

確率の基礎

滝根 哲哉*

目次

1	確率空間	1
1.1	標本空間と事象	1
1.2	確率分布	3
1.3	確率変数とその分布	9
1.4	確率ベクトルと結合分布, 周辺分布	10
1.5	確率変数を引数とする関数と確率変数の畳み込み	12
2	期待値	15
2.1	期待値の定義	15
2.2	積率, 分散, 共分散	19
3	確率変数と分布の収束	23
3.1	確率変数列の概収束と確率収束	23
3.2	大数の法則	25
3.3	分布の弱収束と中心極限定理	27
4	高校までの確率の復習用問題	28
4.1	基本問題	28
4.2	練習問題	29
A	Riemann 積分と Riemann-Stieltjes 積分について	30
A.1	Riemann 積分	30
A.2	Riemann-Stieltjes 積分	32
B	定理の証明	32
B.1	大数の弱法則の証明	32
C	練習問題の略解	34

注意：高校までに習う確率に関しては理解していることを前提に授業を行う。高校レベルの確率の練習問題（標準レベルの大学入試問題）を 4 章に付けているので，不安がある学生はこれで自習しておくこと。

1 確率空間

1.1 標本空間と事象

確率モデル（あるいは確率的実験）は以下の 3 つの要素から構成されていると考えることができる。

- 標本空間 (sample space): 起こり得る結果, すなわち標本 (sample) の集合であり, 実験が行われるところの中の一つだけが実際に起こる
- 事象 (event) の集合: それぞれの事象は標本空間の部分集合
- それぞれの事象へ確率を割り当てる規則: 確率は 0 以上 1 以下の実数

*大阪大学大学院工学研究科電気電子情報工学専攻 (〒565-0871 吹田市山田丘 2-1)
 電話: (06)6879-7740 FAX: (06)6875-5901 電子メール: takine@comm.eng.osaka-u.ac.jp
 URL: <http://www2b.comm.eng.osaka-u.ac.jp/~takine/>

注意 1.1 標本と事象の違いには特に注意を払う必要がある．標本とは実験の結果そのものであり，一つの実験に対して単一の標本が対応する．一方，事象は，一般には，実験の結果を部分的に特徴付けるものであり，実験の結果が（その事象に対応する）標本空間の部分集合に含まれていることを示す． □

実験の最も単純な例は標本の数有限か，あるいは可算無限個（個々の標本に対して $\omega_1, \omega_2, \dots$ というふうに番号をつけることができる）の場合である．この場合は，それぞれの標本に対応する事象を考えることで，標本そのものに対して確率を割り当てることができる．また，このとき，任意の事象の確率は，その事象を構成する標本の確率を全て足し合わせたもので与えられる．さらに全ての標本に割り当てられた確率の総和は 1 である．

例 1.1 コインを 2 回投げた実験を考える．ここで標本空間は四つの標本 $\{(表, 表), (表, 裏), (裏, 表), (裏, 裏)\}$ から成るとし，それぞれの標本に $1/4$ の確率を割り当てる．標本空間の部分集合はそれぞれ事象である．例えば「最初に表が出る」という事象は $\{(表, 表), (表, 裏)\}$ であり，確率 $1/4 + 1/4 = 1/2$ をもつ． □

注意 1.2 標本空間を上記のように選ぶということは，コインが立つ，あるいはコインが 2 回目に投げられる前になくなってしまうといった現象を無視したことを意味する．また，確率を上記のように選ぶということはコインに歪みがあるという可能性を排除したことになる．このように，実際の物理現象と確率モデルの違いに注意しなければならない． □

例 1.1 をみると事象という概念を持ち出す必要はなく，個々の標本に確率を割り当てれば十分と思えるかも知れない．しかし，次の例では，個々の標本に確率を割り当てるということは意味がなく，なぜ，事象という概念を導入しなければならないかが明らかになる．

例 1.2 $[0, 1]$ の実数からなる標本空間を考え，実験から得られる結果は区間 $[0, 1]$ 内で一様に分布しているとす¹．もし，ある標本に正の確率 $p > 0$ が割り当てられているならば，一様に分布しているという仮定より，他の標本も同じ確率 p を持つことになるが，一方，標本の数は無数にあるので，確率の総和は $p \times \infty = \infty$ となってしまう矛盾する．よって各標本に割り当てられる確率は 0 でなければならない．しかし，確率 0 をもつ標本を幾つ足し合わせても確率の総和は 0 にしかならず，個々の事象に対する確率（例えば標本が区間 $[0, 1/2]$ 内に入る確率）を得る合理的な方法はない． □

