確率論入門I

服部哲弥

START

１．ガイダンス

フェルマーとパスカル（17 世紀半ば，確率論の記録上最初の議論）

問 2–1：『2 人が何回戦か行うカードのゲームで，一方はあと 2 回，もう一方はあと 3 回，先に勝ったほうが賞金を総取りするところで，中止になった．賞金をどう分配すべきか？各回五分五分引き分け無し』 ◇

確率論

現実には存在しないかもしれない未来や過去についての大きさ（量）矛盾のない，ただ一つの数値を与える考え方

注意： 確立ではなく確率（減点の対象）

数学としての確率論　－　量を測ること

考え方としての確率論（数学）

数学：　正しい理解＋良い理解

数学（確率論）の立場：実現しなかった「何か」（集合）の「大きさ」（測度）が持つべき性質を定義して，結論を研究

公理的確率論（コルモゴロフ）

可算和と補集合に関して閉じている集合族（集合の集合）を定義域とする，σ加法性を持つ非負値（集合）関数を測度といい，全体集合の測度が 1 である測度を確率（測度）と言う

確率を決めて計算するよりも，どの確率も共通に持つ性質は?という問

期待される資質：

集合の演算に習熟していること

抽象的な定義を身近に感じること

社会の中の確率論　－　バラツキの理解

ＡＴＭや窓口が複数ある時,（空港や銀行などで見かけるように）一列に並んで空いた窓口を利用する一列並びと,（スーパーで多く見かけるように）窓口ごとに列ができる並列並びがある．一列並びが並列並びより良いのはどういう点か？

問６−１：一列並びが並列並びより良いのはどういう点か？　◇

待ち時間の**期待値**は（ほぼ）等しい
→　平均だけで考えるとどちらでも良いはず
バラツキ（**分散**）が０でないことに意味（教科書「統計と確率の基礎 第３版」２章）
決定論と**分布**を持つ法則の決定的な違い

正しい理解・良い理解　－　高度な結果（への道筋）

正しい・良い理解　→　遙か彼方の複雑な問題にも応用できる

株価の変化を観測期間全体で 1 本の時間の関数と見て，関数が袋から無作為に取り出された結果が現実世界と見る
関数の集合上の確率という見方（確率過程）

数学としての確率論は集合の上の測度と抽象化して定義するので，一般的に成り立つ定理は複雑な対象でも成り立つ
種々の応用

春学期の内容の予定

フェルマーとパスカル	初等確率論
集合の大きさを測る	測度としての確率
モンティ・ホール問題	条件付き確率
試験問題の総数	確率変数と期待値
なぜ一列並びか	分散
２項定理再び	独立な確率変数
いつか表が出る	σ加法性
強運の持ち主	ポワッソン分布
バスはなぜ来ないか	
保険が成り立つ理由	大数の法則
確率変数列の収束	

秋学期の内容の予定

秋学期は，統計学など実用上の確率の計算，の基礎．特に，
非可算無限集合の上の確率測度，特に，積分を用いた確率

- ランダムウォーク
- 非可算無限集合の上の確率測度
- 実数上の確率測度
- ガウス積分と正規分布
- 母関数と特性関数
- 同時分布
- 独立性と特性関数
- 多次元正規分布
- 中心極限定理
- 特性関数の収束
- その他

教科書等

教科書：　服部哲弥「統計と確率の基礎 第３版」，学術図書，2014年

- １年間で本書前半が読める程度が目標
- 教科書にある問題は解答を教科書参照ですませる
- 試験範囲（の一部）

ウェブ ：　服部哲弥／ホーム／「統計と確率の基礎」／補足と訂正
確率論入門　（この講義スライドと過去問）

レポート：　keio.jp ３回の予定（紙掲示無し，講義とkeio.jp要確認）

成績・単位：　学期末定期試験

- 教科書の問題や説明（講義等で明示する範囲）＋レポート
- ウェブの過去問（略解付き）を参照

- 服部担当の水曜3,4限のはどちらを聴いても同じです（試験も同一）

（特殊な登録（例：春３限秋４限）は念のため学生部に問い合わせて下さい）

履修登録の判断の参考

◯　**数学の選択科目**なので，数学を勉強したい諸君に勧めます．

・　春学期Iは，数学技術的には多くは要求しませんが，思っていた**確率**と違ってみえるかもしれません．抽象化して論理の構造を考える**もののみかた**は**合う・合わない**があります．

・　**大学での成長が単位**の前提（高校水準では不可）です．

ー　数学が得意な人：　行きがけの駄賃（単位）に，どうぞ．

ー　数学にご縁が薄かったけどこれ（単位）を機に勉強する気の人：　歓迎します（ただし，下記理由で一定の採点水準を維持しますので，不運な巻き添えを食わないよう，過去問やレポートの勉強など，成績を取る勉強もよろしく．）

ー　最近の春学期は統計的に考えにくい受講者数で，**不合格＋不受験が３割以上**あり，出回ってることが（統計的に）推測される**裏情報は誤り**です．不適切な受講や受験は試験監督や部屋確保で多方面に迷惑をかけます．数学科目の１つとして，誤答はばっさり０点，の方針を強化します．

◯　秋学期IIは他科目・実用も意識し，**計算技術**も難しくなります

２．フェルマーとパスカル

- 数学としての確率論（コルモゴロフの公理）
- 集合関数としての確率
- 確率の簡単な計算（初等確率論）

確率空間を決めて計算する，という考え方から，
どの確率も共通に持つ性質は？という問へ（公理的確率論）

教科書の付録Ａ §1 の一部を今回と次回の２回で説明します

初等確率論

文章題（現象）を数学の問題として定式化（確率空間の設定）

問 2–1：『2 人が何回戦か行うカードのゲームで，一方はあと 2 回，もう一方はあと 3 回，先に勝ったほうが賞金を総取りするところで，中止になった．賞金をどう分配すべきか？各回五分五分引き分け無し』 ◇

確率空間：コルモゴロフの公理系（20 世紀前半）　　5

公理主義的確率論

全体集合 Ω：考察の対象全体（母集団）
まず，Ω が有限集合の場合の簡単版

P が Ω 上の確率（確率測度）であるとは

0．Ω の部分集合 $(A \subset \Omega)$ に対して実数値 $\mathrm{P}[\,A\,]$ を与える関数

$\mathcal{F}$：Ω の部分集合を全て（簡単版）集めた集合（集合族）

0'．**集合関数**： $\mathrm{P}:\ \mathcal{F} \to \mathbb{R}$ 　　cf. 普通の関数は $f:\ \Omega \to \mathbb{R}$

1．**非負値**： $\mathrm{P}[\,A\,] \geqq 0,\ A \in \mathcal{F}$

2．**加法性**：$A, B \in \mathcal{F}$ が $A \cap B = \emptyset \ \Rightarrow\ \mathrm{P}[\,A \cup B\,] = \mathrm{P}[\,A\,] + \mathrm{P}[\,B\,]$

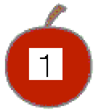
3．全測度 1（確率）：$\mathrm{P}[\,\Omega\,] = 1$

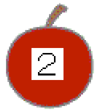
$(\Omega, \mathcal{F}, \mathrm{P})$：**確率空間**

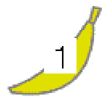
参考：条件 3. を要求しない集合関数を**測度**と呼ぶ．確率は全測度 1 の測度である．
コルモゴロフの公理（20 世紀前半）　→　現代確率論

15

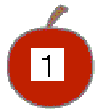
確率空間の例

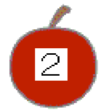
$\Omega = \{$ 

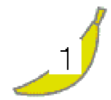


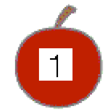


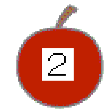
$\}$ （「買い物の総額を表す確率空間」）後述

$\mathcal{F} = \{\{\ \},\ \{$ 

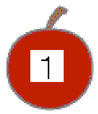
$\},\ \{$ 

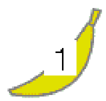
$\},\ \{$ 

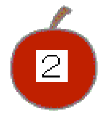
$\},\ \{$ 

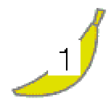


$\},$

$\{$ 

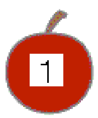


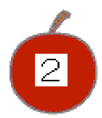
$\},\ \{$ 



$\},\ \Omega\}$

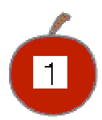
注：**空集合** $\emptyset = \{\ \}$ **と全体集合** Ω **は必ず定義域に加える**：$\mathrm{P}[\,\Omega\,] = 1$ **（確率の定義）.**
$A \cap \emptyset = \emptyset$ **なので加法性から** $\mathrm{P}[\,A \cup \emptyset\,] = \mathrm{P}[\,A\,] + \mathrm{P}[\,\emptyset\,]$.
$A \cup \emptyset = A$ **なので** $\mathrm{P}[\,\emptyset\,] = 0$

$\mathrm{P}[\,\{$ 

$\}\,] = p,\ \mathrm{P}[\,\{$ 

$\}\,] = q.$

問2–2：(1) **Pの定義域** $\mathcal{F}$ **の全ての集合** A **に対して** $\mathrm{P}[\,A\,]$ **を計算せよ.**

(つまり $\mathrm{P}[\,\emptyset\,] = ?,\ \mathrm{P}[\,\{$ 

$\}\,] = ?,\ \ldots,$ **全てに答えよ.)**

(2) **Pが確率であるために,** p **と** q **が満たすべき条件を求めよ.** ◇

最初のヒント：**バナナ1個の値段（確率）は？**

20

集合関数で定義する理由

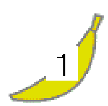
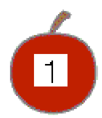
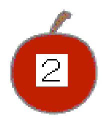
Ω が有限集合ならば，$\Omega \simeq \{1, 2, \dots, N\}$

$i = 1, 2, \dots, N$ に対して $p(i) = \mathrm{P}[\,\{i\}\,]$ 根元事象とおけば，加法性によって，$\mathrm{P}[\,A\,] = \sum_{i \in A} p(i)$ （有限集合では集合関数は不要）

集合関数で定義する理由：0．そもそも集合の大きさを測るのが測度

1．例えば積分で定義される確率 $\mathrm{P}[\,A\,] = \int_A \rho(x)\,dx$ は $p(i)$ の和で書けない．統一的に扱うため（今から慣れるため）

2．定義域をわざと制限した別の確率を考える

$\mathcal{F}_1 = \{\{\ \},$ $\{$  $\},$ $\{$   $\},$ $\Omega\},$ （$\mathrm{E}_1[\,X\,] = \mathrm{E}[\,X \mid \mathcal{F}_1\,]$）

→条件付き確率→ある時に分かっている情報で条件付けてその後の確率や期待値を計算（確率過程，確率連鎖→ファイナンス）

25

集合と要素，和集合と共通部分

集合の**要素** $a \in A,\ b \notin A \quad 1 \in \{1,2\},\ 2 \in \{1,2\},\ 3 \notin \{1,2\}$

和集合 $A \cup B \quad \{1,2\} \cup \{2,3\} = \{1,2,3\}$

共通部分 $A \cap B \quad \{1,2\} \cap \{2,3\} = \{2\}$

$(A \cup B) \cup C = A \cup (B \cup C) \quad (A \cap B) \cap C = A \cap (B \cap C)$

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

証明：$x \in A \cap (B \cup C) \ \Leftrightarrow\ x \in A$ かつ（$x \in B$ または $x \in C$） $\Leftrightarrow$

$x \in A \cap B$ または $x \in A \cap C$ $\square$

空集合 $\emptyset = \{\}$ （加法性の公理で既出，確率測度の定義域に含む）

$A \cap \emptyset = \emptyset \quad A \cup \emptyset = A \quad \{1,2\} \cap \{3\} = \emptyset$ 35

補集合，そして集合算の公式

全体集合：確率論ではΩが使われることが多い（が，決まってない）

補集合 A^c　$\Omega=\{1,2,3\}$, $A=\{1,2\}$のとき $A^c=\{3\}$

$A\cap A^c=\emptyset \quad A\cup A^c=\Omega \quad A\cap\Omega=A \quad A\cup\Omega=\Omega$

証明：$x\in A\cap A^c \Leftrightarrow x\in A$ かつ $x\notin A \Rightarrow x$ は存在しない　$\square$

ド・モルガンの法則　$(A\cap B)^c=A^c\cup B^c \quad (A\cup B)^c=A^c\cap B^c$

証明：$x\in(A\cap B)^c \Leftrightarrow x\notin A\cap B \Leftrightarrow x\notin A$ または $x\notin B \Leftrightarrow x\in A^c\cup B^c$ $\square$

差 $A-B=A\cap B^c$　$\Omega=\{1,2,3\}$の時 $\{1,2\}-\{2,3\}=\{1,2\}\cap\{1\}=\{1\}$

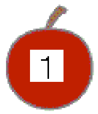
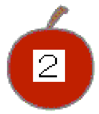
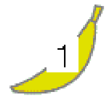
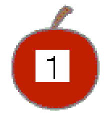
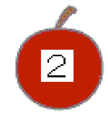
$(A-B)\cup(A\cap B)=A \quad (A-B)\cap(A\cap B)=\emptyset$

問 2–3：今日の講義の公式を用いて上記を証明せよ　$\diamond$

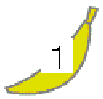
$A-B$ は B と和集合を取って $A\cup B$ になるうち最小の集合

例題解く時間 20 分 45–65

問２の答え

問2–2：$\Omega = \{$    $\}$, $\mathrm{P}[\,\{$  $\}\,] = p$, $\mathrm{P}[\,\{$  $\}\,] = q$.

(1) **Pの定義域$\mathcal{F}$の全ての集合Aに対して**$\mathrm{P}[\,A\,]$**を計算せよ．**

$\mathrm{P}[\,\{$  $\}\,] = 1 - p - q$ **（後略）（根元事象の確率と加法性で定まる）** □

(2) **Pが確率であるために，pとqが満たすべき条件は？**

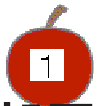
$p \geqq 0,\ q \geqq 0,\ p + q \leqq 1$ **（有限集合では根元事象の確率の非負値性だけ）** □

問2–3：$(A - B) \cup (A \cap B) = A \quad (A - B) \cap (A \cap B) = \emptyset$

$$(A - B) \cup (A \cap B) = (A \cap B^c) \cup (A \cap B)$$
$$= ((A \cap B^c) \cup A) \cap ((A \cap B^c) \cup B) = A \cap ((A \cup B) \cap (B^c \cup B))$$
$$= A \cap ((A \cup B) \cap \Omega) = A \cap (A \cup B) = A$$ **（後略）** □

自習

問2–2の補足

P[{  }] $= p$ は「リンゴ1の値段がp」と同じ意味
（全部買い占める値段を1とデノミ）
P[$\emptyset$] $= 0$ は「何も買わなければただ」
（「測度としての確率」は値段や個数数えと同じということ）

「何も買わない確率が正」という話との関係？
確率空間が全く違う：選ぶ$=1$，選ばない$=0$と対応させて
$\Omega = \{(0,0,0),(1,0,0),(0,1,0),(1,1,0),\ldots,(1,1,1)\}$
「どれも買わない事象」は$\{(0,0,0)\}$で，空集合でないから
「何も買わない確率が正」つまりP[$\{(0,0,0)\}$] > 0 はかまわない
考えたい内容に沿った確率空間を選ぶ必要．
与えられた確率空間の現実的意味（直感）を正確に読み取る必要．

注1．小学校の鶴亀算→中学校の連立一次方程式．鶴と亀の匹数をxとyとおいて，一方の変数を機械的に消せば解ける．鶴の数とその足の数をxとyとしては解けない．
注2．高校教科書の繰り返し試行で出るが，高校には抽象的として確率空間を書かないでごまかした．そのツケ）後述

問2–1の答え

問2–1：『Aはあと2回，Bはあと3回，先に勝ったほうが賞金を総取りするところで，中止になった．賞金をどう分配すべきか？』

4世紀前の解：あと4回やれば決着．全て書き出す．

$$\Omega = \{AAAA, AAAB, AABA, AABB, ABAA, ABAB, ABBA, ABBB, BAAA, BAAB, BABA, BABB, BBAA, BBAB, BBBA, BBBB\}$$

Aが勝つ**事象**（集合）をA勝とおくと，$\mathrm{P}[\,A\text{勝}\,] = \dfrac{11}{16}$.

