 = 6\}$$

$$F = \{(d_1, d_2) : d_1 = 4\}$$

とするとき $P(E|F)$ と書いて、「事象 F が起こったという条件の下での事象 E の条件つき確率」とよぶ。なお、 $P(E|F)$ を考えるときの標本空間は、 $\Omega = \{(1,1), (1,2), \dots, (6,6)\}$ ではなく、 F であることに注意されたい。



さて、上の例では、

$$P(E|F) = \frac{Card(E \cap F)}{Card(F)} = \frac{\frac{Card(E \cap F)}{Card(\Omega)}}{\frac{Card(F)}{Card(\Omega)}} = \frac{P(E \cap F)}{P(F)}$$

となるが、 $Card(\Omega) = \infty$ のときには、この議論は適切ではない。ちなみに、 $Card(A)$ は集合 A の濃度 (A が有限集合の場合は A に含まれる要素数) を表す記法である。

そこで、 $Card(\Omega) = \infty$ の場合も扱うことができ、 $Card(\Omega) < \infty$ の場合には上の議論と同一の結果が得られるようにしたい。そこで、条件つき確率をつぎのように定義することにしよう。

Def. $P(E|F) \triangleq \frac{P(E \cap F)}{P(F)}$ (ただし $P(F) \neq 0$ とする)

問 1. ある家庭には 2 人の子供がいる。男女の出生率は等しいと仮定して、以下の問に答えよ。

(1) 2 人とも男の子である確率はいくらか。

(2) 1 人は男の子であると分かったとする。2 人とも男の子である確率はいくらか。

(3) 年長の子が男の子であると分かっている場合なら、2 人とも男の子である確率はいくらか。

問 2. 区別できる 2 個の正しく作られたサイコロを投げる。各々の目を d_i ($i = 1, 2$) とする。 $d_1 + d_2 = 6$ であるとわかった状況で、いずれか一方のサイコロの目が 2 である条件つき確率はいくらか。

問 3. 10 枚のカードがある. その各々には 1 から 10 までの数のいずれかひとつが書き込まれている. なお, 異なるカードには異なる数が書き込まれているものとする. さて, これら 10 枚のカードが入った箱の中から, A 君はランダムに 1 枚のカードを引いた. もし, 「いま引いたカードに書かれている数は 5 以上です」といわれたとき, そのカードに書かれた数が 10 である条件つき確率はいくらか.

問 4. 赤球 7 個と白球 5 個が入った箱がある. この中から, 非復元抽出により, 順次 2 個を取り出すものとする. このとき, 2 個とも赤球である確率はいくらか.

条件つき確率に関する性質

$$(P1) \quad P(E \cap F) = P(E|F)P(F) = P(F|E)P(E) \quad (\text{乗法公式という})$$

$$(P2) \quad P(E_1 \cap \dots \cap E_n) = P(E_1)P(E_2|E_1)\dots P(E_n|E_1 \cap E_2 \cap \dots \cap E_{n-1})$$

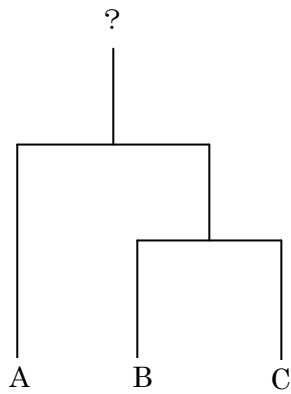
$$(P3) \quad \bigcup_{i=1}^n E_i = \Omega, \quad E_i \cap E_j = \phi \quad (i \neq j) \text{ のとき}$$

$$P(F) = \sum_{i=1}^n P(F|E_i)P(E_i) \quad (\text{全確率の公式という})$$

$$(P4) \quad \bigcup_{i=1}^n E_i = \Omega, \quad E_i \cap E_j = \phi \quad (i \neq j) \text{ のとき, } \sum_{i=1}^n P(E_i|F) = 1$$

問 5. 条件つき確率は確率の公理を満たす (すなわち, 「条件つき確率は確率である」) ことを証明せよ.

問 6.



左図のトーナメントで、引き分けはないとする.

A vs B において A の勝つ確率は $1/2$

A vs C において A の勝つ確率は $2/3$

B vs C において B の勝つ確率は $1/2$

このとき、A が優勝する確率はいくらか.