例 1.2 の場合，任意の区間に対して，区間長に等しい確率を割り当て，互いに交わらない区間の和に対応する確率は，それぞれの区間の確率の和で表すことが合理的であると思われる．例えば標本が区間 $[0, 1/4]$ 内に入る確率は $1/4$ であり，区間 $[0, 1/4]$ あるいは区間 $[3/4, 1]$ のいずれかに含まれる確率は $1/4 + 1/4 = 1/2$ である．また，区間 $[0, 1/2]$ に含まれる確率は，区間 $[0, 1/4]$ あるいは区間 $[1/4, 1/2]$ のいずれかに含まれる確率に等しいはずだが，区間長に等しい確率が割り当てられていれば矛盾は生じない．

この例から分かるように，確率は，一般には，標本ではなく事象に対して割り当てなければならない．さらに，割り当てられた確率は非負の値であり，標本空間全体に対しては 1 であり，かつ，任意の互いに交わらない事象の列 E_1, E_2, \dots に対してこれらの事象の和集合の確率はそれぞれの事象の確率の和で与えられるようなものでなくては我々が通常の生活で使っている「確率」という概念と一致しない．

標本空間が可算の場合，上の議論は事象に対する確率を事象に含まれる個々の標本に対する確率の和として定義できることを示している．しかし，より一般的な例 1.2 のような場合においては，全ての事象に対して確率が決定できるようにするためには，どのような事象に確率を割り当てておけば十分かという問題に直面することになる．これを解決するため，事象は以下のようにして定義される．

定義 1.1 (事象と事象の集合) 標本空間 Ω の部分集合の集まり \mathcal{F} が，条件

1. $\Omega \in \mathcal{F}$ (標本空間 Ω は事象である)
2. $A \in \mathcal{F} \Rightarrow A^c = \Omega - A \in \mathcal{F}$ (A が事象ならば A の余事象 A^c も事象である)
3. $A_i \in \mathcal{F} (i = 1, 2, \dots) \Rightarrow \cup_{i=1}^{\infty} A_i \in \mathcal{F}$ (事象の無限列に対してそれらの和事象も事象となる²)

¹一様に分布しているとは取りうるどの値も同様に確からしいということである．

² $\cup_{i=1}^{\infty} A_i = \{\omega \mid \text{ある自然数 } i \text{ が存在して } \omega \in A_i\}$ ．

を満たすとき, \mathcal{F} を Ω に対する事象の集合といい, \mathcal{F} の要素を事象と呼ぶ. □

直観的には, 標本空間 Ω を分割して得られる集合が事象であり, それらを集めたものが事象の集合 \mathcal{F} である. 定義より, 以下のことが成り立つ.

補題 1.1 (事象の性質)

1. $\phi \in \mathcal{F}$ (空集合 ϕ も事象である)
2. $A, B \in \mathcal{F} \Rightarrow A \cup B \in \mathcal{F}, A \cap B \in \mathcal{F}, A - B \in \mathcal{F}$ (事象 A, B の和事象 $A \cup B$, 積事象 $A \cap B$, 差事象 $A - B$ も事象である)
3. $A_i \in \mathcal{F} (i = 1, 2, \dots) \Rightarrow \bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$ (事象の無限列に対してそれらの積事象も事象となる³)

□

1. $\phi = \Omega^c \in \mathcal{F}$.
2. $A_1 = A, A_2 = B, A_i = \phi (i \geq 3)$ で 3. を適用. $A \cap B = (A^c \cup B^c)^c$. $A - B = A \cap B^c$.
3. $\bigcap_{i=1}^{\infty} A_i = (\bigcup_{i=1}^{\infty} A_i^c)^c$.

標本空間 Ω に対する事象の集合は一意ではない. 例えば, $\mathcal{F}_1 = \{\phi, \Omega\}$ や, ある $A \in \Omega$ に対して $\mathcal{F}_2 = \{\phi, A, A^c, \Omega\}$, あるいは $\mathcal{F}_3 = \{\Omega \text{ の全ての部分集合}\}$ は全て標本空間 Ω に対する事象の集合である. 以下では

$$2^\Omega = \{\Omega \text{ の全ての部分集合}\}$$

という記号を用いる. 全ての部分集合には空集合 ϕ と全体集合 Ω も含まれていることに注意する.