架空の世界(Ω)での賞金$X(\omega) = 1_{A\text{勝}}(\omega)$;

$$1_C(\omega) = \begin{cases} 1 & \omega \in C \\ 0 & \omega \notin C \end{cases}.$$

（Ω上の関数：確率変数）　（後述）

配分率＝期待値$\mathrm{E}[\,X\,]$（後述）　$= \mathrm{E}[\,1_{A\text{勝}}\,]$

$= 1 \times \mathrm{P}[\,A\text{勝}\,] + 0 \times \mathrm{P}[\,A\text{負}\,] = \mathrm{P}[\,A\text{勝}\,] = \dfrac{11}{16}$. AとBに11 : 5で配分．□

注．勝ちが決まったところでやめて良い：$\Omega = \{AA, ABA, ABBA, ABBB, BAA, BABA, BABB, BBAA, BBAB, BBB\}$

$$\mathrm{P}[\,A\text{勝}\,] = \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{8} + \frac{1}{16} + \frac{1}{16} = \frac{11}{16}.$$

4世紀前：同程度に確からしい根元事象→場合の数で計算（今の高校の確率と同じ）

まとめ

全体集合Ω：考察の対象全体（今日は有限集合），
$\mathcal{F}$：Ωの部分集合を全て（簡単版）集めた集合
空集合$\emptyset = \{\}$も定義域に含む（$\mathrm{P}[\,\emptyset\,] = 0$）
PがΩ上の確率（$(\Omega, \mathcal{F}, \mathrm{P})$が確率空間）であるとは，非負値と加法性（簡単版）を持ち$\mathrm{P}[\,\Omega\,] = 1$を満たす集合関数 $\mathrm{P}:\ \mathcal{F} \to \mathbb{R}$.

集合の記号：要素 $a \in A$, $b \notin A$，和集合$A \cup B$，共通部分$A \cap B$，補集合A^c, 差$A - B = A \cap B^c$.

公式：$A \cap \Omega = A$, $A \cup \Omega = \Omega$ $A \cap \emptyset = \emptyset$, $A \cup \emptyset = A$, $A \cap A^c = \emptyset$, $A \cup A^c = \Omega$, $(A - B) \cup (A \cap B) = A$, $(A - B) \cap (A \cap B) = \emptyset$, 分配法則，ド・モルガンの法則

積み残したこと：
測度は「集合の大きさ」，確率は全測度1の測度
積分で定義される確率（集合関数が自然な理由），無限集合上の確率論
定義域の制限（条件付き確率），確率過程，数理ファイナンス

３．集合の大きさを測る

- 集合の大きさを測る集合関数（測度）

- 測度としての確率

注意：確立ではなく**確率**　　速度ではなく**測度**（減点の対象）

- 集合算と確率の基礎公式

経験的な大数の法則（賭博師，保険）で値が決まるという見方から
定義から値が決まる数学へ（現実に騙されない定式化と強力な算法）

前回に続いて，教科書の付録Ａ §1 の一部を説明します

確率論の二重の抽象性

現場・現象（あやふや，騙される）　→（抽象）→　確かなモノ
1．数学（測度論）として扱うための抽象化（全ての数学に共通）
2．対象の抽象性「存在しなかった過去や未来」

問 3–1：プロ野球日本シリーズ開幕２連勝は 2010 年度までで 29 回，うち 23 回がそのまま先に４勝（日本一）．ニュース解説『連勝は良い流れをもたらし，２勝以上の価値』．解説は正しいか？ ◇

- 深く考えないと間違えやすい　→　数学的取り扱いが役立つ
- 抽象性　→　具体的な計算は見方を変えるアイデアも多い

測度

確率を決めて計算するよりも，**どの確率も共通に**持つ性質は？という問

全体集合 X の全ての部分集合（簡単版）を集めた**集合族** $\mathcal{F}$．

μ が $(X, \mathcal{F})$ 上の**測度**とは，

0．**集合関数**： $\mu: \ \mathcal{F} \to \mathbb{R} \cup \{+\infty\}$ で $\mu(\emptyset) = 0$

1．**非負値**： $\mu(A) \geqq 0, \ A \in \mathcal{F}$

2．**加法性（簡単版）**： $A, B \in \mathcal{F}$ が $A \cap B = \emptyset \ \Rightarrow \ \mu(A \cup B) = \mu(A) + \mu(B)$

$(X, \mathcal{F}, \mu)$：**測度空間**

測度は**集合の大きさを測る量**に共通する性質を抽象化した数学的概念

（20 世紀初め，ルベーグ）（後述）

確率は**全測度**が 1（$\mathrm{P}[\,\Omega\,] = 1$）の測度

　→　**集合の大きさ**を測る量

先に確率の公式

確率の基礎性質

$\mathrm{P}[\,\Omega\,] = 1$，$\mathrm{P}[\,\emptyset\,] = 0$（既出）

補集合の確率：$\mathrm{P}[\,A^c\,] = 1 - \mathrm{P}[\,A\,]$

証明：$A \cup A^c = \Omega,\ A \cap A^c = \emptyset$（既出）だから，

加法性から $\mathrm{P}[\,A\,] + \mathrm{P}[\,A^c\,] = \mathrm{P}[\,A \cup A^c\,] = \mathrm{P}[\,\Omega\,] = 1$ $\square$

A と B に共通部分があっても使える公式を探す

包含 $A \subset B,\ A \supset B$ $\{1\} \subset \{1,2\},\ \{1,2\} \supset \{1\},\ \emptyset \subset A \subset \Omega$

注：等号を含む． $A \subset B$ **かつ** $A \supset B \ \Leftrightarrow\ A = B$

$A \cap B \subset A \subset A \cup B$

証明：$x \in A \cap B$ ならば $x \in A$ なので $A \cap B \subset A$．後半も同様． $\square$

例：$A - B = A \cap B^c \subset A$

$A \supset B \ \Rightarrow\ A \cup B = A$，かつ，$A \cap B = B$

例： $A \cup (A \cap B) = A$ かつ $A \cap (A \cap B) = A \cap B$

確率の非負値性・単調性・劣加法性

単調性：$A \supset B \ \Rightarrow\ \mathsf{P}[\,A\,] \geqq \mathsf{P}[\,B\,]$

証明：$(A-B)\cup(A\cap B)=A,\ (A-B)\cap(A\cap B)=\emptyset$ （問2–3）

だから，確率の加法性と非負値性から

(1) $\mathsf{P}[\,A\,]=\mathsf{P}[\,A-B\,]+\mathsf{P}[\,A\cap B\,]\geqq\mathsf{P}[\,A\cap B\,]$.

$A\supset B\ \Rightarrow\ A\cap B=B$（前回）だから $\mathsf{P}[\,A\cap B\,]=\mathsf{P}[\,B\,]$ □

例：$A\cup B\supset B$（前回）なので $\mathsf{P}[\,A\cup B\,]\geqq\mathsf{P}[\,B\,]$

問2–3で A に $A\cup B$ を代入すると

$(A-B)\cup B=A\cup B,\ (A-B)\cap B=\emptyset$

(2) $\mathsf{P}[\,A\cup B\,]=\mathsf{P}[\,A-B\,]+\mathsf{P}[\,B\,]$ （加法性から）

これと $A-B=A\cap B^c\subset A$（前回）と単調性から

劣加法性：$\mathsf{P}[\,A\cup B\,]\leqq\mathsf{P}[\,A\,]+\mathsf{P}[\,B\,]$

和集合の確率

再掲（前頁）

(1) $\mathsf{P}[\,A\,] = \mathsf{P}[\,A - B\,] + \mathsf{P}[\,A \cap B\,]$

(2) $\mathsf{P}[\,A \cup B\,] = \mathsf{P}[\,A - B\,] + \mathsf{P}[\,B\,]$

$$\mathsf{P}[\,A \cup B\,] = \mathsf{P}[\,A\,] + \mathsf{P}[\,B\,] - \mathsf{P}[\,A \cap B\,]$$

注：加法性 $A \cap B = \emptyset \;\Rightarrow\; \mathsf{P}[\,A \cup B\,] = \mathsf{P}[\,A\,] + \mathsf{P}[\,B\,]$ **だけから証明できた．加法性のほうが特別な場合なので，同値．**

狭い性質を確率の定義にしておけば，ある集合関数が確率測度であることをチェックするのが楽．使うときは，上記公式を有効利用（数学の議論展開の常道）

例：和集合の確率の公式の B **に** $B \cup C$ **を代入してこれまでの公式も駆使すると，**

$\mathsf{P}[\,A \cup B \cup C\,] = \mathsf{P}[\,A\,] + \mathsf{P}[\,B\,] + \mathsf{P}[\,C\,] - \mathsf{P}[\,A \cap B\,] - \mathsf{P}[\,B \cap C\,] - \mathsf{P}[\,C \cap A\,] + \mathsf{P}[\,A \cap B \cap C\,]$

（別解を後述）

測度としての確率

測度:非負値性と加法性(簡単版)を持つ集合関数$\mu: \mathcal{F} \to \mathbb{R} \cup \{+\infty\}$

長さ($X=$直線上の点の集合),面積($X=$平面上の$\cdots$),体積

非負値関数の積分も面積で定義する

個数も測度

測度:集合の大きさを測ることに共通する性質

・集合関数によって定義する必要性:点の長さは0,線分$[0,1]$の長さは1.点をいくら集めても(加法性によれば)0.

・測度論の非凡:$\mu([a,b])=b-a$と非負値性とσ加法性が矛盾せず,唯一のμを定義する(1次元ルベーグ測度)(ルベーグ,20世紀初頭)

・測度としての確率:確率も実現しなかった過去や未来の大きさを測る.何を測っているのか分からなくても矛盾のないただ一つの値を得る(コルモゴロフ,20世紀前半)

・計算結果が(ほぼ)実現する状況もある(大数の法則=保険や賭博経営の根拠)

例題解く時間30分

問3–1の答え

問3–1：プロ野球日本シリーズ開幕２連勝は2010年度まででで29回，うち23回がそのまま先に４勝（日本一）．ニュース解説『連勝は良い流れをもたらし，２勝以上の価値』．解説は正しいか？

五分五分として２勝０敗から始めて先に４勝になる確率

$$= \mathrm{P}[\{AA, ABA, BAA, ABBA, BABA, BBAA, ABBBA, BABBA, BBABA, BBBAA\}]$$

$$= \frac{1}{4} + \frac{2}{8} + \frac{3}{16} + \frac{4}{32} = \frac{26}{32} = 0.81\ldots > 0.79 = \frac{23}{29}.$$

現実は先勝したほうが五分五分より分が悪い． □

注：2011年度終了時点 $\frac{23}{30} = 0.77$．さらに分が悪くなった

まとめ

全体集合 X，X の部分集合を全て集めた集合族（簡単版）$\mathcal{F}$，$(X, \mathcal{F}, \mu)$ が測度空間とは，μ が非負値性と加法性（簡単版）を満たす集合関数
確率は $\mathrm{P}[\ \Omega\ ] = 1$ となる測度

包含関係に関する公式：$\emptyset \subset A \subset \Omega$, $A \subset B$ **かつ** $A \supset B \ \Leftrightarrow\ A = B$, $A \cap B \subset A \subset A \cup B$, $A - B = A \cap B^c \subset A$, $A \supset B \ \Rightarrow\ A \cup B = A$ **かつ** $A \cap B = B$, $A \cup (A \cap B) = A$ **かつ** $A \cap (A \cap B) = A \cap B$.

確率の公式:$\mathrm{P}[\ A^c\ ] = 1 - \mathrm{P}[\ A\ ]$, **単調性, 劣加法性,** $\mathrm{P}[\ A\ ] = \mathrm{P}[\ A - B\ ] + \mathrm{P}[\ A \cap B\ ]$, $\mathrm{P}[\ A \cup B\ ] = \mathrm{P}[\ A - B\ ] + \mathrm{P}[\ B\ ]$, $\mathrm{P}[\ A \cup B\ ] = \mathrm{P}[\ A\ ] + \mathrm{P}[\ B\ ] - \mathrm{P}[\ A \cap B\ ]$.

測度は集合の大きさを測ることに共通する性質
長さ，面積，体積,（積分,）個数，確率は測度（確率は実現しない事の大きさを測る）
集合関数によって定義する必要性
定義は同値な性質のうち見かけ上条件の厳しい性質に取ってある．

積み残し：大数の法則，σ 加法族，σ 加法性，ルベーグ測度の存在と一意性

４．モンティ・ホール問題

- 条件付き確率
- ベイズの定理

今回は関心の維持のためにも前回までよりもややくだけた話をします

レポート１ ＠ keio.jp

パラドックス？

問 4–1：3 つのドアが舞台に登場するテレビの番組がある．1 つだけが当たりで，開けると景品の新車がある．挑戦者がドアを 1 つ選んだ後で，司会者のモンティ・ホールが残り 2 つのうち外れのドアを 1 つ開ける．挑戦者はドアの選択を変更できるが，変更したほうが得か？　◇

素朴な考え 1：現在の選択肢は確率 $1/3$，変更先は確率 $1/2$
素朴な考え 2：元々対等だったはずだから対等なままのはず

条件付き確率

事象（集合）B が起きる条件の下で A が起きる**条件付き確率**：

$$\mathrm{P}[\,A \mid B\,] = \frac{\mathrm{P}[\,A \cap B\,]}{\mathrm{P}[\,B\,]}$$

情報を得て事象の確率がどの程度**確実**になるか，の数理モデル
（　→　**確率過程・確率連鎖**：確率的現象の時間発展）

- 追加情報無し（**先験的確率**）→ A が起きる確率は $\mathrm{P}[\,A\,]$
- B が起きたとき A が起きる確率は $\mathrm{P}[\,A \mid B\,]$

事象 A と B が**独立**とは $\mathrm{P}[\,A \cap B\,] = \mathrm{P}[\,A\,]\,\mathrm{P}[\,B\,]$
（$\mathrm{P}[\,B\,] \neq 0$ のとき $\Leftrightarrow \mathrm{P}[\,A \mid B\,] = \mathrm{P}[\,A\,]$）

条件付き確率が確率であること

$Q[\cdot] = P[\cdot \mid B]$ で集合関数 Q を定義

$$Q[A] + Q[A^c] = \frac{\mathsf{P}[\,A \cap B\,] + \mathsf{P}[\,A^c \cap B\,]}{\mathsf{P}[\,B\,]} = 1$$

P が確率測度ならば，
Q も非負，加法性，全測度 1 を満たすので確率測度．
全体集合 Ω，定義域 $\mathcal{F} = \{\emptyset,\ A,\ A^c,\ \Omega\}$

今後このような確認を省略するが，各自意識するほうが良い．

条件付き確率の例

例：工場Bは不良品率1%工場Cは3% ，同数ずつ製造．不良品を買う事象A，工場B,Cの製品という事象B, C
$C = B^c$, $\mathrm{P}[\,B\,] = \mathrm{P}[\,C\,] = 0.5$, $\mathrm{P}[\,A \mid B\,] = 0.01$, $\mathrm{P}[\,A \mid B^c\,] = 0.03$
$\mathrm{P}[\,A\,] = \mathrm{P}[\,A \cap B\,] + \mathrm{P}[\,A \cap B^c\,] = \mathrm{P}[\,A \mid B\,]\mathrm{P}[\,B\,] + \mathrm{P}[\,A \mid B^c\,]\mathrm{P}[\,B^c\,] = 0.01 \times 0.5 + 0.03 \times 0.5 = 0.02$

$\mathrm{P}[\,A\,]$	0.02
$\mathrm{P}[\,A \mid B\,]$	0.01

$\mathrm{P}[\,A\,]$	0.02
$\mathrm{P}[\,A \mid B^c\,]$	0.03

Bでも$B^c = C$でも，情報が増えたと考える（数値の大小ではなく，原因の情報によって結果の評価が精密になった）

相関の尺度の例：$r_{AB} = \dfrac{\mathrm{P}[\,A \cap B\,] - \mathrm{P}[\,A\,]\mathrm{P}[\,B\,]}{\sqrt{\mathrm{P}[\,A\,]\mathrm{P}[\,A^c\,]\mathrm{P}[\,B\,]\mathrm{P}[\,B^c\,]}}$ **とおくと** $-1 \leqq r_{AB} \leqq 1$
（相関係数，後述）$A \perp B \Leftrightarrow r_{AB} = 0$ **（事象の独立の時のみ．一般には$\Rightarrow$だけ）**
今の例では$r_{AB} = -\dfrac{1}{14} = -7.1\%$ **（但し，めったにないことに相関係数を使うのは良くない）**