問 7. あるパーティでの余興の話である. パーティの主催者であるサリエリ氏は、シルクハットを差し出し、3人のビジネスマン (A さん, B さん, C さん) に対して、「この中にご自分の名刺を入れてください」といった. サリエリ氏は、シルクハットの中の名刺を十分にかき混ぜてから、3人のビジネスマンに目隠しをさせ、「A さん, B さん, C さんの順で、シルクハットの中から名刺を1枚ずつ引いてください」といった. 3人のうち少なくともひとりが自分の名刺を引き当てる確率はいくらか.

1.5 独立事象

Def. 1 事象 E と F が独立 $\Leftrightarrow P(E \cap F) = P(E)P(F)$

あるいは、つぎのように定義しても同義である。

事象 E と F が独立 $\Leftrightarrow P(E|F) = P(E)$ または $P(F|E) = P(F)$

「独立性」の概念を直感的に把握するための例をあげておこう。

(例 1) 「確率論」と「力学」は一般には異なる学問体系に属する科目である。両科目に対する学生の嗜好や興味を持つ(持たない)理由も同一ではないのがふつうであろう。したがって、「確率論」の試験で優秀な成績をおさめることが、「力学」の試験で優秀な成績をおさめることを意味することはない。すなわち、「確率論」でよい成績を修めることと、「力学」でよい成績を修めることは、独立事象であると考えてもよい。

(例 2) 「確率論」を担当している I 教授が茶目っ気を起こして、今年の「確率論」の試験問題は、今年度の「力学」の講義内容に即した応用問題の文脈で作成してみようと考えたなら、話は別である。そのような「確率論」の試験問題は、「力学」が得意な学生にとっては願ってもないものであろう。試験後に TWINS にアクセスして成績を調べたとき、まだ「確率論」の成績が入力されていないものの「力学」の評点が「A」となっていることが確認できれば、やがて「確率論」の評点欄に「A」が記載される可能性は高いと期待できよう。すなわち、「確率論」で好成績を修めることと、「力学」で好成績を修めることが独立事象であるとは考えにくい。

コンピュータシステム、原子力プラント、航空機をはじめ、社会にはさまざまなシステムがある。これらのシステムが故障することなく所定の機能を発揮することができる確率を信頼性 (reliability) とよぶ。システムの信頼性評価において、「独立性」の概念は大きな役割を果たしている。

(例 3) 大型航空機には、複数のエンジンが搭載されている。ボーイング 747-400 やエアバス A380 なら 4 基、ボーイング 787 やエアバス A320 なら 2 基といった具合である。飛行中にこれらのエンジンのひとつが故障しても、ふつうは他のエンジンが故障しやすくなるということはない。実際、双発機がひとつだけのエンジンで数時間の飛行をすることは可能である。システム信頼性評価では、基本的にはエンジン故障は「独立事象」として扱うことが多い。

ただし、「エンジンの故障は、つねにたがいに独立であるか」といえば、実はそうではない。

(例4) ボーイング 747-400 の4基のエンジンがすべて飛行中に停止したことがあった。上空に巻き上げられた火山灰を吸い込んだためである。ひとつの原因(たとえば火山灰)のために、複数のエンジンが同時に故障したのである。これを、「共通原因故障」(common cause failure)という。共通原因(common cause)のもとでは、エンジンの故障は「独立事象」ではなくなる。

(例5) 飛行前のエンジンを点検したエンジニアが手順を間違えていたため、飛行機のエンジンが上空で不調になり、やがてすべてが停止してしまったという事例も知られている。これも共通原因故障の典型である。あるエンジンが不調に陥ったとき、「同一の手順で『整備』された他のエンジンは、やがて停止する可能性が高い」と考えて対応することは、機長の重要な責務である。

システム信頼性評価において共通原因故障を考えておくことは、むずかしいが重要なことである。

問1. 2個の正しく作られたサイコロの目を d_1, d_2 とする。 $d_1 + d_2 = 6$ である事象を E 、 $d_1 = 4$ である事象を F とするとき、 E と F は独立か?