標本空間の中で応用上特に重要なものに, 有限あるいは可算無限個の標本からなる Ω と, 実数全体 ($= \mathcal{R}$) からなる Ω がある. 有限あるいは可算無限個の標本からなる Ω に対しては, 通常, $\mathcal{F} = 2^\Omega$ とする. 一方, 実数全体 \mathcal{R} からなる Ω に対しては

$\mathcal{F} = \mathcal{B}(\mathcal{R}) =$ 全ての半開区間 $(a, b]$ を要素にもち, かつ, 定義 1.1 の性質をみたす集合

とする. このように事象の集合を選ぶことで, 確率を矛盾なく定義することができることが知られている. この事象の集合 $\mathcal{B}(\mathcal{R})$ には 1 点集合 $\{a\}$, 开区間 (a, b) , 任意の閉区間 $[a, b]$ や (a, ∞) , $(-\infty, b)$, $[a, \infty)$, $(-\infty, b]$ が全て要素として含まれる.

$$\{a\} = \bigcap_{n=1}^{\infty} (a - 1/n, a] \in \mathcal{B}(\mathcal{R}), (a, b) = \bigcup_{n=1}^{\infty} (a, b - 1/n) \in \mathcal{B}(\mathcal{R}), [a, b] = \bigcap_{n=1}^{\infty} (a - 1/n, b] \in \mathcal{B}(\mathcal{R}) \text{ etc}$$

1.2 確率分布

次に標本空間と事象の集合の組 (Ω, \mathcal{F}) が与えられたとき, 各事象に対して確率分布を定義する. 以前に述べたように確率は以下の三つの公理を満足する必要がある.

定義 1.2 (確率の公理) 以下の性質をもつ, \mathcal{F} 上で定義された実数値関数 P は確率分布 (probability distribution) と呼ばれる⁴.

1. 全ての $A \in \mathcal{F}$ に対して $P(A) \geq 0$ (確率は非負の値を取る)
2. $P(\Omega) = 1$ (標本空間全体に対する確率は 1)
3. $A_i \in \mathcal{F} (i = 1, 2, \dots)$ かつ $A_i \cap A_j = \phi (i \neq j)$ ならば

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

が成立する (排反な事象の無限列の和事象の確率はそれぞれの事象の確率の和に等しい)

³ $\bigcap_{i=1}^{\infty} A_i = \{\omega \mid \text{全ての自然数 } i \text{ に対して } \omega \in A_i\}$.

⁴ P は Pr と書かれることもある.

□

確率の公理より以下が成立する .

補題 1.2 (確率の性質) 事象 $A, B \in \mathcal{F}$ に対して

1. $P(\phi) = 0$
2. $A \cap B = \phi$ ならば $P(A \cup B) = P(A) + P(B)$
3. $P(A^c) = 1 - P(A)$
4. $A \subset B$ ならば $P(A) \leq P(B)$

□

性質 1. は $A_i = \Omega, A_i = \phi (i = 2, 3, \dots)$ とすれば公理の 3. より $P(\Omega) = P(\Omega) + \sum_{i=2}^{\infty} P(\phi)$, すなわち $\sum_{i=2}^{\infty} P(\phi) = 0$. 公理の 1. より $P(\phi) \geq 0$ なので $P(\phi) = 0$. 性質 2. は $A_1 = A, A_2 = B, A_i = \phi (i = 3, 4, \dots)$ として公理の 3. と性質 1. を用いる . 性質 3. は $A_1 = A, A_2 = A^c, A_i = \phi (i = 3, 4, \dots)$ として公理の 2., 3. と性質 1. を用いる . 性質 4. は $B = A \cup (B - A)$ かつ $A \cap (B - A) = \phi$ に留意して性質 2. を適用し, 公理 1. を用いる .

$P(A)$ は事象 A の確率 (probability) と呼ばれる . 特に $P(A) = 1$ のとき, 事象 A は確率 1 で成り立つ (with probability 1, w.p.1) という . $P(A) = 1$ は $A = \Omega$ を意味しないことに注意する .

例 1.3 例 1.2 において A を標本が開区間 $(0,1)$ 内にある事象とすれば $P(A) = 1$ であるし, B を標本が閉区間 $[0,1]$ から高々可算個の点 x_1, x_2, \dots を取り除いた残りの区間に含まれる確率とすれば $P(B) = 1$ である . □

言い換えると, ある事象 C に対して $P(C) = 0$ であるからといって, その事象に含まれる標本 $\omega \in C$ が決して起こり得ないかということ, 必ずしもそうではない . たとえば, 例 1.2 において, 実験を一回行えば, 必ずある標本 $\omega \in [0, 1]$ が得られるが, $P(\{\omega\}) = 0$ である . このように標本空間が連続な場合, 我々が生活で用いている「確率」という概念はそのまま適用することが出来ない .