ベイズの定理

結果 A から原因 C についての情報を得る．

定理：$$\mathrm{P}[\,C \mid A\,] = \frac{\mathrm{P}[\,A \mid C\,]\mathrm{P}[\,C\,]}{\mathrm{P}[\,A \mid C\,]\mathrm{P}[\,C\,] + \mathrm{P}[\,A \mid C^c\,]\mathrm{P}[\,C^c\,]} \qquad \diamond$$

証明：$\mathrm{P}[\,C \mid A\,] = \dfrac{\mathrm{P}[\,C \cap A\,]}{\mathrm{P}[\,A\,]} = \dfrac{\mathrm{P}[\,A \mid C\,]\mathrm{P}[\,C\,]}{\mathrm{P}[\,A\,]}$ の分母について，既出事項，$(A \cap C) \cap (A \cap C^c) \subset C \cap C^c = \emptyset$ と
$D \cap E = \emptyset \;\Rightarrow\; \mathrm{P}[\,D \cup E\,] = \mathrm{P}[\,D\,] + \mathrm{P}[\,E\,]$ （加法性；定義），から $\mathrm{P}[\,A\,] = \mathrm{P}[\,A \mid C\,]\mathrm{P}[\,C\,] + \mathrm{P}[\,A \mid C^c\,]\mathrm{P}[\,C^c\,]$ $\square$

$\mathrm{P}[\,C \mid A\,]$ と $\mathrm{P}[\,A \mid C\,]$ は数学的にはどちらで条件付けるかの違いに過ぎないが，因果関係・時間的前後関係があるとき応用上の意味が違う
$\mathrm{P}[\,C\,]$ を**事前確率**，$\mathrm{P}[\,C \mid A\,]$ を**事後確率**，と言うことがある

ベイズの定理の例

（再掲）．工場Bは不良品率1%工場Cは3% ，同数ずつ製造．不良品を買う事象 A，工場B,Cの製品という事象 B, C

例：買った品が不良品だったとき工場Cの製品だった確率

$\mathrm{P}[\,A \mid C\,] = 0.03$, $\mathrm{P}[\,A \mid C^c\,] = \mathrm{P}[\,A \mid B\,] = 0.01$, $\mathrm{P}[\,C\,] = \mathrm{P}[\,C^c\,] = 0.5$

ベイズの定理から，$\mathrm{P}[\,C \mid A\,] = \dfrac{0.03 \times 0.5}{0.03 \times 0.5 + 0.01 \times 0.5} = 75\%$

・不良品だったという結果から，不良品の多く出るCの製品の可能性をより大きく疑うのは自然，という考えの定量化

先験的確率の問題点

問 4–2：人口 10 万人の市で殺人事件が起こり X 氏が容疑者となった．犯人が残した血痕の血液型は 100 人に 1 人の割合である．X 氏は同じ血液型だった．氏が犯人の確率はいくらか？ ◇

B を X 氏が犯人である事象，A を X 氏の血液型が犯人のと一致する事象，X 氏が犯人であることの事前確率 $p = \mathrm{P}[\,B\,]$ とおく．$p = 0.5$（五分五分）とするとどうなるか？

・ベイズ統計学＋大規模データ機械処理の現代

ベイズ統計学復権を強力に進めた 20 世紀前半の統計学者ですら…

ワルド：間違った意志決定に基づく行動が質的に異なる事態に至る，たとえばある選択肢の誤りは経済的損失で別の選択肢の誤りは死に至るとき，（非ベイズの）制約型の条件が意味を持つ．

サヴェッジ：漂流するボートを日没までに探さないと間に合わないときは（さすがに）区間推定を要する．

・仕事と割り切って客観的・機械的に処理するには向く（写真の自動焦点や入園顔認証＝間違えてもご愛敬・小規模金銭損失）が，判断の人や社会への影響が重大なことに使うべきでない．

モンティ・ホール問題の類題

問 4–3：３人の囚人Ａ，Ｂ，Ｃが牢屋にいる．Ａは３人のうち２人が翌日処刑され１人が釈放されることを知ったが，本人の運命は事前に本人には教えない決まりだった．そこでＡは看守に「3人のうち２人が処刑されるから，１人分の情報なら自分のことは決まらない．ＢとＣのうち処刑される者の名前を１人だけ教えてほしい.」と言った．看守は納得して「Ｂは処刑される」と答えた．Ａはこれを聞いて，釈放される可能性があるのは自分とＣだけになったので，釈放の確率が1/3から1/2に増えたと喜んだ．これは正しい判断か？ ◇

例題解く時間30分

問へのコメント

問4–1（含む問4–3）と問4–2の答えは,「統計と確率の基礎 第３版」練習問題11問1と11章§2説明を参照

ヒント：
問4–1（モンティ・ホール問題）は情報を得た後のドアＡとＣの比較，問4–3（3囚人問題）は囚人Ａについて，情報を得る前後の比較．同じベイズの定理が使える．
・Ａが正解の時に司会者がＢとＣどちらを開けるか（Ａが処刑される時に看守がＢとＣどちらを指すか）の癖（確率 p）によって，事後確率が $1/3$ から $1/2$ まで変わる．

問4–2は，血痕の一致を考慮してもＸ氏が犯人の確率0.1% に満たない．
・証拠に合致する無実の人に会うほうが犯罪者に出会うよりもはるかに大きい確率．

めったにない結果に見えてもより珍しい原因の証拠とはできない．
例：地震雲（地震が無い普段の状態で起こる現象は役立たない）
・１つの情報では先験的確率に引きずられる危険（先入観）

まとめ

事象 B の下で A が起きる**条件付き確率** $\mathrm{P}[\,A \mid B\,] = \dfrac{\mathrm{P}[\,A \cap B\,]}{\mathrm{P}[\,B\,]}$

情報による更新：先験的（事前）確率 $\mathrm{P}[\,A\,]$→事後確率 $\mathrm{P}[\,A \mid B\,]$

A と B が**独立**：$\mathrm{P}[\,A \cap B\,] = \mathrm{P}[\,A\,]\mathrm{P}[\,B\,]$

（$\mathrm{P}[\,B\,] \neq 0$ のとき $\Leftrightarrow \mathrm{P}[\,A \mid B\,] = \mathrm{P}[\,A\,]$）

ベイズの定理：$\mathrm{P}[\,C \mid A\,] = \dfrac{\mathrm{P}[\,A \mid C\,]\mathrm{P}[\,C\,]}{\mathrm{P}[\,A \mid C\,]\mathrm{P}[\,C\,] + \mathrm{P}[\,A \mid C^c\,]\mathrm{P}[\,C^c\,]}$

功：パラドックスの一つの解決（モンティ・ホール問題，3囚人問題）
罪：珍しい結果に見えてもより珍しい原因の**証拠**とはできない
罪：先験的確率に引きずられる危険（**先入観**）

積み残し：確率過程・確率連鎖，相関係数

5．試験問題の総数

- 確率変数
- 集合の定義関数 1_A
- 期待値
- 期待値の線形性
- 確率変数の独立

前々回に続いて，いろんな基礎定義と基礎性質

確率変数と期待値

問 5–1：1 回 50 問からなるある就活関連 web の適性試験を 2 回受けると平均してちょうど 1 問がかぶる．この web テストの問題の総数はいくらか？ ◇

確率変数：「ばらつく量」，確率測度を固定して確率変数間で比較
期待値：「中を取った目安」，確定値の数学化
・期待値の**線形性** → 複雑な数学的処理に見通し

確率変数とは確率空間 $(\Omega, \mathcal{F}, \mathrm{P})$ 上の関数（**簡単版**） $X: \ \Omega \to \mathbb{R}$

精密には**可測関数**（$X^{-1}((-\infty, a]) = \{\omega \in \Omega \mid X(\omega) \leq a\} \in \mathcal{F},\ a \in \mathbb{R}$） （秋）

確率変数 X の**期待値** $\mathrm{E}[\,X\,] = \sum_{\omega \in \Omega} X(\omega) \mathrm{P}[\,\{\omega\}\,]$

注：根元事象がある場合の**簡単版**．$\Omega = \mathbb{R}$ では不可→**積分** （秋）

期待値の例といくつかの記号

Ωが有限集合→Xの値域$R_X = \{X(\omega) \mid \omega \in \Omega\}$ も有限集合

$$\mathsf{E}[\,X\,] = \sum_{x \in R_X} x\,\mathsf{P}[\,\{\omega \mid X(\omega) = x\}\,]$$ 簡単版だが一般化少し容易

以後の略記　例：Xが性質Aを満たす確率

$$\mathsf{P}[\,\{\omega \in \Omega \mid X(\omega) \in A\}\,] = \mathsf{P}[\,X \in A\,]$$

- $$\mathsf{E}[\,X\,] = \sum_{x \in R_X} x\,\mathsf{P}[\,X = x\,]$$

例：サイコロの目$\Omega = \{1,2,3,4,5,6\}$,

$$\mathsf{P}[\,\{1\}\,] = \mathsf{P}[\,\{2\}\,] = \mathsf{P}[\,\{3\}\,] = \mathsf{P}[\,\{4\}\,] = \mathsf{P}[\,\{5\}\,] = \mathsf{P}[\,\{6\}\,] = \frac{1}{6}$$

$$X(\omega) = \omega,\ \omega \in \Omega,\ \mathsf{E}[\,X\,] = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = 3.5$$

確率変数$X:\ \Omega \to \mathbb{R}$と関数$f:\ \mathbb{R} \to \mathbb{R}$の合成関数は確率変数

$f \circ X:\ \Omega \to \mathbb{R};\ (f \circ X)(\omega) = f(X(\omega))$　　$f(X)$と書くこと多い

例：n次モーメント $= \mathsf{E}[\,X^n\,]$　（$f(x) = x^n$との合成X^nは確率変数）

期待値の線形性と集合の定義関数

- **線形性：a と b が定数，X と Y が確率変数のとき，**

$\mathsf{E}[\,aX + bY\,] = a\mathsf{E}[\,X\,] + b\mathsf{E}[\,Y\,]$　（黒板）

X と Y の関係にも，有限集合か無限集合かにも，一切関係なく期待値があれば常に成立

このことからさらに **非負値確率変数の期待値は非負**（期待値の非負値性）

特に，$X \geqq Y$ ならば $X - Y \geqq 0$ なので，期待値の線形性と合わせると，

- $X \geqq Y \ \Rightarrow\ \mathsf{E}[\,X\,] \geqq \mathsf{E}[\,Y\,]$　（期待値の単調性）

事象 $A \in \mathcal{F}$ の定義関数：$\mathbf{1}_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A, \end{cases}$ で定義される確率変数 $\mathbf{1}_A$：　$\Omega \to \mathbb{R}$　（教科書では δ_A や χ_A）

- $\mathbf{1}_{A\cap B}(\omega) = \mathbf{1}_A(\omega)\,\mathbf{1}_B(\omega),\ \mathbf{1}_{A^c}(\omega) = 1 - \mathbf{1}_A(\omega)$

ドモルガンの法則と合わせると，集合算を関数の加法や乗法で書ける！

- $\mathsf{E}[\,\mathbf{1}_A\,] = 1 \times \mathsf{P}[\,A\,] + 0 \times \mathsf{P}[\,A^c\,]\ = \mathsf{P}[\,A\,]$

集合関数としての確率測度から確率変数の期待値へ

期待値の線形性の例

例：事象 $A_i,\ i=1,\dots,n,$ の少なくとも一つが起こる事象 $B=\bigcup_{i=1}^{n}A_i$ の確率 $\mathsf{P}[\,B\,]$．$B^c=\bigcap_{i=1}^{n}A_i^c$（ド・モルガンの法則）だから，

$$\mathsf{P}[\,B\,]=1-\mathsf{P}[\,\bigcap_{i=1}^{n}A_i^c\,]=1-\mathsf{E}[\,\mathbf{1}_{\bigcap_{i=1}^{n}A_i^c}\,]=1-\mathsf{E}[\,\prod_{i=1}^{n}\mathbf{1}_{A_i^c}\,]$$

$$=1-\mathsf{E}[\,\prod_{i=1}^{n}(1-\mathbf{1}_{A_i})\,])=\sum_i\mathsf{E}[\,\mathbf{1}_{A_i}\,]-\sum_{i<j}\mathsf{E}[\,\mathbf{1}_{A_i}\mathbf{1}_{A_j}\,]+\dots$$

$$=\sum_i\mathsf{P}[\,A_i\,]-\sum_{i<j}\mathsf{P}[\,A_i\cap A_j\,]+\dots$$ （包除原理）

$n=2$ のとき $\mathsf{P}[\,A\cup B\,]=\mathsf{P}[\,A\,]+\mathsf{P}[\,B\,]-\mathsf{P}[\,A\cap B\,]$ （既出）
$n=3$ のとき $\mathsf{P}[\,A\cup B\cup C\,]=\mathsf{P}[\,A\,]+\mathsf{P}[\,B\,]+\mathsf{P}[\,C\,]-\mathsf{P}[\,A\cap B\,]-\mathsf{P}[\,B\cap C\,]-\mathsf{P}[\,C\cap A\,]+\mathsf{P}[\,A\cap B\cap C\,]$ （既出）
既出の解き方に比べ，期待値の線形性を使えば展開と積で機械的に出る

確率変数の独立

確率変数 $X_1, \ldots, X_n$ が**独立**：任意の実数 $a_1 \ldots, a_n$ について

$$\mathsf{P}[\, X_1 \leqq a_1, \ldots, X_n \leqq a_n \,] = \prod_{k=1}^{n} \mathsf{P}[\, X_k \leqq a_k \,]$$

・$X_1, \ \ldots, X_n$ が独立 $\Rightarrow \ \mathsf{E}\left[\, \prod_{k=1}^{n} X_k \,\right] = \prod_{k=1}^{n} \mathsf{E}[\, X_k \,]$ （逆は不成立）

・X と Y が独立ならば任意の実数値関数（簡単版）f と g について $f(X)$ と $g(Y)$ も独立　（一般には可積分関数，逆も成り立つ）

問5–2：確率変数たちの値域が有限集合の場合に

$$\mathsf{E}[\, f(X)\, g(Y) \,] = \mathsf{E}[\, f(X) \,]\, \mathsf{E}[\, g(Y) \,]$$

を期待値の定義に戻って直接証明せよ $\diamond$

例：$X =$ サイコロ１の目，$Y =$ サイコロ２の目．目の積の期待値．直接積を計算すると大変だが，$\mathsf{E}[\, XY \,] = \mathsf{E}[\, X \,]\mathsf{E}[\, Y \,] = 3.5 \times 3.5 = \dfrac{49}{4}$

確率と期待値についての公式

命題． X が非負値 $(X \geqq 0)$ ならば $\mathsf{E}[\,X\,] = \int_0^\infty \mathsf{P}[\,X > x\,]\,dx$ ◇

簡単版だけでなく，一般に成り立つ公式　（フビニの定理の例題）
証明（値域が非負値有限集合 $R_X = \{x_0 = 0, x_1, \ldots, x_r\}$（$x_0 < \ldots$）の場合）：
$x_{j-1} \leqq y < x_j \;\Rightarrow\; \mathsf{P}[\,X > y\,] = \sum_{i=j}^{r} \mathsf{P}[\,X = x_i\,]$ なので（黒板）$\int_0^\infty \mathsf{P}[\,X > y\,]\,dy = \sum_{j=1}^{r}\sum_{i=j}^{r}(x_j - x_{j-1})\mathsf{P}[\,X = x_i\,] = \sum_{i=1}^{r}\mathsf{P}[\,X = x_i\,]\sum_{j=1}^{i}(x_j - x_{j-1})$

$$= \sum_{i=1}^{r} x_i \mathsf{P}[\,X = x_i\,] = \sum_{i=0}^{r} x_i \mathsf{P}[\,X = x_i\,] = \sum_{x \in R_X} x\mathsf{P}[\,X = x\,] = \mathsf{E}[\,X\,] \quad \Box$$

- X の分布関数 $\mathsf{P}[\,X \leqq y\,] = 1 - \mathsf{P}[\,X > y\,]$ から直接期待値を求める公式（フビニは重積分の積分順序の交換）

確率変数の独立と積の期待値について

（再掲）X と Y が独立ならば $\mathsf{E}[\,XY\,] = \mathsf{E}[\,X\,]\mathsf{E}[\,Y\,]$

前ページの命題から，

$$\mathsf{E}[\,X\,]\mathsf{E}[\,Y\,] = \int_0^\infty \int_0^\infty \mathsf{P}[\,X > x\,]\mathsf{P}[\,Y > y\,]\,dx\,dy$$

$$= \int_0^\infty \int_0^\infty \mathsf{P}[\,X > x,\ Y > y\,]\,dx\,dy$$

命題と同様に２つの確率変数の値の組で場合分けして計算すると，右辺は $\mathsf{E}[\,XY\,]$ に等しい． ◇

• X,Y の値域 R_X,R_Y が有限集合のとき，X,Y が独立 $(X \perp Y) \Leftrightarrow$ 任意の $a \in R_X$ と $b \in R_Y$ に対して事象 $\{X = a\}$ と $\{Y = b\}$ が独立．