問2. 箱の中に5個の白球と3個の赤球が入っており、各々の球は等確率で取り出されるものとする。

$E_1 \triangleq$ 2個の球を引き続いて取り出したとき、最初の球は白球

$E_2 \triangleq$ 2個目の球は白球

とするとき、つぎの2通りのケースにおいて、 E_1 と E_2 は独立か。

(1) 1個を取り出した後、その球は箱に戻さないとき(非復元抽出)

(2) 1個を取り出した後、その球は箱に戻すとき(復元抽出)

問 3. 事象 A と B が独立であるとき, (1) A^c と B , (2) A と B^c , (3) A^c と B^c も独立であることを証明せよ.

3つ以上の事象の独立性の定義は, やや複雑である. たとえば, 3つの事象 E, F, G を考えたとき, 2つずつ取り出すと独立である (Def.1 で述べた式が成立) ように見えても, 3つ全体に対しては独立ではない (Def.1 を3つの事象に拡張した式が成立しない) といったことが起こり得るからである. そのような例をあげておこう.

(例6) 箱の中に1, 2, 3, 4のいずれかの番号が記されたボールが入っている. ただし, 同じ番号のボールはない. 各々のボールは等確率で取り出されるものとする. このとき,

$$E = \{\text{ボール1またはボール2が取り出される}\},$$

$$F = \{\text{ボール1またはボール3が取り出される}\},$$

$$G = \{\text{ボール1またはボール4が取り出される}\}$$

とすると,

$$P(E \cap F) = P(E)P(F) = \frac{1}{4},$$

$$P(E \cap G) = P(E)P(G) = \frac{1}{4},$$

$$P(F \cap G) = P(F)P(G) = \frac{1}{4}$$

であり, 2つずつの事象は独立であるといえる.

しかし,

$$\frac{1}{4} = P(E \cap F \cap G) \neq P(E)P(F)P(G) = \frac{1}{8}$$

となることから, 3つの事象は全体としては独立ではないことがわかる.

このことから, 3つ以上の事象に対する独立性は, つぎのように定義する必要がある.

Def. 2 n 個の事象 E_1, E_2, \dots, E_n が独立 \Leftrightarrow

これらの任意の部分集合 $E_{i_1}, E_{i_2}, \dots, E_{i_r}$ ($r \leq n$) について次式が成立

$$P(E_{i_1} \cap E_{i_2} \cap \dots \cap E_{i_r}) = P(E_{i_1}) \dots P(E_{i_r})$$

問 4. 2 個の正しく作られたサイコロを投げる. $\Omega = \{HH, HT, TH, TT\}$. つぎの 3 事象は独立か.

$$A \triangleq \{\text{第1のコインは}H\}, \quad B \triangleq \{\text{第2のコインは}H\}, \quad C \triangleq \{\text{1ヶのコインだけ}H\}$$

1.6 ベイズの定理

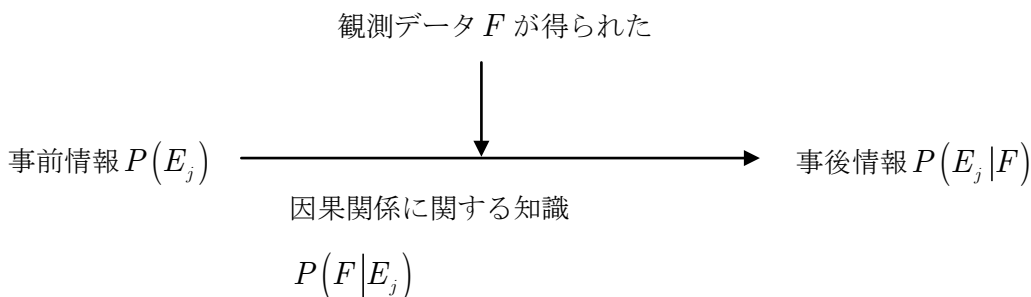
外国へ出かけるのが大好きなコロンブス君は、ある会社の海外旅行キャンペーンに応募した。50人しか応募できないが、応募できさえすれば、そのうち30人が当選するという。発表当日、一足先に発表を見てきたというマゼラン君から、「君は当選していたよ。おめでとう」といわれた。ただ、マゼラン君は茶目っ気が旺盛であり、3回のうち2回はウソをいって人をからかう。ほんとうにコロンブス君が当選している可能性はどれくらいなのだろうか。