標本空間 Ω , 事象の集合 \mathcal{F} , 確率分布 (事象に確率を割り当てる規則) P の三つ組 (Ω, \mathcal{F}, P) を確率空間 (probability space) という . 確率空間は, 標本空間が有限または可算な場合 (離散型確率空間) とそうでない場合に大別することができる . 離散型確率空間の場合は, $\Omega = \{\omega_1, \omega_2, \dots\}$, $\mathcal{F} = 2^\Omega$ に対して, $p_i (i = 1, 2, \dots)$ を総和が 1 である非負の実数としたとき, 任意の事象 $A \in \mathcal{F}$ に対して

$$P(A) = \sum_{\omega_i \in A} p_i$$

とすれば, (Ω, \mathcal{F}, P) は確率空間となる . このような p_i は確率関数 (probability function, または probability mass function) と呼ばれる .

一方, 標本空間 Ω が実数全体の場合 ($\Omega = \mathcal{R}$), 明らかに, 各標本と確率を 1 対 1 に対応させる方法では確率空間を作り出すことは出来ない . このような場合, 確率空間は実数値関数を用いて構築される . 実数値関数 $F(x) (x \in \mathcal{R})$ が以下の性質をもつと仮定する .

定義 1.3 (分布関数の性質)

1. $\lim_{x \rightarrow -\infty} F(x) = 0$ ⁵
2. $F(x)$ は非減少 ($x < y \Rightarrow F(x) \leq F(y)$) かつ右連続 ($\lim_{\epsilon \rightarrow 0+} F(x + \epsilon) = F(x)$ ⁶)
3. $\lim_{x \rightarrow \infty} F(x) = 1$ ⁷

このような性質をもつ関数 $F(x)$ は分布関数 (distribution function) あるいは単に分布 (distribution) と呼ば

⁵任意の正数 ϵ に対して正数 δ が存在し, $x < -\delta$ ならば $|f(x) - a| < \epsilon$ となるとき $\lim_{x \rightarrow -\infty} f(x) = a$ と書く .

⁶ $\lim_{\epsilon \rightarrow 0+}$ は ϵ を正の側から a へ近付けることを意味する . よって, $\lim_{\epsilon \rightarrow 0+} f(x + \epsilon) = f(x)$ は, 任意の正数 ϵ に対して正数 δ が存在し, $y \in [x, x + \delta)$ ならば $|f(y) - f(x)| < \epsilon$ となるということである .

⁷任意の正数 ϵ に対して正数 δ が存在し, $x > \delta$ ならば $|f(x) - a| < \epsilon$ となるとき $\lim_{x \rightarrow \infty} f(x) = a$ と書く .

れる .

□

ここで , $\Omega = \mathcal{R}$, $\mathcal{F} = \mathcal{B}(\mathcal{R})$ とする . ある分布関数 $F(x)$ が与えられたとき ,

$$P((-\infty, x]) = F(x) \quad (1)$$

を満たすような確率分布 P が一意に定まることが知られている . よって , 標本空間が実数全体の場合 , 確率空間を定めることと分布関数を定めることは等価であり , 式 (1) を満たすように確率分布 P を定めると , $(\Omega, \mathcal{F}, P) = (\mathcal{R}, \mathcal{B}(\mathcal{R}), P)$ は確率空間となる . 式 (1) ならびに定義 1.2 で与えた確率の公理 3. より , 任意の a, b ($a < b$) に対して

$$F(b) = P((-\infty, b]) = P((-\infty, a] \cup (a, b]) = P((-\infty, a]) + P((a, b]) = F(a) + P((a, b])$$

が成立する . すなわち

$$P((a, b]) = F(b) - F(a), \quad a < b \quad (2)$$

となる .