（再掲）事象 $A,\ B$ が独立：$\mathsf{P}[\,A \cap B\,] = \mathsf{P}[\,A\,]\mathsf{P}[\,B\,]$

問の答え

問 5–2 の答えは,「統計と確率の基礎 第３版」練習問題３問１を参照

問 5–1 ： １回 50 問からなるある就活関連 web の適性試験を２回受けると平均してちょうど１問がかぶる．この web テストの問題の総数はいくらか？ ◇
（2012 年３年加藤裕太君の問題）

Web サイトの問題の集合（ストック）を $S=\{1,2,\ldots,N\}\subset\mathbb{R}$，１回目の問の集合を $A\subset\Omega$ $(\sharp A=50)$，$i=1,\ldots,50$ について，２回目の問 i を $X_i:\ \Omega\to S$ とおくと，

$$1=\mathrm{E}[\,\sum_{i=1}^{50}1_{X_i\in A}\,]=\sum_{i=1}^{50}\mathrm{P}[\,X_i\in A\,]=50\times\frac{50}{N}$$ だから，

$N=2500$．

（Ω は S から２回目の 50 問を取り出す組み合わせ全部を集めた集合で，その上の確率を等確率で定義した.）

まとめ

期待値： $\mathsf{E}[\,X\,] = \sum_{x\in R_X} x\,\mathsf{P}[\,X = x\,]$ （X の値域 R_X）

線形性， $\mathsf{E}[\,\mathbf{1}_A\,] = \mathsf{P}[\,A\,]$

分布関数 $F(x) = \mathsf{P}[\,X \leqq x\,] = \mathsf{E}[\,\mathbf{1}_{X\leqq x}\,]$，$n$ 次**モーメント** $\mathsf{E}[\,X^n\,]$

$X \geqq 0$ ならば $\mathsf{E}[\,X\,] = \int_0^\infty \mathsf{P}[\,X > y\,]\,dy$，包除原理

記号：X が性質 A を満たす確率 $\mathsf{P}[\,\{\omega\in\Omega \mid X(\omega)\in A\}\,] = \mathsf{P}[\,X\in A\,]$

$$\mathbf{1}_A(\omega) = \begin{cases} 1, & \omega\in A, \\ 0, & \omega\notin A, \end{cases} \quad (\mathbf{1}_{A\cap B}(\omega) = \mathbf{1}_A(\omega)\,\mathbf{1}_B(\omega),\ \mathbf{1}_{A^c}(\omega) = 1 - \mathbf{1}_A(\omega))$$

確率変数が独立：それぞれある値をとる積事象の確率が個別にその値をとる確率の積に等しいこと（簡単版）

独立ならば，積の期待値は期待値の積に等しい（逆は不成立）

$X \perp Y \quad\Rightarrow\quad f(X)\perp g(Y)$ （$\rightarrow \mathsf{E}[\,f(X)g(Y)\,] = \mathsf{E}[\,f(X)\,]\mathsf{E}[\,g(Y)\,]$）

積み残し：連続値確率変数の場合の期待値，可測関数，フビニの定理

６．なぜ一列並びか

- 分散
- 共分散，相関係数
- 独立確率変数の分散の加法性

教科書の第２章に基づいて，前回の定義を少し応用します

決まった量とばらつく量の決定的な違い

問6–1：ATMや窓口が複数ある時，(空港や銀行などで見かけるように）一列に並んで空いた窓口を利用する一列並びと，(スーパーで多く見かけるように）窓口ごとに列ができる並列並びがある．一列並びが並列並びより良いのはどういう点か？ ◇

待ち時間の**期待値**は（ほぼ）等しいので，平均だけで考えるとどちらでも良いはず

決定論的な法則と**分布**を持つ法則の決定的な違い
バラツキ（**分散**）が0でないことの意味

分散

バラツキの（もっとも素朴・基礎的な）目安

確率変数 X の分散 $\mathsf{V}[\,X\,] = \mathsf{E}[\,(X - \mathsf{E}[\,X\,])^2\,]$

注：期待値の定義は簡単版を紹介したが，分散を含めて，期待値を用いて紹介した定義は全て一般に通用する

期待値の線形性から，$\mathsf{V}[\,X\,] = \mathsf{E}[\,X^2\,] - \mathsf{E}[\,X\,]^2$

標準偏差 $\sigma_X = \sqrt{\mathsf{V}[\,X\,]} = \sqrt{\mathsf{E}[\,(X - \mathsf{E}[\,X\,])^2\,]}$

$\mathsf{E}[\,X\,]$ との比が物差しや秤の単位を変えても変わらない量が実用的（確率論の数学では分散，データに当てはめる統計学では標準偏差）

k 定数→ $\mathsf{V}[\,k\,X\,] = k^2\mathsf{V}[\,X\,]$ $\quad \sigma_{k\,X} = |k|\,\sigma_X$ $\quad \mathsf{E}[\,k\,X\,] = k\mathsf{E}[\,X\,]$

注．加法性 $\mathsf{E}[\,X + Y\,] = \mathsf{E}[\,X\,] + \mathsf{E}[\,Y\,]$ は分散では通常は不成立．

分布の平均と分散

・確率変数 X の分布：$Q(A) = \mathsf{P} \circ X^{-1}(A) = \mathsf{P}[\, X \in A \,]$ で決まる $\mathbb{R}$ 上の確率

確率の定義（全確率 1 と非負値性と加法性）は P が確率であることから

注．X は「起きなかったこと」の集合 Ω の上の関数，分布 Q は X の**値域**「取り得る実数」の上の確率 $(\Omega = \mathbb{R})$

・実数上の確率測度（分布）Q があるとき $Q(\cdot) = \mathsf{P}[\, X \in \cdot \,]$ を満たす X を，分布 Q に従う確率変数と言う

・分布の平均と分散＝その分布に従う確率変数の期待値と分散

・有限集合上の分布 Q の平均 m と分散 v：

$$m = \mathsf{E}[\, X \,] = \sum_{x \in R_X} x \mathsf{P}[\, X = x \,] = \sum_{x \in R_X} x Q(\{x\})$$

$$v = \mathsf{E}[\, (X - m)^2 \,] = \sum_{x \in R_X} (x - m)^2 Q(\{x\})$$

問 6–2：$v = \displaystyle\sum_{x \in R_X} x^2 Q(\{x\}) - m^2$ を示せ． ◇

共分散と相関係数

分散：X のバラツキの目安

共分散と相関係数：X と Y のバラツキ方の類似の目安

共分散 $C(X,Y) = \mathsf{E}[\,(X - \mathsf{E}[\,X\,])(Y - \mathsf{E}[\,Y\,])\,]$

相関係数 $r(X,Y) = \dfrac{C(X,Y)}{\sqrt{\mathsf{V}[\,X\,]\mathsf{V}[\,Y\,]}}$

(1) **単位を変えても変わらない量が実用的（係数，指数）**

(2) $-1 \leqq r(X,Y) \leqq 1$ **が常に成り立つ．**

証明：$\mathsf{E}[\,((X - \mathsf{E}[\,X\,]) - t\,(Y - \mathsf{E}[\,Y\,]))^2\,] \geqq 0$ **（既出）**

左辺を展開して期待値の線形性を用い，$t = \sqrt{\dfrac{\mathsf{V}[\,X\,]}{\mathsf{V}[\,Y\,]}}$ **を代入すると**

$\mathsf{V}[\,X\,] - 2\sqrt{\dfrac{\mathsf{V}[\,X\,]}{\mathsf{V}[\,Y\,]}}C(X,Y) + \mathsf{V}[\,X\,] \geqq 0$ **となるので，**$r(X,Y) \leqq 1$**．**

最初で $-$ **の代わりに** $+$ **とすれば** $r(X,Y) \geqq -1$**．** □

相関係数の性質

$\mathsf{V}[\,X+Y\,]=\mathsf{V}[\,X\,]+\mathsf{V}[\,Y\,]+2\,C(X,Y) \qquad \mathsf{V}[\,X\,]=C(X,X)$

$C(X,Y)=\mathsf{E}[\,XY\,]-\mathsf{E}[\,X\,]\mathsf{E}[\,Y\,] \qquad \mathsf{V}[\,X\,]=\mathsf{E}[\,X^2\,]-\mathsf{E}[\,X\,]^2$

例：$r(\mathbf{1}_A,\mathbf{1}_B)=\dfrac{\mathsf{P}[\,A\cap B\,]-\mathsf{P}[\,A\,]\mathsf{P}[\,B\,]}{\sqrt{\mathsf{P}[\,A\,]\mathsf{P}[\,A^c\,]\mathsf{P}[\,B\,]\mathsf{P}[\,B^c\,]}}=r_{AB}$ **（既出）**

X, Y **が独立ならば** $C(X,Y)=\mathsf{E}[\,X-\mathsf{E}[\,X\,]\,]\mathsf{E}[\,Y-\mathsf{E}[\,Y\,]\,]=0$.

よって， X, Y **が独立ならば** $\mathsf{V}[\,X+Y\,]=\mathsf{V}[\,X\,]+\mathsf{V}[\,Y\,]$.

独立ならば分散にも加法性がある

問 6–3：n **個のさいころを振ったとき，1の目の出る個数** X **と2の目の出る個数** Y **の相関係数** $r(X,Y)$ **を求めよ．** ◇

一列並びと並列並び

問 6–1：一列並びが並列並びより良いのはどういう点か？ ◇

単純な数理モデル：窓口 M 個 $(M > 1)$，自分が待ち N 人目，客 i の処理時間 S_i は確率変数で，その期待値 $\tau = \mathsf{E}[\, S_i \,]$ は客にも窓口にもよらない．処理以外の移動などの時間は無視する．

一列並びのときの自分の待ち時間 $T^{(1)} = \dfrac{1}{M}(S_1 + \cdots + S_N)$

並列並びのときの自分の待ち時間 $T^{(2)} = S_{i_1} + \cdots + S_{i_{N/M}}$ **（i_j たちは自分の並ぶ窓口の待ち人たち）**

結論 1．期待値の線形性（既出）から，

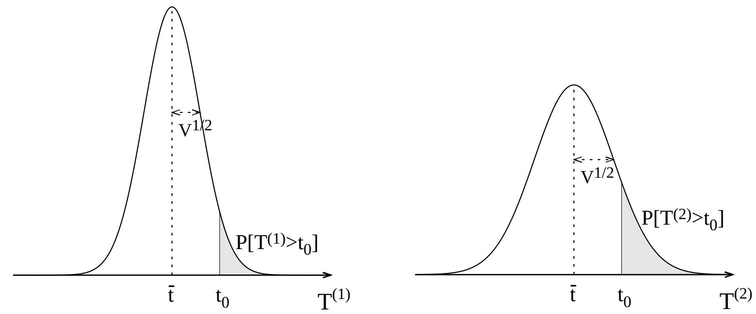
$$\mathsf{E}[\, T^{(1)} \,] = \mathsf{E}[\, T^{(2)} \,] = \frac{N}{M}\tau =: \bar{t}$$

平均（期待値）だけを見るならば，並列並びと一列並びに優劣はない

（平均待ち時間の短縮は起きない）

バラツキは均せるとは限らない

待ち時間の余裕 t_0
買えぬ確率 $= \mathsf{P}[\, T^{(\alpha)} > t_0 \,]$
その**目安** $\mathsf{V}[\, T^{(\alpha)} \,]$



「統計と確率の基礎 第３版」２章図９

モデルへの追加：平均的には余裕を見込んで並ぶ場合 $(t_0 > \bar{t})$，
$v = \mathsf{V}[\, S_i \,]$ は客にも窓口にもよらない．
結論２．**独立**確率変数の**分散の加法性**（と，２次同次性）**（既出）**から

一列並び：$\mathsf{V}[\, T^{(1)} \,] = \mathsf{V}[\, \frac{1}{M}(S_1 + \cdots + S_N) \,] = \frac{1}{M^2} \sum_{i=1}^{N} \mathsf{V}[\, S_i \,] \;\; = \frac{N}{M^2} v$

並列並び：$\mathsf{V}[\, T^{(2)} \,] = \mathsf{V}[\, \sum_{j=1}^{N/M} S_{i_j} \,] = \mathsf{V}[\, S_{i_1} \,] + \cdots + \mathsf{V}[\, S_{i_{N/M}} \,] \;\; = \frac{N}{M} v$

一列並びのほうがバラツキ（**分散**）が小さい！
バラツキが小さいので，**平均的には余裕を見込んで並ぶ人にとっては一列並びのほうが優れる**（運悪く間に合わない確率が減る）

問へのコメント

問6–2と問6–3の答えは,「統計と確率の基礎 第3版」練習問題1問1と練習問題3問2を参照.

問6–3のヒント：k 個目のサイコロの目を Z_k とおくと,

$$X = \sum_{k=1}^{n} 1_{Z_k=1},\ Y = \sum_{k=1}^{n} 1_{Z_k=2}.$$

注：教科書練習問題3問2とその略解では $1_{Z_k=1}$ のことを $\delta_{Z_k,1}$ などと書いている．両者は同じ意味．

$1_{X=a}$（$\delta_{X,a}$）は，$X(\omega) = a$ となる $\omega \in \Omega$ に対して1，そうでない ω に対して0となる Ω 上の関数（確率変数）$1_{X=a}:\ \Omega \to \mathbb{R}$.

元の問題に戻ると，$\mathsf{E}[\,X\,] = \mathsf{E}[\,\sum_{k=1}^{n} 1_{Z_k=1}\,] = \sum_{k=1}^{n} \mathsf{E}[\,1_{Z_k=1}\,]$（期待値の線形性）に既出の性質 $\mathsf{E}[\,1_{Z=a}\,] = \mathsf{P}[\,\{Z = a\}\,]$ を当てはめることになる

問へのコメント（続き）

$\mathsf{E}[\ X^2\ ] = \mathsf{E}[\ (\sum_{k=1}^{n} \mathbf{1}_{Z_k=1})^2\ ]$**について．**

$(\sum_{k=1}^{n} \mathbf{1}_{Z_k=1})^2 = (\mathbf{1}_{Z_1=1} + \cdots + \mathbf{1}_{Z_n=1})^2$

$= \mathbf{1}_{Z_1=1}^2 + \mathbf{1}_{Z_1=1}\mathbf{1}_{Z_2=1} + \mathbf{1}_{Z_1=1}\mathbf{1}_{Z_3=1} + \cdots + \mathbf{1}_{Z_1=1}\mathbf{1}_{Z_n=1}$

$+ \mathbf{1}_{Z_2=1}\mathbf{1}_{Z_1=1} + \cdots + \cdots + \mathbf{1}_{Z_n=1}^2$

2つの性質

- $\mathbf{1}_{Z_1=1}^2 = \mathbf{1}_{Z_1=1}$
- $\mathbf{1}_{Z_1=1}$ **と** $\mathbf{1}_{Z_2=1}$ **は独立なので**

$\mathsf{E}[\ \mathbf{1}_{Z_1=1}\mathbf{1}_{Z_2=1}\ ] = \mathsf{E}[\ \mathbf{1}_{Z_1=1}\ ]\mathsf{E}[\ \mathbf{1}_{Z_2=1}\ ] = \frac{1}{6} \times \frac{1}{6}$

を用いる（以下教科書略解参照）

まとめ

分散 $\mathsf{V}[\,X\,] = \mathsf{E}[\,(X - \mathsf{E}[\,X\,])^2\,]$, **標準偏差** $\sigma_X = \sqrt{\mathsf{V}[\,X\,]}$

X の**分布** $Q(A) = \mathsf{P}[\,X \in A\,]$ $(\Omega = \mathbb{R})$ （Q に**従う**確率変数）