実は、人はこのような推論が得意ではない。「50人のうち30人が当選する」のだから、当選の可能性のほうが高い。しかし、必ずしも信頼できない人から、「当選していましたよ」といわれたとき、実際に当選している可能性がどれほど影響を受けるかを推測することは、それほど簡単ではない。直感が働かないときに助けてくれる「強い味方」が、つぎに述べるベイズの定理である。

Thm. (Bayes) $\bigcup_{i=1}^n E_i = \Omega$, $E_i \cap E_j = \phi$ ($i \neq j$) とするとき、次式が成立する。

$$P(E_j|F) = \frac{P(F|E_j)P(E_j)}{\sum_{i=1}^n P(F|E_i)P(E_i)}$$

(解釈) E_1, \dots, E_n のうち、いずれかひとつだけが起こりうる。その各々の起こりやすさを $P(E_1), \dots, P(E_n)$ と考えていたところ、観測データ F を得た。この観測データ F のもとで、 E_1, \dots, E_n の起こりやすさをどのように更新すればよいだろうか。因みに、因果関係に関する知識 $P(F|E_j)$ は持ち合わせているとする。



問 1. 上の **Thm** (ベイズの定理またはベイズの公式という) を証明せよ.

(注) ベイズの定理は, 記憶してはならない. つねに導出できるようにしておくことが大切.

問 2. 本節冒頭に述べた海外旅行キャンペーンに応募したコロンブス君が, 発表当日, 一足先に発表を見てきたというマゼラン君から, 「君は当選していたよ. おめでとう」といわれた. ただし, マゼラン君は茶目っ気が旺盛であり, 3回のうち2回はウソをいって人をからかうものとする.

(1) このキャンペーンの応募者はコロンブス君を含めて 50 名であり, そのうちの 30 人が当選すると分かっているとき, コロンブス君が当選している可能性はいくらか.

(2) このキャンペーンの応募者はコロンブス君を含めて 50 名であり, そのうちの 40 人が当選すると分かっているときなら, コロンブス君が当選している可能性はどれくらい高くなるか.

問3. m 個の選択肢があるクイズがあり，博識で名高いポアロ君がそれに挑戦しようとしている．クイズに参加する前からポアロ君は出題領域に関する豊富な知識を持っており，はじめからクイズの正解を知っている（と思われる）確率を p とする．一方，クイズの出題領域に関する知識の持ち合わせがなく，はじめからクイズの正解を知っているとは思えない確率を $1-p$ とする．出題されたクイズに対して正解を知らないとき，ポアロ君は， m 個の選択肢の中から 1 個をランダムに選んで「ファイナルアンサー」とすることになっている．

(1) 出題されたクイズに対してポアロ君が選んだ答は正解であり，ポアロ君は莫大な賞金を手に入れた．このとき，もともとポアロ君がクイズに対する正解を知っていた（と思われる）確率 p^* を求めよ．

(2) ポアロ君がクイズに対する正解を知っている（と思われる）事前確率 p を 0.2 としてみよう．ポアロ君がクイズに正解したことを知ったとき，そのクイズが 3 択問題 ($m=3$) であるか，10 択問題 ($m=10$) であるかによって，事後確率 p^* はどのように違うかを調べてみよ．

問4. A, B, Cは同じ製品を作る機械であり, それぞれ製品全体の20%, 30%, 50%を生産している.
製品の不良率は, Aが5%, Bが5%, Cが3%であることが経験的に知られている.
今, 製品全体からサンプルを1個取り出して調べてみると, 不良品であった.

(1) このサンプルが機械Aで作られた確率を求めよ.

(2) このサンプルが機械Bで作られた確率を求めよ.

(3) このサンプルが機械Cで作られた確率を求めよ.

問5. ある製品が「欠陥品である」と, 消費者からクレームがついた. 本当に欠陥があるときにクレームがつくのは確率0.45 (我慢強い人もいる), 欠陥はないのにクレームがつくのは確率0.05 (怖い人もいる) とする. 過去のデータから, この製品の不良率は10%であることが分かっていると, 消費者がクレームをつけた製品がほんとうに欠陥品である確率はいくらか.