ある分布関数 $F(x)$ が与えられたとき

$$\int_{-\infty}^x f(y)dy = F(x) \quad (3)$$

が全ての实数 x に対して成り立つような関数 $f(x)$ が存在すれば , $f(x)$ は $F(x)$ の密度関数 (density function) と呼ばれる . 定義より , 分布関数 $F(x)$ が微分可能ならば密度関数 $f(x)$ が存在し , $\epsilon > 0$ に対して

$$f(x) = \frac{d}{dx}F(x) = \lim_{\epsilon \rightarrow 0} \frac{F(x + \epsilon) - F(x)}{\epsilon} = \lim_{\epsilon \rightarrow 0} \frac{F(x) - F(x - \epsilon)}{\epsilon}$$

が成立する . 式 (2) より ,

$$F(x + \epsilon) - F(x) = P((x, x + \epsilon])$$

であり , $F(x + \epsilon)$ を x の周りでテーラー展開すると

$$F(x + \epsilon) = F(x) + \frac{d}{dx}F(x)\epsilon + \frac{d^2}{dx^2}F(x)\frac{\epsilon^2}{2} + \frac{d^3}{dx^3}F(x)\frac{\epsilon^3}{6} + \dots$$

となる . よって $f(x)$ は

$$P((x, x + \epsilon]) = f(x)\epsilon + o(\epsilon)$$

を満たす⁸ . $o(\epsilon)$ の定義から上式は

$$f(x)\epsilon = P((x, x + \epsilon]) + o(\epsilon)$$

と等価であることに注意する .

同様に ,

$$F(x) - F(x - \epsilon) = P((x - \epsilon, x])$$

であり , $F(x - \epsilon)$ を x の周りでテーラー展開すると

$$F(x - \epsilon) = F(x) + \frac{d}{dx}F(x) \cdot (-\epsilon) + \frac{d^2}{dx^2}F(x)\frac{(-\epsilon)^2}{2} + \dots$$

となる . よって $f(x)$ は

$$P((x - \epsilon, x]) = f(x)\epsilon + o(\epsilon)$$

すなわち

$$f(x)\epsilon = P((x - \epsilon, x]) + o(\epsilon)$$

も満たす .

⁸ $o(\epsilon)$ は $\lim_{\epsilon \rightarrow 0} o(\epsilon)/\epsilon = 0$ となる項を表す . この授業では , ϵ で級数展開したとき , ϵ に関する 2 次以上の項をまとめたものに対応する .

以上より, 十分小さな ϵ に対しては

$$f(x)\epsilon = f(x)\frac{\epsilon}{2} + f(x)\frac{\epsilon}{2} = P((x - \epsilon/2, x]) + P((x, x + \epsilon/2]) + o(\epsilon)$$

となるため

$$f(x)\epsilon = P((x - \epsilon/2, x + \epsilon/2]) + o(\epsilon) \quad (4)$$

とも解釈できる. また $F(\infty) = 1$ なので, 非負の値を取る関数 $f(x)$ が

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

を満たすならば, 式 (3) で定義された関数 $F(x)$ は $f(x)$ を密度関数にもつ分布関数となる.

問 1.1 例 1.2 の実験に対応する分布関数を求めよ. また, 密度関数はどのようなになるか.

標本空間が連続な場合と同様に, 標本空間 Ω が $\Omega \subset \mathcal{R}$ であるような離散型確率空間 $(\Omega, 2^\Omega, P)$ に対しても分布関数は定義できる. すなわち, 可算個の $x_1, x_2, \dots \in \Omega$, に対して確率関数 $p_i = P(x_i)$ が与えられたとき,

$$F(x) = \sum_{x_i \leq x} p_i$$

とすれば, $F(x)$ は分布関数となる.

問 1.2 $x_1 < x_2 < x_3$ に対して, $p_1 = p_2 = p_3 = 1/3$ のとき, 分布関数 $F(x)$ を求めよ.

上記では標本空間が 1 次元の分布関数を考えたが, これを多次元に拡張することができる. 標本空間が n 次元実ベクトル空間 \mathcal{R}^n であるとする. このとき, \mathcal{R}^n 上の全ての長方形領域 $\cap_{i=1}^n (a_i, b_i]$ を要素のもつ集合 $\mathcal{B}(\mathcal{R}^n)$ を \mathcal{R}^n 上の事象の集合と呼ぶ. 1 次元の場合と同様に以下のような性質をもつ関数 $F(x_1, x_2, \dots, x_n)$ を考える.