分布の平均mと分散v＝その分布に従う確率変数の期待値と分散

共分散 $C(X,Y) = \mathsf{E}[\,(X - \mathsf{E}[\,X\,])(Y - \mathsf{E}[\,Y\,])\,]$,

相関係数 $r(X,Y) = \dfrac{C(X,Y)}{\sqrt{\mathsf{V}[\,X\,]\mathsf{V}[\,Y\,]}}$

公式：$\mathsf{V}[\,X\,] = \mathsf{E}[\,X^2\,] - \mathsf{E}[\,X\,]^2$, $\mathsf{V}[\,kX\,] = k^2\mathsf{V}[\,X\,]$, $\sigma_{kX} = |k|\sigma_X$, $\mathsf{V}[\,X+Y\,] = \mathsf{V}[\,X\,] + \mathsf{V}[\,Y\,] + 2C(X,Y)$, $\mathsf{V}[\,X\,] = C(X,X)$, $C(X,Y) = \mathsf{E}[\,XY\,] - \mathsf{E}[\,X\,]\mathsf{E}[\,Y\,]$, $-1 \leqq r(X,Y) \leqq 1$, $m = \sum xQ(\{x\})$, $v = \sum (x-m)^2 Q(\{x\}) = \sum x^2 Q(\{x\}) - m^2$

独立確率変数ならば $\mathsf{V}[\,X + \cdots + Y\,] = \mathsf{V}[\,X\,] + \cdots + \mathsf{V}[\,Y\,]$
(cf. $\mathsf{E}[\,kX\,] = k\mathsf{E}[\,X\,]$, $\mathsf{E}[\,X + \cdots + Y\,] = \mathsf{E}[\,X\,] + \cdots + \mathsf{E}[\,Y\,]$)

積み残し：**分散は** $\mathsf{P}[\,X \geqq a\,]$ **（偏差）の目安**

７．２項定理再び

- 確率測度の最初の具体例　－　硬貨投げ
- 直積確率測度と独立性
- ２項分布とその性質

高校教科書で曖昧だった独立試行を事象で書くためのヒント：
確率空間の定義（公理）を満たすように現実を当てはめる

教科書第１章に戻ります

レポート２ ＠ keio.jp

集合の直積

$0 < p < 1$.「確率 p で表」の確率空間：Ω ($=$ {裏, 表}) $= \{0,1\}$,
$\mathcal{F} = \{\emptyset, \{0\}, \{1\}, \Omega\}$, $\mathrm{P}[\,\{1\}\,] = p$, $\mathrm{P}[\,\{0\}\,]$, $\mathrm{P}[\,\emptyset\,]$, $\mathrm{P}[\,\Omega\,]$ は各自

前回まで（果物の集合の例等を除き）具体的な確率分布を殆ど使わなかった．定義（非負値，加法性，全確率 1）だけで多くの結論を得る．極限を使えばさらにたくさんの結論を得る（後述）（確率論の数学としての豊かさの例）．しかし，
数値を得るには具体的な確率測度（分布）が必要．最単純が硬貨投げ．
次がその直積（独立な繰り返し）：ベルヌーイ試行．

集合 A と B の直積 $A \times B = \{(\omega_1, \omega_2) \mid \omega_1 \in A,\ \omega_2 \in B\}$
例：$A = \{$太郎, 次郎$\}$, $B = \{$花子, 恵子$\}$,
$A \times B = \{$(太郎, 花子), (太郎, 恵子), (次郎, 花子), (次郎, 恵子)$\}$
- $(x, y) \in A \times B \iff x \in A$ かつ $y \in B$

確率の直積

$(\Omega, \mathcal{F}, \mathrm{P})$ が $(\Omega_1, \mathcal{F}_1, \mathrm{P}_1)$ と $(\Omega_2, \mathcal{F}_2, \mathrm{P}_2)$ の**直積**確率空間とは，

1．$\Omega = \Omega_1 \times \Omega_2 \ = \{(\omega_1, \omega_2) \mid \omega_i \in \Omega_i,\ i = 1, 2\}$,

2．$\mathcal{F} = \sigma[\{A \times B \mid A \in \mathcal{F}_1,\ B \in \mathcal{F}_2\}]$,

（部分集合の直積集合から和集合と補集合で作られる集合）

3．$\mathrm{P}[\,A \times B\,] = \mathrm{P}_1[\,A\,]\,\mathrm{P}_2[\,B\,]$ （直積確率測度 $\mathrm{P} = \mathrm{P}_1 \times \mathrm{P}_2$）．

• 分布 Q_i に従う確率変数 $X_i:\ \Omega \to \mathbb{R}\ (i = 1, 2)$ について X_1, X_2 が独立 $\Leftrightarrow$ (X_1, X_2) の分布（$\mathbb{R}^2$ 上の確率，**結合分布**）が $Q_1 \times Q_2$

例：ベルヌーイ試行（硬貨投げの繰り返し）

証明：独立なら $Q(A \times B) = \mathrm{P}[\,(X_1, X_2) \in A \times B\,] = \mathrm{P}[\,X_1 \in A,\ X_2 \in B\,] = \mathrm{P}[\,X_1 \in A\,]\mathrm{P}[\,X_2 \in B\,] = Q_1(A)\,Q_2(B)$．逆はこの式を中ほどで切って左右をつなげばわかる．　□

ベルヌーイ試行と確率変数列の独立

表の確率 p の硬貨を n 回投げる（ベルヌーイ試行）：

$\Omega_n = \{(i_1,\dots,i_n) \mid i_k = 0, 1,\ k = 1,\dots,n\}$,

$\mathsf{P}[\,\{(1,\dots,1)\}\,] = p^n$ 等 $\Leftrightarrow$ (Ω_n, P) は 1 回硬貨投げ n 個の直積

k 回目の結果を表す確率変数 $X_k:\ \Omega_n \to \mathbb{R}$; $X_k(i_1,\dots,i_n) = i_k$

X_k の（値の）分布 = 1 回硬貨投げ (k 回目)

→ 確率変数列 $X_1,\dots,X_n$ は独立 （略記： $\{X_i\}_{i=1}^n \perp$）

1 回硬貨投げ：$\mathsf{E}[\,X_k\,] = 1\times\mathsf{P}[\,X_k = 1\,] + 0\times\mathsf{P}[\,X_k = 0\,] = p$,

$\mathsf{V}[\,X_k\,] = \mathsf{E}[\,X_k^2\,] - p^2 = 1^2 \times \mathsf{P}[\,X_k = 1\,] - p^2 = p(1-p)$

n 回投げたときの表の枚数 $N_n = X_1 + \ldots + X_n$;

期待値と分散の加法性から

$\mathsf{E}[\,N_n\,] = \mathsf{E}[\,X_1\,] + \cdots + \mathsf{E}[\,X_n\,] = np$,

$\mathsf{V}[\,N_n\,] = \mathsf{V}[\,X_1\,] + \cdots + \mathsf{V}[\,X_n\,] = np(1-p)$

高校では独立性からではなく２項分布を経由して習う（次頁）

２項分布

ベルヌーイ試行におけるおもての回数（高校の方法の復習）

$N:\ \Omega_n \to \mathbb{R};\ N(i_1,\ldots,i_n) = i_1 + \cdots + i_n$

その分布 $\mathrm{Binom}_{n,p}(\{k\}) = \mathsf{P}\circ N^{-1}(\{k\}) = \mathsf{P}[\,N = k\,]$

$Q(\{k\}) = \mathrm{Binom}_{n,p}(\{k\}) = {}_n\mathrm{C}_k p^k (1-p)^{n-k}$

（Ω_n **ではなく）**$\Omega'_n = \{1,2,\ldots,n\}$ **上の確率測度が決まる：２項分布**

２項定理 $(p+q)^n = \displaystyle\sum_{k=0}^{n} {}_n\mathrm{C}_k p^k q^{n-k}$

$q = 1-p$ **：**$1 = \sum Q(\{k\}) = Q(\Omega'_n)$ **（全確率１）．**

p **で** $1,2$ **回微分して** $q = 1-p$ **とおいて整理すると，**

平均 $m = \sum k\,Q(\{k\}) = np,$

分散 $v = \sum (k-m)^2 Q(\{k\}) = np(1-p)$

（分布の図などは教科書「統計と確率の基礎 第３版」p.6–p.7）

個体数の推定

例（湖の良く動く**魚の個体数**推定）：作業(1) 500匹釣って印を付けて放流．作業(2) バラバラの時と場所を選んで合計1000匹釣ったところ10匹に印が付いていた．魚の個体数Nはどの程度か？ ◇

作業(2)で1匹釣るごとに印が付いている確率$p = \dfrac{500}{N}$．それを1000回繰り返すベルヌーイ試行と見ると，(2)で釣り上げる印付き魚の個体数の分布は$\mathrm{Binom}_{1000,p}$に従う．概ね平均$m = 1000\,p$程度と期待．それが10だから$p = \dfrac{10}{1000}$．従って$N = 5$万程度と見積もる． □

（ウェブテストの問題ストックの推定も同じ）

2項分布のその他の例

例（視聴率）：視聴率 p を，各視聴者が共通に確率 p で見る傾向を持つ（番組がその大きさの人間的魅力を持つ）と理解して，n 人をサンプル調査すると実際に見た人数は $\mathrm{Binom}_{n,p}$ に従う．上記例と同様にサンプル調査から視聴率 p を逆算して見積もることができる．

参考：バラツキも考慮するには，統計的推定を行う　（当該科目参照）

問 7–1：表の出る確率 0.5 の硬貨を $n = 12$ 回投げて表の出た回数（確率変数）を N とおく．期待値 $\mathrm{E}[\,N\,] = 6$ から離れた回数 $A = \{1, 2, \ldots, 6-k, 6+k, 7+k, \ldots, 12\}$ 表が出る事象の確率 $\mathrm{P}[\,N \in A\,]$ が 0.1 以下（$\mathrm{P}[\,|N-6| \geqq k\,] \leqq 0.1$）になる k を求めよ．　◇

ベルヌーイ試行の極限

$\lambda = \mathsf{E}[\,N_n\,] = np$ を一定にして $n \to \infty (p = \dfrac{\lambda}{n} \to 0)$ とすると，

$$\mathsf{P}[\,N_n = k\,] = \mathsf{Binom}_{n,\frac{\lambda}{n}}(\{k\}) = {}_n\mathsf{C}_k \lambda^k n^{-k} (1 - \frac{\lambda}{n})^{n-k}$$

$$= \frac{\lambda^k}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k} (1 - \frac{\lambda}{n})^{n-k} \to e^{-\lambda} \frac{\lambda^k}{k!}$$ **（微積分学既出）**

「１要因あたり確率が目に見えない程小さい（例：１交差点の事故）が要因が独立に多数ある（交差点数）結果，平均 $\lambda > 0$ が目に見える大きさ（毎日の管内交通事故数の１年間）の分布」（ポワッソン分布）

補足

問 7–1 の答えは,「統計と確率の基礎 第３版」練習問題５問１を参照.
参考１：問 7–1 は次の統計的検定の問題と同じ答えである.『表の出た回数にもとづいて帰無仮説 H_0: $p = 0.5$ を両側検定する．表が何回出たとき H_0 を危険率 0.1 で棄却できるか？』
参考２：問 5–1 も,『1 回 50 問からなる就活関連 web の適性試験を２回受けてみたところ，そのうち１問がかぶっていた．Web テストの問題数を推定せよ』という問題．期待値の線形性で答えたが，２項分布の最尤推定量になる.

まとめ

集合の直積 $A \times B$**，直積確率測度**，$X_1 \perp X_2 \Leftrightarrow$ **結合分布が分布の直積**

２**項分布** $\mathsf{P}[\,\{k\}\,] = {}_n\mathsf{C}_k p^k (1-p)^{n-k}$,

ベルヌーイ試行の表の回数 $N(i_1, \ldots, i_n) = \sum_{k=1}^{n} i_k$ **の分布**

k **回目の硬貨投げ** $X_k :\ \Omega_n \to \mathbb{R};\ X_k(i_1, \ldots, i_n) = i_k$

$\{X_k\}_{k=1}^{n} \perp$**，**　$N = \sum_{k=1}^{n} X_k$

$\mathsf{E}[\,N_n\,] = \sum \mathsf{E}[\,X_k\,] = np$,

$\mathsf{V}[\,N_n\,] = \sum \mathsf{V}[\,X_k\,] = np(1-p)$.

８．いつか表が出る

- $\mathbb{Z}_+$（可算集合）上の確率測度
- （可算）無限個の集合算
- 確率の定義（公理）を拡張（加法性→σ加法性，和→級数の和）
- ポワッソン分布

無限集合上の確率論です

例題（ポワッソン分布）

問８：ある駅で１日あたり平均１件忘れ物がある．（忘れ物件数は期間によらずポワッソン分布に従うとして）以下を求めよ．

(i)　１週間忘れ物がない確率．

(ii)　１週間の忘れ物件数の長期間にわたる度数分布においてもっとも度数の高い件数．

非負整数上の確率測度

$\Omega = \mathbb{Z}_+$（または，$\mathbb{Z}_+$ で番号付けられる集合），$\mathcal{F}$ が Ω の部分集合を全て（簡単版）集めた集合族のとき，P がその上の**確率測度**とは，**全測度 1** で**非負値**な**集合関数** $\mathsf{P}: \mathcal{F} \to \mathbb{R}$ で，**σ 加法性**

$\{A_1, A_2, \ldots\} \subset \mathcal{F}$ が $(i \neq j \Rightarrow A_i \cap A_j = \emptyset)$ ならば $\mathsf{P}[\bigcup_{i=1}^{\infty} A_i] = \sum_{i=1}^{\infty} \mathsf{P}[A_i]$

を満たすもの．右辺は普通の級数，つまり，部分和 $s_n = \sum_{i=1}^{n} \mathsf{P}[A_i]$

の作る数列 $s_1, s_2, \ldots$ の**極限**．

$\bigcup_{i=1}^{\infty} A_i = \{x \in \Omega \mid \exists i \in \mathbb{N};\ x \in A_i\}$ ＝どれかの A_i の要素になっている Ω の要素を全て集めた集合

注：$\exists x; \cdots$ は「$\cdots$ を満たす x が存在する」．
集合算の ∞ には極限の意味は**ない**．「どれかの i」「すべての i」

無限個の集合算

$\displaystyle\bigcap_{i=1}^{\infty} A_i = \{x \in \Omega \mid (\forall i \in \mathbb{N})\ x \in A_i\}$ は全ての A_i に共通する要素を全て集めた集合（$(\forall x)\cdots$ は「全ての x に対して…が成り立つ」）

- 和集合，共通部分は**結合法則**が成立（和を取る順序によらない等）
- **分配法則** $(\bigcup A_i) \cap B = \bigcup (A_i \cap B)$, $(\bigcap A_i) \cup B = \bigcap (A_i \cup B)$
- $\displaystyle(\forall n \in \mathbb{N})\ \bigcap_{i=1}^{\infty} A_i \subset A_n \subset \bigcup_{i=1}^{\infty} A_i$, 無限個の $\emptyset$ の和集合は $\emptyset$，
- **ド・モルガンの法則**：$\displaystyle\left(\bigcup_{i=1}^{\infty} A_i\right)^c = \bigcap_{i=1}^{\infty} {A_i}^c$, $\displaystyle\left(\bigcap_{i=1}^{\infty} A_i\right)^c = \bigcup_{i=1}^{\infty} {A_i}^c$

証明：前半 $\displaystyle x \notin \bigcup_{i=1}^{\infty} A_i \ \Leftrightarrow\ (\forall i)\ x \notin A_i$，後半 $\displaystyle x \notin \bigcap_{i=1}^{\infty} A_i \ \Leftrightarrow\ (\exists i)\ x \notin A_i$ $\square$

備考：級数の $\displaystyle\sum_{i=1}^{\infty}$ は部分和の極限という有限の言葉で定義された数．集合算の $\displaystyle\bigcup_{i=1}^{\infty}$ は無限個ある集合に各要素が入っているかどうかという（原理的に無限の確認を要する事項が）数学的帰納法などによって有限の言葉で実行できることを前提にする．

σ加法性と確率測度の性質

（再掲）σ加法性：可算個の（自然数で番号付けられる）事象が排反ならば和集合の確率は各集合の確率の級数の和．排反でなくても成り立つ性質に興味．狭い性質を確率の定義にしておけば，集合関数が確率測度であることを証明するのが楽．使うときは公式を利用（数学の議論展開の常道）

$A_1 = A_2 = \cdots = \emptyset$ ならば排反かつ $\bigcup A_i = \emptyset$ なので，σ加法性から $\mathrm{P}[\,\emptyset\,] = \sum_{i=1}^{\infty} \mathrm{P}[\,\emptyset\,]$．$\mathrm{P}[\,\emptyset\,]$ は非負実数．正だと右辺発散で等号不成立．よって $\mathrm{P}[\,\emptyset\,] = 0$．

$A_1 = A,\ A_2 = B,\ A \cap B = \emptyset,\ i \geqq 3$ のとき $A_i = \emptyset$ とすると，σ加法性から $\mathrm{P}[\,A \cup B\,] = \mathrm{P}[\,A\,] + \mathrm{P}[\,B\,]$（初回簡単版の有限加法性）．よって，既出の $\mathrm{P}[\,A - B\,]$ 関連公式，包除原理，非負値性，単調性，有限劣加法性，ベイズの定理，条件付き確率，確率変数（含む期待値等々とその公式）は $\mathbb{Z}_+$ 上の確率測度でも成立（定義の拡張）

確率の連続性と劣加法性

確率の連続性：$B_1 \subset B_2 \subset \cdots$ ならば $\mathrm{P}[\bigcup_{i=1}^{\infty} B_i] = \lim_{n\to\infty} \mathrm{P}[B_n]$,

$B_1 \supset B_2 \supset \cdots$ ならば $\mathrm{P}[\bigcap_{i=1}^{\infty} B_i] = \lim_{n\to\infty} \mathrm{P}[B_n]$.