問6. 箱が2つある. 箱1には, 白球2個と赤球7個が入っている. 一方, 箱2には, 白球5個と赤球6個が入っている. さて, 正しく作られたコインを投げ, 表が出たら箱1から, 裏が出たら箱2から, ランダムに球を1個取り出すものとする. さて, 取り出された球は白球であった. このとき, コインに表が出ていた条件つき確率はいくらか.

問7. 怖い病気が蔓延する気配を見せている. 今, この病気にかかっているのは全人口の0.5%である.

(1) この病気への感染の有無を99%の正確さで検出できる血液検査が開発されたというので早速検査を受けてみたところ, 「陽性である」といわれた. 本当にこの病気にかかっている確率はいくらだと考えればよいか.

(2) 検査の正確さが95%であれば, 上の議論と様相が大きく変わることを確認せよ.

(注) 100点満点の試験での99点と95点なら, どちらも「良い成績」なのだが…

付録1 確率空間

n 変数実数値関数 f に対して, つぎのような記法がある.

$$f: x \rightarrow f(x) \quad \text{あるいは} \quad f: R^n \rightarrow R^1$$

1.3 で定義した実数値集合関数 P についても同様に書くと,

$$P: E \rightarrow P(E) \quad \text{あるいは} \quad f: \mathcal{F} \rightarrow R^1$$

両表現の対応から明らかなように, $x \in R^n$ に対して $E \in \mathcal{F}$ である. すなわち, \mathcal{F} は標本空間 Ω の部分集合 E (事象) を要素とする集合族である. \mathcal{F} にはどのような性質を要求すればよいだろうか.

E, F が事象であれば, $E \cup F, E \cap F, E^c, E - F$ も事象であってほしい. すなわち,

- ◆ 事象 E_1, \dots, E_n に有限回の集合演算を施したものを事象として考えられるようにしたい.

さらに一般化して,

- ◆ 事象 E_1, E_2, \dots に可算無限回の集合演算を施したものを事象として考えられるようにしたい.

そこで, 集合族 \mathcal{F} に, つぎの3つの条件を要求する.

- 1° $\phi \in \mathcal{F}$ (または $\Omega \in \mathcal{F}$) (空でない集合族を考えようということ)
- 2° $E \in \mathcal{F}$ ならば, $E^c \in \mathcal{F}$
- 3° $E_1, E_2, \dots \in \mathcal{F}$ ならば, $\bigcup_{i=1}^{\infty} E_i \in \mathcal{F}$

上の性質を持つ \mathcal{F} を, Ω 上の σ -加法族 (σ -集合体, σ -代数) とよぶ.

(注1) 上述の2°, 3°において2種の演算 $\cup, ^c$ のみを扱った理由については, 1.2を参照のこと.

(注2) 以後, E が事象であることを, $E \in \mathcal{F}$ と表すことがある.

(注3) σ -加法族は, 有限加法族 ($E_1, \dots, E_n \in \mathcal{F}$ ならば, $\bigcup_{i=1}^n E_i \in \mathcal{F}$ であるとするもの).

σ -加法族 \mathcal{F} を記述に入れて確率の公理1°~3°を厳密に書くと、つぎのようになる。

標本空間 Ω の部分集合から成るひとつの σ -加法族 \mathcal{F} の上で定義される実数値関数 P が以下の1°~3°を満たすとき、 $P(E) (E \in \mathcal{F})$ を「 E の確率」とよぶ。

$$1^\circ \quad P(E) \geq 0 \quad (E \in \mathcal{F})$$

$$2^\circ \quad P(\Omega) = 1$$

$$3^\circ \quad E_i \in \mathcal{F} \quad (i = 1, 2, \dots) \text{ かつ } E_i \cap E_j = \phi \quad (i \neq j) \text{ のとき, } P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$

こうして定まる (Ω, \mathcal{F}, P) を「確率空間 (probability space)」とよぶ。

(補注1) 測度論では、 (Ω, \mathcal{F}) を可測空間 (measurable space), $E (E \in \mathcal{F})$ を可測集合 (measurable set), (Ω, \mathcal{F}, P) を測度空間 (measure space) とよぶ。ただし、一般には測度 P に有界性は要求しない。換言すれば、確率論は $0 \leq P(E) \leq 1$ という条件を満たす有界測度 P を扱う測度論である。ただし、確率論独自の概念もあるため、「確率論は測度論に包含されている」とまではいえない。