1. $\lim_{x_1 \rightarrow -\infty} \lim_{x_2 \rightarrow -\infty} \cdots \lim_{x_n \rightarrow -\infty} F(x_1, x_2, \dots, x_n) = 0$
2. $F(x_1, x_2, \dots, x_n)$ は各引数 x_i に対して非減少かつ右連続
3. $\lim_{x_1 \rightarrow \infty} \lim_{x_2 \rightarrow \infty} \cdots \lim_{x_n \rightarrow \infty} F(x_1, x_2, \dots, x_n) = 1$

このような性質をもつ関数 $F(x_1, x_2, \dots, x_n)$ は n 次元分布関数と呼ばれる. n 次元分布関数 $F(x_1, x_2, \dots, x_n)$ が与えられると, 標本空間と事象の集合の組 $(\mathcal{R}^n, \mathcal{B}(\mathcal{R}^n))$ に対して

$$P(\cap_{i=1}^n (-\infty, x_i]) = F(x_1, x_2, \dots, x_n), \quad -\infty < x_i < \infty$$

によって確率分布 P を決定することができる. n 次元分布関数によって決定される確率分布は n 次元確率分布, または, n 次元分布と呼ばれる.

また,

$$\int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \cdots \int_{-\infty}^{x_n} f(y_1, y_2, \dots, y_n) dy_1 dy_2 \cdots dy_n = F(x_1, x_2, \dots, x_n)$$

を満たす関数 $f(x_1, x_2, \dots, x_n)$ を $F(x_1, x_2, \dots, x_n)$ の密度関数といい,

$$\sum_{i_1 \leq x_1} \sum_{i_2 \leq x_2} \cdots \sum_{i_n \leq x_n} p_{i_1, i_2, \dots, i_n} = F(x_1, x_2, \dots, x_n)$$

を満たす関数 p_{i_1, i_2, \dots, i_n} を $F(x_1, \dots, x_n)$ の確率関数という.

事象 A_i ($i = 1, \dots, n$) が全ての i, j ($i \neq j$) に対して $A_i \cap A_j = \phi$ であるとき, 事象 A_i ($i = 1, \dots, n$) は排反 (exclusive, disjoint) と言われる. 一般に, n 個の事象 A_1, A_2, \dots, A_n が次の二つの条件を満たすとき, 事象の組 $\{A_1, A_2, \dots, A_n\}$ を標本空間 Ω の分割 (partition) という.

$$A_i \cap A_j = \phi, \quad i \neq j \quad (\text{各事象は互いに排反})$$

$$A_1 \cup A_2 \cup \cdots \cup A_n = \Omega \quad (\text{各事象の和集合は標本空間となる})$$

例えば、事象 A が与えられたとき、 $\{A, A^c\}$ は標本空間 Ω の分割である。

問 1.3 ([3]) 今、ある標本空間 Ω における二つの事象 A, B がある。以下の問いに答えよ。

- (a) 事象 A, B によって定まる四つの事象からなる集合を用いて標本空間 Ω を分割することができる。このような分割を求めよ (ヒント: ベン図を書いてみよ)。
- (b) $P(A) = a, P(B) = b, P(A \cap B) = c$ であるとする。前問 (a) の分割を与える各事象の確率を a, b, c を用いて表せ。
- (c) a, b, c がこのような確率として正しく定義できるための条件を示せ (ヒント: 確率は非負であり、かつ総和は 1)。

問 1.4 ([3]) 事象 A, B に対して $P(A) = 3/4, P(B) = 1/3$ であるとする。前問の結果を利用して以下の問いに答えよ。

- (a) $P(A \cup B)$ が取り得る可能性のある最大ならびに最小の値を求めよ。
- (b) $P(A \cup B)$ が最大であるとき、 $P(A \cap B)$ の値を求めよ。また、 $P(A \cup B)$ が最小であるとき、 $P(A \cap B)$ の値を求めよ。

次に条件付き確率を定義する。

定義 1.4 (条件付き確率) 事象 $A, B \in \mathcal{F}$ に対して、 $P(B) > 0$ のとき、

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

を事象 B が起こったという条件の下での事象 A の条件付き確率 (conditional probability) という。 □

定義より、積事象に対する確率 $P(A \cap B)$ は条件付き確率 $P(A | B)$ を用いて

$$P(A \cap B) = P(A | B)P(B) \tag{5}$$

と表すことができる。この関係を繰り返し用いると

$$P(A \cap B \cap C) = P(A \cap (B \cap C)) = P(A | B \cap C)P(B \cap C) = P(A | B \cap C)P(B | C)P(C) \tag{6}$$

という関係も得られる。さらに、条件付き確率の定義から

$$P(A \cap B | C) = \frac{P((A \cap B) \cap C)}{P(C)} = \frac{P(A \cap B \cap C)}{P(C)} \tag{7}$$