証明：$A_1 = B_1$, $i \geqq 2$ のとき $A_i = B_i \setminus B_{i-1}$ とおくと，$i \neq j \Rightarrow A_i \cap A_j = \emptyset$, $\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i$ なので，σ加法性と級数の定義と有限加法性から $\mathrm{P}[\bigcup_{i=1}^{\infty} B_i] = \mathrm{P}[\bigcup_{i=1}^{\infty} A_i] = \sum_{i=1}^{\infty} \mathrm{P}[A_i] = \mathrm{P}[B_1] + \lim_{n\to\infty} \sum_{i=2}^{n} (\mathrm{P}[B_i] - \mathrm{P}[B_{i-1}]) = \mathrm{P}[B_1] + \lim_{n\to\infty} (\mathrm{P}[B_n] - \mathrm{P}[B_1]) = \lim_{n\to\infty} \mathrm{P}[B_n]$．後半はドモルガン □

包含関係もない一般の場合，**劣加法性**：$\mathrm{P}[\bigcup_{n=1}^{\infty} A_n] \leqq \sum_{n=1}^{\infty} \mathrm{P}[A_n]$.

証明：$B_i = \bigcup_{n=1}^{i} A_n$ と有限劣加法性と確率の連続性から □

非負整数上の確率測度

・**標語**：「（適切な）数列があると，$\Omega = \mathbb{Z}_+$ 上の確率測度が作れる」

・数列 $\{a_k \mid k \in \mathbb{Z}_+\}$ が**非負**実数列で級数の**和があって 0 でない**ならばＯＫ

・式で書くと，$a_k \geqq 0,\ k = 0, 1, 2, \ldots,$ かつ $0 < z := \sum_{k=1}^{\infty} a_k < \infty$

・作り方：　根元事象の確率 $\mathrm{P}[\ \{k\}\ ] = \dfrac{a_k}{z},\ k \in \mathbb{Z}_+$

一般の事象 $A \subset \mathbb{Z}_+$ の確率 $\mathrm{P}[\ A\ ] = \displaystyle\sum_{k \in A} \frac{a_k}{z}$

・証明：上で定義された集合関数 $\mathrm{P} :\ 2^{\mathbb{Z}_+} \to \mathbb{R}$ は，非負値性，σ加法性，全測度 1 を満たすので確率測度．

（σ加法性は互いに共通部分を持たない $A_i,\ i = 1, 2, \ldots,$ に対して，$\mathrm{P}[\ \bigcup_{i=1}^{\infty} A_i\ ] = \sum_{i=1}^{\infty} \mathrm{P}[\ A_i\ ]$ を言えば良いが，非負値収束級数は絶対収束するので級数の和の順序を自由に交換しても和が変わらないからＯＫ）　（証明終わり）

ポワッソン分布

$\Omega = \mathbb{Z}_+$ 上の確率測度の例：ポワッソン分布：

$\lambda > 0$ に対して，$\mathrm{P}[\,\{k\}\,] = e^{-\lambda}\dfrac{\lambda^k}{k!}$ （次回詳述）

$\displaystyle\sum \frac{\lambda^k}{k!} = e^{\lambda},\ m \ = \sum k\, e^{-\lambda}\frac{\lambda^k}{k!} \ = \lambda,$

$\displaystyle v \ = \sum k(k-1)\, e^{-\lambda}\frac{\lambda^k}{k!} + m - m^2 \ = \lambda.$

２項分布との比較（前回）から,「1 要因あたり確率が目に見えない程小さい（例：1 交差点の事故）が要因が独立に多数ある（交差点数）結果，平均 $\lambda > 0$ が目に見える大きさ（毎日の管内交通事故数の１年間）の分布」と思える

問８の答えとコメント

問８の答え

１週間の忘れ物件数 k の分布は平均７のポワッソン $\mathrm{P}[\,\{k\}\,] = \dfrac{7^k}{k!}e^{-7}$

(i)　１週間忘れ物がない確率は $\mathrm{P}[\,\{0\}\,] = e^{-7} \fallingdotseq 0.0009$

注：１日平均１件なので１日あたり忘れ物０の確率 e^{-1}．異なる日で独立だと７日間０の確率は e^{-7} で，答え一致←ポワッソン分布に従う独立な確率変数の和は平均を和とするポワッソン分布に従う

(ii)　$\mathrm{P}[\,\{6\}\,] = \dfrac{7^5}{5!}e^{-7} \times \dfrac{7}{6} > \mathrm{P}[\,\{5\}\,]$，$\mathrm{P}[\,\{7\}\,] = \dfrac{7^8}{8!}e^{-7} \times \dfrac{8}{7} > \mathrm{P}[\,\{8\}\,]$，

などから，$k = 6, 7$ のとき $\mathrm{P}[\,\{k\}\,] = \dfrac{7^6}{6!}e^{-7}$ で最大．

ポワッソン分布のグラフ，統計的検定の例題，は
「統計と確率の基礎 第３版」６章§２．例題２ を参照．

まとめ

$\Omega = \mathbb{Z}_+$（または，$\mathbb{Z}_+$ で番号付けられる集合），の上の確率測度とは，全測度 1 非負値集合関数で **σ 加法性**をみたすもの．

記号：$\bigcup_{i=1}^{\infty} A_i$, $\bigcap_{i=1}^{\infty} A_i$, $\exists$, $\forall$

可算個の集合算：$\bigcap_{i=1}^{\infty} A_i \subset A_n \subset \bigcup_{i=1}^{\infty} A_i$，結合／分配／ド・モルガンの法則

確率測度の性質：有限加法的確率測度の性質，**連続性**，**劣加法性**

$\{a_k \mid k \in \mathbb{Z}_+\}$ が非負実数列で $\sum_{k=1}^{\infty} a_k = 1$ を満たせば，根元事象の確率 $\mathrm{P}[\,\{k\}\,] = a_k$, $k \in \mathbb{Z}_+$，と $\mathrm{P}[\,A\,] = \sum_{k \in A} \mathrm{P}[\,\{k\}\,]$ で確率測度が決まる．

ポワッソン分布：$\mathrm{P}[\,\{k\}\,] = e^{-\lambda} \dfrac{\lambda^k}{k!}$. $m = v = \lambda$

９．強運の持ち主

- ポワッソン分布の性質と意味
- 少数の法則
- 非負整数上の分布の他の例
- 幾何分布，負の２項分布（パスカル分布）

$\mathbb{Z}_+$ 上の分布に慣れるために具体例の性質に少し立ち入ります

無限集合上の分布を考える必要性

問9–1：$0 < p < 1$とする．表の出る確率がpの硬貨を初めて表が出るまで投げるとき，裏が出た回数の合計の期待値と分散を求めよ．◇

問9–2：aを自然数とし，$0 < p < 1$とする．表の出る確率がpの硬貨を表がa回出るまで投げるとき，裏が出た回数の合計の期待値と分散を求めよ．◇

裏が何回でも続けて出る可能性がある＝全ての自然数kに対して$\mathrm{P}[\,\{k\}\,] > 0$である．

- 自然数は無数にあるから，有限集合上の確率では正確に表せない．
- 無限和は級数の和として数学的に有限の手続きで定義する．

解析：素朴には無限の手続き（例：無限和）を要する内容を有限の手続きで定義．
極限：適切な「無限個の」不等式をもって等式と定義．

ポワッソン分布

（再掲）ポワッソン分布：$\lambda > 0$ に対して，$\mathrm{P}[\,\{k\}\,] = e^{-\lambda}\dfrac{\lambda^k}{k!},\ k \in \mathbb{Z}_+$，と $\mathrm{P}[\,A\,] = \sum_{k\in A} \mathrm{P}[\,\{k\}\,]$ で決まる $\Omega = \mathbb{Z}_+$ 上の確率

２項分布の極限から，「１要因あたり確率が目に見えない程小さいが要因が独立に多数ある結果，平均が目に見える大きさの分布」と思える
例：１交差点の１日の事故確率は極めて小さいが交差点数が多いので，ある管内の交通事故が平均（例えば）１日１件．実際は事故が起きない日も２件以上起きる日もあり，その何日もにわたる分布をとれば，ポワッソン分布に近いと思われる．

例２：宝くじフリークの集団の一人あたりの購入枚数，人数，人 i の当たりくじ，を交差点の数，日数，調査 i 日目の事故数，と対応させると，当たり数の集団における分布はポワッソン分布に近いだろう．特に，大勢の中には『くじ運の強い人』もいるが，その人の特性とは言えない（くじ運の強い人＝分布の裾野という偶然）．

ポワッソン分布の平均と分散

（再掲）$\lambda > 0$. $\mathsf{P}[\,\{k\}\,] = e^{-\lambda}\dfrac{\lambda^k}{k!}$, $k \in \mathbb{Z}_+$. $\displaystyle\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{\lambda}$ （微分積分既出）

- 確率になること：$\mathsf{P}[\,\Omega\,] = \displaystyle\sum_{k=0}^{\infty}\mathsf{P}[\,\{k\}\,] = \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = 1$

（根元事象に対して非負値で定義されていて全測度1なので，$\mathsf{P}[\,A\,] = \displaystyle\sum_{k\in A}\mathsf{P}[\,\{k\}\,]$ で確率測度 $\mathsf{P}:\ \mathcal{F} \to \mathbb{R}$ が決まる）

平均 $m = \displaystyle\sum_{k=0}^{\infty} k\,e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^k}{(k-1)!} = e^{-\lambda}\sum_{\ell=0}^{\infty}\frac{\lambda^{\ell+1}}{\ell!} = \lambda e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}$
$= \lambda$

分散 $v = \displaystyle\sum_{k=0}^{\infty} k^2\,e^{-\lambda}\frac{\lambda^k}{k!} - m^2 = \sum_{k=0}^{\infty} k(k-1)\,e^{-\lambda}\frac{\lambda^k}{k!} + \sum_{k=0}^{\infty} k\,e^{-\lambda}\frac{\lambda^k}{k!} - m^2 =$
$\displaystyle e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^k}{(k-2)!} + m - m^2 = \lambda^2 e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} + m - m^2 = \lambda^2 + \lambda - \lambda^2 \ = \lambda$

少数の法則

２項分布で平均 $\lambda = n\,p_n$ を固定して $n \to \infty$ とすると（根元事象の確率について）ポワッソン分布に収束（既出）

定理（ポワッソンの少数の法則）：$i = 1, 2, \ldots, k_n$，$n = 1, 2, \ldots,$ に対して $X_{n,i}$ は確率 $p_{n,i}$ で１（表・成功），それ以外は０なる独立確率変数，$S_n = X_{n,1} + \cdots + X_{n,k_n}$（表・成功の回数）とする．

$\lim\limits_{n\to\infty} \sum\limits_{i=1}^{k_n} p_{n,i} = \lambda$，$\lim\limits_{n\to\infty} \max\limits_{1\leqq i\leqq k_n} p_{n,i} = 0$，$\lim\limits_{n\to\infty} k_n = \infty$ を満たすならば，

S_n の分布はポワッソン分布に収束する．

すなわち，任意の $A \subset \mathbb{Z}_+$ に対して $\lim\limits_{n\to\infty} \mathrm{P}[\, S_n \in A \,] = \sum\limits_{k\in A} e^{-\lambda}\dfrac{\lambda^k}{k!}.$ ◇

・「１要因あたり確率が目に見えない程小さいが要因が独立に多数ある結果，平均が目に見える大きさの分布」と思える $k_n = n$，$p_{n,i} = \lambda/n$ の場合が２項分布の極限．

もう少し一般化すると同値条件も得るが，煩雑なので省略

$\mathbb{Z}_+$上の分布（確率測度）の例

ポワッソン分布（既出）：$\mathsf{P}[\,\{k\}\,] = e^{-\lambda}\dfrac{\lambda^k}{k!},\ k \in \mathbb{Z}_+,$

と$\mathsf{P}[\,A\,] = \displaystyle\sum_{k\in A} \mathsf{P}[\,\{k\}\,]$（以下同様）

幾何分布：$\mathsf{P}[\,\{k\}\,] = p(1-p)^k,\ k \in \mathbb{Z}_+$

負の２項分布：$a \in \mathbb{N}$．$\mathsf{P}[\,\{k\}\,] = {}_{a+k-1}\mathsf{C}_k p^a (1-p)^k,\ k \in \mathbb{Z}_+$

幾何分布とパスカル分布

幾何分布：$0 < p < 1$ とする．$\mathsf{P}[\,\{k\}\,] = Ge_p(\{k\}) = p\,(1-p)^k$

$\displaystyle\sum_{k=0}^{\infty} p\,(1-p)^k = \frac{p}{1-(1-p)} = 1$ なので $\mathbb{Z}_+$ 上の確率が定まる □

表の確率 p の硬貨で，初めて表が出るまでに裏の出る回数の分布

負の２項分布（パスカル分布）：$0 < p < 1$，$a \in \mathbb{N}$ とする．

$\mathsf{P}[\,\{k\}\,] = {}_{a+k-1}\mathsf{C}_k\, p^a\,(1-p)^k,\ k \in \mathbb{Z}_+.$

$x = c$ の周りのテイラー展開 $(1-x)^{-a} = \displaystyle\sum_{k=0}^{\infty} {}_{a+k-1}\mathsf{C}_k \frac{(x-c)^k}{(1-c)^{a+k}}$ において，

$x = 0,\ c = 1 - p^{-1}$ とおくと $\displaystyle\sum_{k=0}^{\infty} \mathsf{P}[\,\{k\}\,] = 1$ を得るので確率を定める． □

表が a 回出るまでに裏の出る回数の分布

$a + k - 1$ 回のうち $a - 1$ 回表，k 回裏で，最後が表だから □

級数と$\mathbb{Z}_+$上の分布

収束する非負項級数→$\mathbb{Z}_+$上の確率測度が作れる（第８回）

←マクローリン展開 $f(x)=\sum_{k=0}^{\infty}\frac{x^k}{k!}f^{(k)}(0)$

- **指数関数の展開** $f(x)=e^x=\sum_{k=0}^{\infty}\frac{x^k}{k!}=1+x+\frac{1}{2}x^2+\frac{1}{6}x^3+\cdots$

→ポワッソン分布： $\mathrm{P}[\,\{k\}\,]=e^{-\lambda}\frac{\lambda^k}{k!},\ k\in\mathbb{Z}_+$

- **等比級数** $f(x)=\frac{1}{1-x}=\sum_{k=0}^{\infty}x^k\ ;\quad x=1-p$

→幾何分布： $\mathrm{P}[\,\{k\}\,]=p(1-p)^k,\ k\in\mathbb{Z}_+$

- **等比級数の**$a=2,3,\ldots$**への拡張** $f(x)=\frac{1}{(1-x)^a}=\cdots$

→負の２項分布： $\mathrm{P}[\,\{k\}\,]={}_{a+k-1}\mathrm{C}_k p^a(1-p)^k,\ k\in\mathbb{Z}_+$

問へのコメント

問9–1：答えは「統計と確率の基礎 第3版」練習問題2問4を参照．ただし，計算はテキストのとおりやる必要はなく，平均と分散の定義に戻って計算すればよい．定義に基づくと，求める回数は幾何分布 Ge_p に従うので，期待値 m と分散 v は Ge_p のそれぞれ平均と分散

$$m = \sum_{k=0}^{\infty} k\,p\,(1-p)^k,\ v = \sum_{k=0}^{\infty} (k-m)^2\,p\,(1-p)^k.$$