(補注2) 厳密に言えば「標本空間 Ω の部分集合ならば、どのようなものでも事象である」とはいえない。これは「 Ω の部分集合で可測でないものが存在する」という測度論の結果による。例えば、円周 Ω に対して、 $E_i \cap E_j = \phi (i \neq j)$, $\Omega = \bigcup_{i=1}^{\infty} E_i$ を満たすがいいに合同な集合 E_i の存在が知られているが、 E_i は可測集合ではない。

(理由: E_i が可測集合であるとする、 $1 = P(\Omega) = \sum P(E_i)$ が成り立つ。 E_i は合同であるから、 $P(E_i) = \alpha (\forall i)$ と書くことができるが、 $\alpha > 0$ としても $\alpha = 0$ としても矛盾が生じる。すなわち、 E_i の確率 $P(E_i)$ を考えることはできない。)

そこで、どのようなものを事象とよぶか (どのようなものに確率を考えることができるか) を定めるために導入したのが、 σ -加法族である。

ちなみに、 R^n のすべての区間からできている集合族を A とする。 A より大きい (A を包含する) σ -加法族のうち、最も小さな σ -加法族 A' を考える。 A' に属する集合を、ボレル集合とよぶ。

(補注 3) 公理 1° ~ 3° を見ると、「確率論は、不確実性を扱う唯一無二の理論である」と思えるかもしれない。しかし、「不確実性」には、さまざまなものがある。

例えば、Fuzzy 理論や Dempster-Shafer 理論 (証拠理論 Theory of Evidence ともいう) は、それぞれ確率論では表現しにくい不確実性を扱おうとする理論である。

付録 2 確率の連続定理

Def. 1 事象列 $\{E_i\}$ が単調事象列 $\Leftrightarrow E_1 \subset E_2 \subset \dots$ または $E_1 \supset E_2 \supset \dots$

Def. 2 $E_1 \subset E_2 \subset \dots$ のとき, $\lim_{n \rightarrow \infty} E_n \triangleq \bigcup_{n=1}^{\infty} E_n$

$E_1 \supset E_2 \supset \dots$ のとき, $\lim_{n \rightarrow \infty} E_n \triangleq \bigcap_{n=1}^{\infty} E_n$

Thm. 1 (確率の連続定理) $\{E_i\}$ が単調事象列ならば, $P\left(\lim_{n \rightarrow \infty} E_n\right) = \lim_{n \rightarrow \infty} P(E_n)$.

(証明) $E_i \in \mathcal{F}$ のとき, $\bigcup_{n=1}^{\infty} E_n \in \mathcal{F}$, $\bigcap_{n=1}^{\infty} E_n \in \mathcal{F}$. すなわち, $\{E_i\}$ が単調事象列のとき, $\lim E_n \in \mathcal{F}$.

よって, $P\left(\lim E_n\right)$ は存在する.

以下の 2 通りの各々について, 定理を証明する.

(1) $E_1 \subset E_2 \subset \dots$ のとき,

$$\begin{aligned} P\left(\lim E_n\right) &= P\left(\bigcup_{n=1}^{\infty} E_n\right) \\ &= P\left(E_1 \cup (E_2 - E_1) \cup \dots\right) \\ &= P(E_1) + P(E_2 - E_1) + \dots \\ &= \lim \left\{P(E_1) + P(E_2 - E_1) + \dots + P(E_n - E_{n-1})\right\} \\ &= \lim P\left(E_1 \cup (E_2 - E_1) \cup \dots \cup (E_n - E_{n-1})\right) \\ &= \lim P(E_n) \end{aligned}$$

(2) $E_1 \supset E_2 \supset \dots$ のとき,

$E_1^c \subset E_2^c \subset \dots$ に対して, (1) より $P(\lim E_n^c) = \lim P(E_n^c)$ が成立することを用いる.

(注 1) 左辺の極限操作は事象列に対して行われるが、右辺の極限操作は実数列に対して行われることに注意せよ.

(注 2) 上の等式は、 $P(\cdot)$ と \lim の順序交換が可能であることを示す.