なので、式 (6), (7) から

$$P(A \cap B | C) = P(A | B \cap C)P(B | C) \tag{8}$$

を得る。

今、事象 $B_i (i = 1, \dots, n)$ が標本空間 Ω の分割であるとする。このとき、任意の事象 A に対して

$$P(A) = P(A \cap \Omega) = P\left(A \cap \left(\bigcup_{i=1}^n B_i\right)\right) = P\left(\bigcup_{i=1}^n (A \cap B_i)\right) = \sum_{i=1}^n P(A \cap B_i)$$

となるので、式 (5) を用いると

$$P(A) = \sum_{i=1}^n P(A | B_i)P(B_i) \tag{9}$$

を得る。ここで n は無限大でも良い。これは全確率の法則 (law of total probabilities) と呼ばれる。

また、ある事象 B に対して、事象 B_i ($i = 1, \dots, n$) が排反でかつ $B = \bigcup_{i=1}^n B_i$ であるとする。このとき、任意の事象 A に対して

$$\begin{aligned} P(A | B) &= P(A \cap B | B) = P\left(A \cap \left(\bigcup_{i=1}^n B_i\right) \middle| B\right) = P\left(\bigcup_{i=1}^n (A \cap B_i) \middle| B\right) \\ &= \frac{P\left(\left(\bigcup_{i=1}^n (A \cap B_i)\right) \cap B\right)}{P(B)} = \frac{P\left(\bigcup_{i=1}^n (A \cap B_i \cap B)\right)}{P(B)} = \frac{P\left(\bigcup_{i=1}^n (A \cap B_i)\right)}{P(B)} \\ &= \frac{\sum_{i=1}^n P(A \cap B_i)}{P(B)} = \sum_{i=1}^n P(A \cap B_i | B) \end{aligned}$$

となるので、式 (8) を用いると

$$P(A | B) = \sum_{i=1}^n P(A | B_i \cap B)P(B_i | B) = \sum_{i=1}^n P(A | B_i)P(B_i | B) \quad (10)$$

を得る。

また、式 (5)、式 (9) を用いると

$$P(B_j | A) = \frac{P(B_j \cap A)}{P(A)} = \frac{P(A | B_j)P(B_j)}{\sum_{i=1}^n P(A | B_i)P(B_i)}$$

となる。これはベイズの公式 (Bayes' formula) と呼ばれる。

問 1.5 (ポリヤの壺 (その 1)) r 個の赤玉と b 個の黒玉の入った壺がある。ここで、この壺から玉を一つ取りだし、取り出した玉に、それと同じ色の玉を c 個加えて、合計 $c+1$ 個の玉を壺へ戻すという操作を繰り返す。したがって、この試行を n 回 ($n = 1, 2, \dots$) 完了した後は、箱の中に $(r+b+nc)$ 個の玉がある。ただし、 r, b, c は全て自然数である。 $P(1$ 回目赤)、 $P(2$ 回目赤)、 $P(1$ 回目赤 | 2 回目赤) を求めよ。

問 1.6 ある工場には 3 台の機械 A, B, C があり、それぞれ製品の 40%, 30%, 30% を生産している。各機械で生産される製品の不良品の割合はそれぞれ 1%, 2%, 3% である。この工場で生産された製品からランダムに一つを選んだとき、その製品が不良品であった場合、それが機械 C で生産されたものである確率を求めよ。

問 1.7 (ウイルス感染確率 [3]) あるウイルスに感染している人は人口の 0.1% であると言われている。一方、このウイルスに感染しているか否かの検査を行うと、感染している人、感染していない人が、それぞれ、陽性、陰性と判定される確率は次の表の通りである。この検査によって陽性と判定された人が、実際にウイルスに感染している確率を求めよ。

	陽性	陰性
感染している人	0.9	0.1
感染していない人	0.03	0.97

問 1.8 (囚人の問題) A, B, C の 3 人の死刑囚のうち、二人は明日、刑が執行され、一人は釈放されることになった。そこで囚人 A は看守に「 B, C の内、少なくとも一人は刑を執行されるはずだから、明日、刑を執行される人を一人教えて欲しい」と頼んだ。看守はそれを教えても特別な情報を与えたことにはならないと考え、「 B は明日、刑が執行される」と教えた。これを聞いた A は「教えてもらうまでは自分が釈放される確率は $1/3$ であったが、教えてもらったので釈放される確率が $1/2$ に上がった」と喜んだという。看守が A に与えた情報は A が釈放される確率に影響を与えたか否かを吟味せよ。