問9–2：答えは「統計と確率の基礎 第3版」練習問題1問4を参照．ただし，計算はテキストのとおりやる必要はなく，平均と分散の定義に戻って計算すればよい．定義に基づくと，求める回数は負の2項分布に従うので，期待値 m と分散 v は

$$m = \sum_{k=0}^{\infty} k\,{}_{a+k-1}\mathrm{C}_k p^a (1-p)^k,\ v = \sum_{k=0}^{\infty} (k-m)^2\,{}_{a+k-1}\mathrm{C}_k p^a (1-p)^k.$$

まとめ

ポワッソン分布：$\lambda > 0,\ \mathrm{P}[\,\{k\}\,] = e^{-\lambda}\dfrac{\lambda^k}{k!},\ m = v = \lambda$

『くじ運の強い人』は，その人や環境の特性ではなく，ランダムならば『くじ運の強い人』という偶然が出るものである

少数の法則（特に，2項分布の極限）「1要因あたり確率が目に見えない程小さいが要因が独立に多数ある結果，平均が目に見える」

非負整数上の分布のその他の例：

幾何分布（表が出るまでに裏の出る回数）：$\mathrm{P}[\,\{k\}\,] = p(1-p)^k$

負の2項分布（表がa回出るまでに裏の出る回数の分布）：

$\mathrm{P}[\,\{k\}\,] = {}_{a+k-1}\mathrm{C}_k p^a (1-p)^k.$

１０．バスはなぜ来ないか

- ポワッソン確率過程とポワッソンのランダムな測度
- ポワッソン分布と指数分布

ポワッソン分布の性質の理解のため，数学的に一段深い話の直感的・初等的な紹介をします

レポート３ ＠ keio.jp

ポワッソン分布とポワッソン過程

問 10–1：1ヶ月あたりの世界の航空事故は平均４のポワッソン過程にほぼ従うという．最後の事故がいまから１週間前にあったとき，いまから半月以内に事故のある確率を求めよ． ◇

ポワッソン分布の詳細性質の理解を助ける範囲でポワッソン過程を紹介

確率過程

確率過程：$\mathbb{Z}_+$ や $\mathbb{R}$ などの「時刻」パラメータを持つ確率変数の組 $X_\cdot :$ $\Omega \times \mathbb{R}_+ \to \mathbb{R}$（で，ある程度良い性質を持つもの）．

- 時刻 $t \in \mathbb{R}_+$ を固定する毎に確率変数，**サンプル（試行）**
- $\omega \in \Omega$ を固定する毎に t の関数．

確率過程は実確率変数のたば（前者の意味）＆関数に値を取る確率変数（後者の意味）

しかるべき性質を持った確率過程が存在することの数学的な証明は一般には講義の範囲をはるかに超える．

もっとも単純な例：単純ランダムウォーク　（秋）

ポワッソン過程

$\lambda > 0$．**強度** λ の**ポワッソン過程**：確率過程 $X_\cdot : \ \Omega \times \mathbb{R}_+ \to \mathbb{Z}_+$ で

i)　$\omega \in \Omega$ ごとに関数 $x = X_t(\omega)$ は t について増加

ii)　**独立増分**（$(a,b] \cap (c,d] = \emptyset$ のとき増分 $X_b - X_a$ と $X_d - X_c$ は独立）

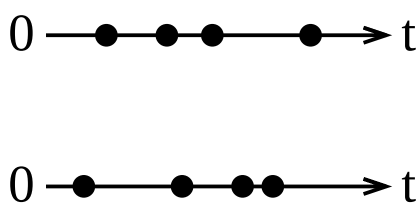
iii)　$b > a$ のとき $X_b - X_a$ の分布は平均 $(b-a)\lambda$ の**ポワッソン分布**

を満たすもの（X_t ＝時刻 t までの累積発生事件（「到着」）数）

注：1) 以上の性質を持つ確率過程（ポワッソン過程）が数学的に存在することが知られている，という意味

2) λ を時刻の関数 $\lambda(t)$ とすることができる（季節差，昼夜差）：$X_b - X_a$ の分布は平均 $\displaystyle\int_a^b \lambda(t)dt$ のポワッソン分布

ポワッソンのランダムな測度



$\mathbb{R}_+$（例：時間軸）上に単位長さ当たり平均 λ で**一様分布**する点（例：事件発生）：強度 λ のポワッソンのランダムな測度．（昼夜差のように平均（頻度）を時刻の関数にもできる.）

1 つのばらまき方（たくさんの点＝$\mathbb{R}_+$ 上の測度）が 1 つのサンプル（試行）．ばらまき方を全部集めた集合（測度の集合）が Ω：**ランダム測度**

区間 $(a, b]$（幅 $b-a$）にある点の数（事件件数）の分布 $\mu((a, b])$：「独立に無数の点を全体の平均 $(b-a)\lambda$ になる程度に置く」＝**平均 $(b-a)\lambda$ のポワッソン分布**

$[0, t]$ の点の数（時刻 t までの事件件数）$X_t = \mu([0, t])$
$X_\cdot :\ \Omega \times \mathbb{R}_+ \to \mathbb{Z}_+$：**ポワッソン過程**
・**ランダムな測度**から見ればポワッソン過程，ポワッソン分布，指数分布（次頁以降）の独立性に関する性質は自然

参考：服部哲弥「Amazon ランキングの謎を解く」化学同人出版

ポワッソン過程と指数分布

強度$\lambda > 0$のポワッソン過程でt時間事件が起きない確率

$= \mathrm{P}[\ X_t = 0\ ] = e^{-t\lambda}$ →**間隔は指数分布**

指数分布：$Q([a,b]) = \int_a^b \lambda e^{-t\lambda}\,dt$ で定まる$\mathbb{R}_+$上の分布

（ルベーグ）積分は非負値性と積分範囲に関する（σ）加法性があるので，測度を定義する　（秋）

$Q([0,t]^c) = Q((t,\infty)) = \left[-e^{-t\lambda}\right]_t^\infty = e^{-t\lambda}$ だから，ポワッソン過程で事件の間隔は指数分布

時刻tから$\triangle t$時間の間に最初の事件が起きる確率 $= Q([t, t+\triangle t])$ はほぼ，$\lambda e^{-t\lambda}\triangle t$．累積分布関数の微分（確率密度）は，その時刻での単位時間あたりの発生率

ポワッソン過程における指数分布の例

「ホテルで滞在日数を統計調査すると，出国時調査の倍の日数を得る」

出国時調査（全数調査）： $\tau_0 = \int_0^\infty t\,\lambda e^{-t\lambda}\,dt = \frac{1}{\lambda}$

ホテル調査（調査に出会う確率が滞在時間に比例）：

$$\tau = \frac{\int_0^\infty t \times a\,t\,\lambda e^{-t\lambda}\,dt}{\int_0^\infty a\,t\,\lambda e^{-t\lambda}\,dt} = \frac{1}{\lambda}\frac{\int_0^\infty t^2 e^{-t}\,dt}{\int_0^\infty te^{-t}\,dt} = \frac{1}{\lambda}\frac{\left[-t^2 e^{-t}\right]_0^\infty + 2\int_0^\infty te^{-t}\,dt}{\int_0^\infty te^{-t}\,dt} = \frac{2}{\lambda}$$

（滞在日数が指数分布に従うとすると説明が付く）

指数分布に関する待ち時間のパラドックス：

バスがランダム＝ポワッソン過程で到着すると，客の待ち時間の平均は，時刻表上の間隔 τ の半分ではなく τ

問10–1へのコメント

**問10–1の答えは「統計と確率の基礎 第３版」練習問題Ａ問４を参照．
そこでは，ポワッソン分布（正確には，ポワッソン過程）に従って事故が起こるとき，その発生時間間隔は指数分布に従うことと，指数分布の無記憶性を用いる．**
指数分布の無記憶性は，時刻t_0まで発生しなかった条件下で時刻$t > t_0$まで発生しない確率が，発生時刻（確率変数）をTとおくとき，

$$\mathrm{P}[\,T \geqq t \mid T \geqq t_0\,] = \frac{\mathrm{P}[\,T \geqq t\,]}{\mathrm{P}[\,T \geqq t_0\,]} = \frac{e^{-w\,t}}{e^{-w\,t_0}} = e^{-w\,(t-t_0)}$$

となって，時刻の原点をt_0に取り直す$(s = t - t_0)$と同じ頻度（平均値）wの指数分布になることを指す．

指数分布の無記憶性

指数分布の無記憶性は，背後にあるポワッソン過程（もしくはポワッソンのランダムな測度）において,「事件（到着）」が重ならない時間区間の間で独立に起こること（独立増分性）に由来

証明：ベイズの定理とポワッソン過程の独立増分性から，

時刻t_0まで起こらなかった条件下で時刻tまで起こらない確率

$$= \frac{\text{時刻}\, t_0\, \text{まで起こらなかった確率} \times \text{時間}\,(t_0, t]\,\text{に起こらない確率}}{\text{時刻}\, t_0\, \text{まで起こらなかった確率}}$$

＝t_0を時刻の原点として時刻tまで（時間$t - t_0$）に起こらない確率

まとめ

・**確率過程**：$X_\cdot: \ \Omega \times \mathbb{R}_+ \to \mathbb{Z}_+$ （関数をサンプルとする確率測度＝時刻をパラメータとする確率変数）

・**強度λのポワッソンのランダムな測度**．
幅cの区間にある点の数は平均$c\lambda$のポワッソン分布に従う．

・ポワッソン過程：時刻tまでの事件件数．
tを止めれば平均$t\lambda$の**ポワッソン分布**に従う．

・事件間隔は平均$1/\lambda$の**指数分布**：$Q([a,b]) = e^{-a\lambda} - e^{-b\lambda}$

・ポワッソンのランダムな測度の独立性
→ポワッソン過程の独立増分性
→指数分布の**無記憶性**
（少数の法則とともに，ポワッソン分布が$\mathbb{Z}$上の分布の中で特別な位置にある理由）

１１．保険が成り立つ理由

- 大数の法則
- 確率変数列の収束（概収束，確率収束）

事実としては広く知られていて，保険などで広く疑われず使われる原理である**大数の法則**の定式化（数学的に理解）を講義の水準の範囲で紹介します

ばらつく量を確定値で置き換えられる理由

問11–1：船乗りシンドバッドの時代に，船を持つ商人たちが，1航海あたり諸経費合計50単位金額を借りて国際貿易を行う．商品を買って無事持ち帰れば100単位の売上だが，独立に平均50航海中1回，積み荷を失って売上0になる．既に他の商売で金持ちだった商人たちは組合を作って，最初に十分多額の金額を出資した上で，1航海ごとにM単位の掛け金を払う一方，事故の際は組合から50単位の保険金を受け取って諸出費に当てることで，破産を免れるようにした．長い期間の後に平均的に出資金を取り崩さずにすむにはMをいくら以上にすれば良いか？50単位の経費は仕入れ・人件費から諸手数料まですべて込みとし，借金やプール金の年利は0とする． ◇

確率変数列の収束

- 確率変数列 $Z_1, Z_2, \ldots$ の確率変数 Y への収束

確率変数は関数．点の収束と違って関数の収束は実質的に異なる無数の基準がある．単純なのは各 $\omega \in \Omega$ ごとに数列の収束 $\lim_{n\to\infty} Z_n(\omega) = Y(\omega)$ ：各点収束（微分積分で関数の収束として既出）．確率論向きの基準は…

- **概収束**：確率１の $\omega \in \Omega$ に対して $\lim_{n\to\infty} Z_n(\omega) = Y(\omega)$

言い換え例:$\mathrm{P}[\ \lim_{n\to\infty} Z_n = Y\ ] = 1$，　記号:$Z_n \to Y$, a.e., $n \to \infty$

記号：$Z_n \to Y$, a.e.

- **確率収束**：全ての正数 ε に対して $\lim_{n\to\infty} \mathrm{P}[\ |Z_n - Y| \geqq \varepsilon\ ] = 0$

記号：$Z_n \to Y$, in prob.

独立確率変数列の大数の法則

大数の弱法則：　**独立な**確率変数列 $X_1, X_2, \ldots$ が分散が有界で，すなわち $\mathsf{V}[\,X_k\,] \leqq v$ が全ての k について同時に成り立つ正定数 v があり，期待値 $m_k = \mathsf{E}[\,X_k\,]$ の平均が収束する（$\displaystyle\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} m_k = m$ となる実定数 m がある）ならば，確率変数の**平均**は m に**確率収束**する：

$$(\forall\varepsilon > 0)\quad \lim_{n\to\infty} \mathsf{P}[\,|\frac{1}{n}\sum_{k=1}^{n} X_k - m| \geqq \varepsilon\,] = 0 \qquad\qquad \diamond$$

・大数の法則＝確率変数の算術平均 $\displaystyle\frac{1}{n}\sum_{k=1}^{n} X_k$ が（ω によらない数に）収束するという定理．

・通常は期待値がある場合のみ扱うので極限は期待値の極限

・大数の弱法則＝大数の法則の収束を確率収束で定式化した定理

大数の法則について

- $m_k = m$のとき，和の期待値はnに比例し，バラツキは相対的に緩やかにしか増えないことを大数の弱法則は意味する．航海毎に期待値より1単位でも多く掛ければ，バラツキを考慮しても確実に（確率1で）（途中で借金の可能性はあるが）プール金は増大，1単位でも少なければ確実に破綻
- このことから，保険の計算の基礎は，期待値（と，問11–1で無視した利率や諸経費）の計算であり，それにバラツキを少し加味
- 定数に収束することから，定額（保険料）で，予測不能な（ばらつきの大きい）事故や病気（損害保険，死亡保険）の出費に対処可能
- まれにしか起きないから1回あたりの掛け金は実際の損害より遥かに少ない（必ず起きるならば別の，社会的に難しい問題（老齢対策））

概収束と確率収束

- 大数の強法則＝（平均が期待値に）概収束すること
- 大数の弱法則＝（平均が期待値に）確率収束すること
- 概収束＝現実を「神が選んだ」１つのサンプルとみるとき「（神が我々だけに意地悪でない限り）収束すること」．確率０例外（我々だけに意地悪な神）は不生起（空集合）と区別できない．概収束はその理論的制約の中で最強

→理論の現実への応用上は素直な（目標としたい）収束

- 確率収束＝極限から一定以上乖離する事象の確率が０に収束
- 概収束ならば確率収束（後者は前者の必要条件だが十分条件でない）

- 確率収束するが概収束しないのは，いつまでも一定量の極限からのずれが確率正のサンプルで起きること．確率収束するならば，ずれる確率は０に収束．特殊な数理的構造が必要．すなおな数理モデルで確率収束を得たならば概収束も可能性がある．
- 逆に，実際は概収束するモデルにおいては，確率収束は研究上の中間点．結果を応用する立場ならば確率収束が分かった時点で利用できるが，確率収束の証明は概収束の証明に使えないほど単純なことが多いので，モデルの性質を理解するためにプロは概収束の証明まで挑戦する．

大数の強法則

- **大数の強法則**：$\mathrm{E}[\,|X_k|\,]$ が有限で $\mathrm{E}[\,X_k\,]=0$ である独立同分布確率変数列 $\{X_k\}$ について，$\displaystyle \frac{1}{n}\sum_{k=1}^{n} X_k \to 0$, a.e., $n\to\infty$. ◇
- 期待値 $\mathrm{E}[\,X_k\,]=0$ がある時点で，$\mathrm{E}[\,|X_k|\,]$ は有限（ルベーグ積分）が前提なので，仮定は（独立同分布の場合に）究極の一般形．
- 大数の法則と**独立性**：極限（期待値）がある前提では，独立確率変数列ならば（追加条件ほぼ無く）大数の強法則が成立＝**独立性**が大数の法則の（ほぼ）十分条件
- 独立でないときは定数に収束する場合もしない場合もある
- **積み残し**：**無限個**の独立確率変数列 ??