(注 3) 解析学で学んだ関数 f の連続性が

$$f(\lim a_n) = \lim f(a_n)$$

と書けたことを想起すれば、Thm. 1 が連続定理とよばれることが理解できるであろう.

(注 4) 「完全加法性」 = 「有限加法性」 + 「連続性」

Def. 2 では単調事象列の極限を「天下り式」に定義したが、その正当性は以下の議論で明らかになる.

Def. 3 上限 $\sup A_n \triangleq \bigcup_{n=1}^{\infty} A_n$

下限 $\inf A_n \triangleq \bigcap_{n=1}^{\infty} A_n$

上極限 $\overline{\lim} A_n \triangleq \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ (可算無限個の A_n に属する点の集合)

下極限 $\underline{\lim} A_n \triangleq \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$ (ある番号 n' から先の全ての A_n に属する点の集合)

Thm. 2 $\overline{\lim} A_n \supset \underline{\lim} A_n$

(証明) $x \in \underline{\lim} A_n$ とすると、適当な番号 n' に対して $x \in \bigcap_{k=n'}^{\infty} A_k$.

すなわち、 $k \geq n'$ なる任意の k に対して $x \in A_k$.

よって、任意の n に対して、 $x \in \bigcup_{k=n}^{\infty} A_k$.

このことは、 $x \in \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \overline{\lim} A_n$ と等価.

Def. 4 $\overline{\lim}A_n = \underline{\lim}A_n$ のとき, このことを $\lim A_n$ 書き, 事象列 $\{A_n\}$ の極限とよぶ.

Thm. 3 $A_1 \subset A_2 \subset \dots$ のとき, $\lim A_n = \bigcup_{n=1}^{\infty} A_n$

(証明) $\overline{\lim}A_n = \underline{\lim}A_n = \bigcup_{n=1}^{\infty} A_n$ を示す.

$$\bigcup_{k=n}^{\infty} A_k = \bigcup_{k=1}^{\infty} A_k \text{ ゆえ,}$$

$$\overline{\lim}A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \bigcup_{n=1}^{\infty} A_n$$

$$\text{一方, } \bigcap_{k=n}^{\infty} A_k = A_n \text{ ゆえ,}$$

$$\underline{\lim}A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k = \bigcup_{n=1}^{\infty} A_n$$

Thm. 4 $A_1 \supset A_2 \supset \dots$ のとき, $\lim A_n = \bigcap_{n=1}^{\infty} A_n$

(証明) Thm. 3 と同様. 各自試みよ.

Thm. 3 と Thm. 4 を用いれば, Thm. 1 は極限が存在する任意の事象列 $\{A_n\}$ に対して成立することがわかる. すなわち, $P(\lim_{n \rightarrow \infty} A_n) = \lim_{n \rightarrow \infty} P(A_n)$.

(証明) $\lim A_n$ が存在したとする. $\lim A_n = \overline{\lim}A_n = \underline{\lim}A_n$ であったから, $\lim A_n \triangleq \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$

と書くことができる. ここで, $B_n \triangleq \bigcap_{k=n}^{\infty} A_k$ とおくと, $B_n \in \mathcal{F}$ かつ $B_1 \subset B_2 \subset \dots$ であることから,

$$\lim B_n = \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k = \lim A_n$$

単調事象列 $\{B_n\}$ に対して $P(\lim B_n) = \lim P(B_n)$ が成立することに注意すると,

$$P(\lim A_n) = P(\lim B_n) = \lim P(B_n) \leq \lim P(A_n) \quad (*)$$

一方, $\lim A_n = \overline{\lim} A_n = \underline{\lim} A_n$ であつたから, $\lim A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$ とも書くことができる.

ここで, $C_n \triangleq \bigcup_{k=n}^{\infty} A_k$ とおくと, $C_n \in \mathcal{F}$ かつ $C_1 \supset C_2 \supset \dots$ であることから,

$$\lim C_n = \bigcap_{n=1}^{\infty} C_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k = \lim A_n$$

単調事象列 $\{C_n\}$ に対して $P(\lim C_n) = \lim P(C_n)$ が成立することに注意すると,

$$P(\lim A_n) = P(\lim C_n) = \lim P(C_n) \geq \lim P(A_n) \quad (**)$$

(*) と (**) より, $P(\lim A_n) = \lim P(A_n)$ を得る.