大数の法則は，1 つの確率空間 $(\Omega,\mathcal{F},\mathrm{P})$ 上に無限個の独立確率変数 X_k, $k\in\mathbb{N}$, がある時を扱う（なければ定理を証明しても無意味）．実は，ある　（秋）

大数の法則の仮定について

- 概収束すれば確率収束するので，前頁の大数の強法則が成り立つからには，同じ仮定で大数の弱法則も成り立つ．
- 特に，独立確率変数列の場合は，分散に対する仮定は大数の弱法則でも不要のはず．
- 今回冒頭で紹介した大数の弱法則（分散有界な独立確率変数列の場合）は，分散有界という強い前提をおく（すなおな振る舞いの確率変数列だけに考察の対象を限る）ことで，この講義の範囲でもわかる証明が可能
- 次回に冒頭で紹介した大数の弱法則の証明を紹介する予定

大数の法則と期待値とバラツキ — 問11–1の答え

(1) 航海あたり50拠出すれば十分（有界な確率変数）だが，船は無事帰ることが多いのでプール金は膨れあがって死蔵された無駄金となる．

(2) 掛け金は50より減らすべきで，問は直感的には当たり前の答え（期待値）を問う．今回のテーマはその数学的根拠（大数の法則）．

(3) 現実は出資金が有限なので，掛け金＝期待値だと，小さい確率で『黒鳥が飛んできて』立て続けに積み荷を失って組合が破産する（バラツキ，保険数理の安全割増）．

(3') 他方，掛け金≠期待値ならば長期にはその差の影響がバラツキの影響を圧倒（答えは知っていても，本当の理由は難しい）．

(4) 分散はバラツキの目安（既出）．精密には（ランダムウォークが原点0に当たる確率なので）確率連鎖の計算．さらに難しい（数理ファイナンス）．

(*) 大数の法則の別の視点：微視的に（自然現象では分子やそれ以下の小さい長さで，社会現象では個人などのバラツキのある行動について）見ると確率現象に見えるものを集計して巨視的に見る（個々の単位が見えない物差しで計り直す）と決定論的な現象に見える．「Amazon ランキングの謎を解く」

まとめ

- 大数の法則（弱法則＝確率収束，強法則＝概収束）：
  平均値は期待値に収束する
  平均のバラツキ（標準偏差）は０に近づく
  微視的には確率現象だが巨視的には決定論的に見える
  期待値があって有界な独立確率変数列について成立
  独立でないときは定数に収束する場合もしない場合もある

１２．確率変数列の収束

- 大数の弱法則の証明
- 確率論の不等式（チェビシェフの不等式）
- 概収束すれば確率収束すること
- やさしい場合の確率論の極限定理の証明の雰囲気を紹介します

大数の弱法則

問 12．分散有界な独立確率変数列の大数の弱法則を証明せよ．

大数の弱法則：独立な確率変数列 $X_1, X_2, \ldots$ が分散が有界で，すなわち $\mathrm{V}[\,X_k\,] \leqq v$ が全ての k について同時に成り立つ正定数 v があり，期待値 $m_k = \mathrm{E}[\,X_k\,]$ の平均が収束する（$\displaystyle\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} m_k = m$ となる実定数 m がある）ならば，確率変数の平均は m に確率収束する：

$$(\forall\varepsilon > 0)\quad \lim_{n\to\infty} \mathrm{P}[\,|\frac{1}{n}\sum_{k=1}^{n} X_k - m| \geqq \varepsilon\,] = 0 \qquad \diamond$$

チェビシェフの不等式

記号：確率変数 X と事象 $A \in \mathcal{F}$ に対して，$\mathsf{E}[\,X;A\,] = \mathsf{E}[\,X\,\mathbf{1}_A\,]$

例：$\mathsf{E}[\,X\,] = \mathsf{E}[\,X;A\,] + \mathsf{E}[\,X;A^c\,]$

命題（チェビシェフの不等式）：確率変数 $X:\ \Omega \to \mathbb{R}$ と非減少な非負値関数 $f:\ \mathbb{R} \to \mathbb{R}_+$ の合成関数 $f(X) = f \circ X$ と実数 a に対して

$\mathsf{E}[\,f(X)\,] \geqq f(a)\mathsf{P}[\,X \geqq a\,]$ ◇

証明：$A = \{X \geqq a\}$ として例を使い，f の非減少非負性と期待値の単調性（非負性）**（既出）**も用いると

$\mathsf{E}[\,f(X)\,] = \mathsf{E}[\,f(X);X \geqq a\,] + \mathsf{E}[\,f(X);X < a\,]$

$\geqq f(a)\mathsf{E}[\,\mathbf{1}_{X \geqq a}\,] = f(a)\mathsf{P}[\,X \geqq a\,]$ □

大数の弱法則の証明

（再掲）定理：独立確率変数列 $\{X_k\}$ が，$\mathrm{V}[\,X_k\,] \leqq v$ を全ての k について同時に満たし $\displaystyle\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} m_k = m$ ならば $(\forall\varepsilon>0)$ $\displaystyle\lim_{n\to\infty}\mathrm{P}[\,|\frac{1}{n}\sum_{k=1}^{n} X_k - m| \geqq \varepsilon\,] = 0$ ◇

証明：$Y_k = X_k - m_k$ とおくと，$\mathrm{E}[\,Y_k\,] = 0$ で，期待値の平均が m に収束するから，$(\forall\varepsilon > 0)$ $\displaystyle\lim_{n\to\infty}\mathrm{P}[\,|\frac{1}{n}\sum_{k=1}^{n} Y_k| \geqq \varepsilon\,] = 0$ を言えばよい．つまり，m と m_k が全て 0 の場合を示せばよい．$f(x) = \begin{cases} x^2, & x \geqq 0 \\ 0, & x < 0 \end{cases}$ とおくと非減少非負値だからチェビシェフの不等式から，

$$\varepsilon^2\mathrm{P}[\,|\frac{1}{n}\sum_{k=1}^{n} Y_k| \geqq \varepsilon\,] \leqq \mathrm{E}[\,(\frac{1}{n}\sum_{k=1}^{n} Y_k)^2\,] = \frac{1}{n^2}\mathrm{E}[\,(\sum_{k=1}^{n} Y_k)^2\,].$$

（続く）

大数の弱法則の証明（続）

（再掲）$\varepsilon^2 \mathrm{P}[\,|\frac{1}{n}\sum_{k=1}^{n} Y_k| \geqq \varepsilon\,] \leqq \mathrm{E}[\,(\frac{1}{n}\sum_{k=1}^{n} Y_k)^2\,] = \frac{1}{n^2}\mathrm{E}[\,(\sum_{k=1}^{n} Y_k)^2\,]$.

右辺の期待値の中を展開し，独立確率変数の積の期待値は期待値の積に等しいことと $\mathrm{E}[\,Y_k\,] = 0$ を使うと右辺 $= \dfrac{1}{n^2}\displaystyle\sum_{k=1}^{n}\sum_{\ell=1}^{n} \mathrm{E}[\,Y_k Y_\ell\,] =$

$$\frac{1}{n^2}\sum_{k=1}^{n} \mathrm{E}[\,Y_k^2\,] \leqq \frac{v}{n} \to 0 \ \ (n \to \infty).$$

挟み撃ちの原理から $\varepsilon^2 \displaystyle\lim_{n\to\infty} \mathrm{P}[\,|\frac{1}{n}\sum_{k=1}^{n} Y_k| \geqq \varepsilon\,] = 0.$ □

概収束するならば確率収束することの証明

定理． 実確率変数列 $Z_1, Z_2, Z_3, \ldots$ が確率変数 Y に概収束すれば，確率収束する． ◇

系． ある仮定の下で大数の強法則が成り立てば，同じ仮定の下で大数の弱法則が成り立つ． ◇

（再掲，第 11 回）
概収束 $(Z_n \to Y,\ \text{a.e.})$：確率 1 の $\omega \in \Omega$ に対して $\lim_{n\to\infty} Z_n(\omega) = Y(\omega)$
確率収束 $(Z_n \to Y,\ \text{in prob.})$：全ての正数 ε に対して $\lim_{n\to\infty} \mathrm{P}[\, |Z_n - Y| \geqq \varepsilon \,] = 0$

証明． 概収束 $\Leftrightarrow$
$(\forall \varepsilon > 0)\ \exists n_0 \in \mathbb{N};\ (\forall n \geqq n_0)\ |Z_n(\omega) - Y(\omega)| < \varepsilon$ が成り立つサンプル $\omega \in \Omega$ の集合の確率が 1 である．集合で書くと，

全ての $\varepsilon > 0$ に対して $\displaystyle\bigcup_{n_0}^{\infty} \bigcap_{n=n_0}^{\infty} \{|Z_n - Y| < \varepsilon\}$ の確率が 1．

（続く）

概収束するならば確率収束することの証明（続）

（再掲）$\displaystyle\bigcup_{n_0}^{\infty}\bigcap_{n=n_0}^{\infty}\{|Z_n - Y| < \varepsilon\}$ の確率が 1.

補集合を取るとドモルガンの法則から $\displaystyle \mathrm{P}[\bigcap_{n_0}^{\infty}\bigcup_{n=n_0}^{\infty}\{|Z_n - Y| \geqq \varepsilon\}] = 0 \quad (*)$

$\displaystyle\bigcup_{n=n_0}^{\infty}\{|Z_n - Y| \geqq \varepsilon\}$ は n_0 について包含関係がある（n_0 が大きいと含まれる）ので，

確率の連続性から，

$$\mathrm{P}[\bigcap_{n_0}^{\infty}\bigcup_{n=n_0}^{\infty}\{|Z_n - Y| \geqq \varepsilon\}] = \lim_{n_0\to\infty}\mathrm{P}[\bigcup_{n=n_0}^{\infty}\{|Z_n - Y| \geqq \varepsilon\}] \quad (**)$$

$(*)$ と $(**)$ と確率の単調性と非負値性と挟み撃ちの原理から

$$\lim_{n_0\to\infty}\mathrm{P}[\{|Z_{n_0} - Y| \geqq \varepsilon\}] = 0$$

これが全ての $\varepsilon > 0$ に対して成り立つので Z_n は Y に確率収束する. □

大数の強法則の証明について

- 　１１回で紹介した「究極の」大数の強法則の証明：多く紹介される方法：マルチンゲール型不等式，初等方法：大きい値を0に置いた補助的確率変数．いずれかにボレル・カンテリの定理を組み合わせる．この講義では省略．
- 　概収束の当否を確かめるために，強い条件をつけて（ややこしい振る舞いの確率変数列は除外して）簡単な証明で概収束を証明することもある
- 　**大数の強法則**：独立同分布確率変数列$\{X_k\}$の分布の平均が0で4次モーメント（既出）が有限$(\mathrm{E}[\,X_k^4\,] < \infty)$ならば$\frac{1}{n}\sum_{k=1}^{n} X_k \to 0$, a.e., $n \to \infty$. ◇
- 　証明はこの講義の知識でも本質は理解可能（「統計と確率の基礎 第３版」第４章章末補足定理10参照）

まとめ

- 分散（２次モーメント）有限の場合の独立同分布確率変数列の大数の弱法則の証明
- チェビシェフの不等式
- 記号：$\mathsf{E}[\,X;A\,]=\mathsf{E}[\,X\,1_A\,]$, $\mathsf{E}[\,X\,]=\mathsf{E}[\,X;A\,]+\mathsf{E}[\,X;A^c\,]$
- 概収束すれば確率収束することの証明

＊．秋のための重積分

確率論入門I（春学期）：有限集合と$\mathbb{Z}_+$上の確率測度

・（秋）無限列の集合の上の確率は根元事象の確率の和で書けるとは限らない

・（秋）実数の集合の上の確率は根元事象の確率の和で書けるとは限らない

（数学的には非可算無限集合の上の確率測度というくくり）

確率測度：σ加法性と非負値性を持ち$\mathrm{P}[\,\Omega\,]=1$を満たす集合関数

公理的確率論（コルモゴロフ）

ガウス積分と正規分布

実数上の確率は，統計学や計量経済学を含め，実用上頻繁に現れる．
→　避けて通れない
数列上の確率（ランダムウォーク）やその「連続極限」であるブラウン運動は数理ファイナンスを含め，実用上頻繁に現れる．　→　避けて通れない

例：正規分布

$$z = \int_0^{\infty} e^{-\frac{1}{2}x^2}\, dx = \lim_{R\to\infty} \int_0^{R} e^{-\frac{1}{2}x^2}\, dx$$

$\rho_{0,1}(x) = \dfrac{1}{2z} e^{-\frac{1}{2}x^2}$ **は$\mathbb{R}$上の確率測度（標準正規分布）**$N = N(0,1)$**の密度関数**

平均m，分散vの正規分布（ガウス分布）$N(m,v)$**の密度関数：**

$$\rho_{m,v}(x) = \frac{1}{2z\sqrt{v}} e^{-\frac{1}{2v}(x-m)^2}$$

積分の変数変換１

正規分布の話以降で，あちこちで使うので．

$x' = x - c$: $\displaystyle\int_a^b f(x-c)\,dx = \int_{a-c}^{b-c} f(x')\,dx'$

例： $\displaystyle\int_a^b (x-2)^3 dx = \int_{a-2}^{b-2} x'^3 dx' = \frac{1}{4}\left[x'^4\right]_{a-2}^{b-2} = \frac{1}{4}\Big((b-2)^4 - (a-2)^4\Big)$

たしかめ：合成関数の微分 $((x-2)^4)' = 4(x-2)^3$ **から，** $\displaystyle\int_a^b (x-2)^3 dx = \frac{1}{4}\left[(x-2)^4\right]_a^b = \frac{1}{4}\Big((b-2)^4 - (a-2)^4\Big)$ **となって成り立つ**

例（ガウス積分）： $\displaystyle\int_{-\infty}^{\infty} x^2 e^{-(x-m)^2/2} dx = \int_{-\infty}^{\infty} (x'+m)^2 e^{-x'^2/2} dx'$

$\displaystyle = \int_{-\infty}^{\infty} (x'^2 + 2mx' + m^2) e^{-x'^2/2} dx' = \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx + m^2 \int_{-\infty}^{\infty} e^{-x^2/2} dx$

『$\infty \pm m = \infty$』**（極限の意味で理解）**

- **重積分でも同様**

積分の変数変換２

$x' = c\,x$: $\displaystyle\int_a^b f(c\,x)\,dx = \int_{ac}^{bc} f(x')\frac{1}{c}\,dx'$

例：$\displaystyle\int_a^b e^{2x}dx = \frac{1}{2}\int_{2a}^{2b} e^{x'}dx' = \frac{1}{2}\left(e^{2b} - e^{2a}\right)$

たしかめ：合成関数の微分 $(e^{2x})' = 2e^{2x}$ から，

$\displaystyle\int_a^b e^{2x}dx = \frac{1}{2}\left[e^{2x}\right]_a^b = \frac{1}{2}\left(e^{2b} - e^{2a}\right)$ となって成り立つ

例：$\displaystyle\int_{-\infty}^{\infty} x^2 e^{-2x^2}dx = \int_{-\infty}^{\infty} (x'/2)^2 e^{-x'^2/2}dx' = \frac{1}{4}\int_{-\infty}^{\infty} x^2 e^{-x^2/2}dx$

『$2\infty = \infty$』（極限の意味で理解）

積分の変数変換３

・重積分では積分変数変換のヤコビ行列式
（１年微分積分の最後でいちおうやったことになっている…）

$$dx\,dy = \left\| \begin{matrix} \dfrac{\partial\, x}{\partial x'} & \dfrac{\partial\, x}{\partial y'} \\ \dfrac{\partial\, y}{\partial x'} & \dfrac{\partial\, y}{\partial y'} \end{matrix} \right\| dx'\,dy'$$

$\displaystyle\int_a^b f(x)\,dx = \frac{1}{c}\int_{ac}^{bc} f(x'/c)\,dx'\ \ (x' = c\,x)$ の，線形代数的にすなおな，拡張！

直交変換：直交行列 O（${}^tOO = E$，E は単位行列）に対して $\overrightarrow{x} = O\,\overrightarrow{x'}$ なる変数変換のときは，$|O| = \pm 1$ なので，$dx_1 \ldots dx_n = dy_1 \ldots dy_n$（ヤコビ行列式のことは忘れることができる）

・以上の復習は，秋学期で使います（春の定期試験の範囲ではありません）

極座標とガウス積分

・**積分変数変換の例：**

$x = r\cos\theta,\ y = r\cos\theta$

$$dx\,dy = \left\| \begin{matrix} \dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial \theta} \\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial \theta} \end{matrix} \right\| dr\,d\theta = r\,dr\,d\theta$$

・**例：**$z = \displaystyle\int_0^\infty e^{-x^2/2}dx$

$$z^2 = \int_0^\infty e^{-x^2/2}dx \int_0^\infty e^{-y^2/2}dy$$

$$= \int_{\text{第１象限}} e^{-(x^2+y^2)/2}dx\,dy$$

$$= \int_{r\in[0,\infty)}\int_{\theta\in[0,\pi/2]} e^{-r^2/2}r\,dr\,d\theta = \frac{\pi}{2}$$

$$z = \frac{1}{2}\sqrt{2\pi}$$

教科書・参考書

服部哲弥，統計と確率の基礎 第３版，学術図書，2014

参考書

服部哲弥，Amazon **ランキングの謎を解く，化学同人，**2011

服部哲弥，確率変数の収束と大数の完全法則，共立出版，2019