問 1.9 (モンティホール問題 [3]) あるテレビ番組では、解答者がクイズに正解すると三つの衝立 A, B, C の中から一つを選び、その背後に隠されている商品を賞品として獲得する。三つの賞品のうち一つは高価な腕時計で

あり、他はほとんど価値のないものである。まず、クイズの解答者は一つの衝立を選ぶ。司会者は残りの二つの衝立のうち、腕時計がおかれていない衝立の一つを外す。ただし、二つとも腕時計が置かれていないならば、そのうちの一つをランダムに選ぶ。司会者は解答者に「今なら選択を変えても構いません。変えますか」と尋ねる。解答者はどう行動するのが合理的か。

定義 1.5 (独立) 事象 $A, B \in \mathcal{F}$ に対して、 $P(A \cap B) = P(A)P(B)$ が成り立つとき、事象 A と事象 B は独立 (independent) であるといわれる。また、一般に n 個の事象 A_i ($i = 1, \dots, n$) が与えられたとき、

$$P\left(\bigcap_{i=1}^n A_i\right) = \prod_{i=1}^n P(A_i)$$

が成り立つならば、事象 A_i ($i = 1, \dots, n$) は互いに独立 (mutually independent) であるといわれる。□

注意 1.3 もし、事象 A_i ($i = 1, \dots, n$) が互いに独立ならば、任意の m ($m = 2, \dots, n$) と任意の $i_j \in \{1, \dots, n\}$ に対して

$$P\left(\bigcap_{j=1}^m A_{i_j}\right) = \prod_{j=1}^m P(A_{i_j})$$

が成立する。しかし、 $m = n$ を除く全ての m ($m = 2, \dots, n-1$) と任意の $i_j \in \{1, \dots, n\}$ に対して上式が成立したとしても、事象 A_i ($i = 1, \dots, n$) は互いに独立とは限らない。□

問 1.10 例 1.1 において、 A_1 を 1 回目表、 A_2 を 2 回目裏、 A_3 を 1 回目、2 回目ともに表あるいは裏という事象とする。 A_i と A_j ($i \neq j$) は互いに独立となるか。また、 A_1, A_2, A_3 は互いに独立となるか。

問 1.11 ([3]) 大相撲の巴戦を考える⁹。3 人の実力は全く同等であるが、各対戦ではその前の対戦で控えていた力士の方が有利で、勝つ確率は p であったとする。このとき、最初に対戦した力士と、控えていた力士では、どちらが優勝する確率が高いか。ただし、最初の対戦だけは両者互いに五分であるとする。

1.3 確率変数とその分布

最初に確率変数を定義する。

定義 1.6 (確率変数) 確率空間 (Ω, \mathcal{F}, P) において、 Ω から実数の集合 \mathcal{R} への関数 $X(\omega)$ が任意の実数 x に対して

$$\{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F}$$

であるならば、 $X = X(\omega)$ を確率変数 (random variable) という。□

すなわち、確率変数とは、各標本に対して実数が対応しており、さらに、その実数が取り得る範囲を限定したとき、対応関係を用いて事象を特定することが可能であり、かつ、そのような事象に対する確率が定まるものである。

確率変数の定義において注意すべき点は、各標本に対して実数が定義されていなければならないことである。

例 1.4 例 1.1 において、 N を最初に表が出るまでにコインを投げた回数とする。このとき (表, 表) なら $N = 1$ 、(表, 裏) なら $N = 1$ 、(裏, 表) なら $N = 2$ であるが (裏, 裏) の場合は値が定義されていない。よって N は確率変数ではない¹⁰。□

⁹巴戦とは、千秋楽に 3 人の力士 A, B, C が同じ勝ち星で並んだとき、その中から一人の優勝者を決める方法である。まず、 A と B が対戦し、その勝者が C と対戦する。ここで連勝すればその力士が優勝、 C が勝てば、前の対戦で負けて控えていた力士と C が対戦し、 C が連勝すれば C の優勝、もし負ければ、勝った力士と控えていた力士が再度対戦し、...。このようにして 2 連勝する力士が現れるまで対戦を続け、最初に 2 連勝した力士が優勝を勝ち取る。

¹⁰このような場合 (裏, 裏) に対して ∞ を割り当て、確率変数 N の取り得る値に無限大を含めて考えることが多い。このように本来の確率変数の定義を拡張した結果、 $P(-\infty < X < \infty) < 1$ となるような確率変数を不完全な (defective) 確率変数という。

以下では、誤解を生じない限り確率空間を特に指定せず、単に「 X は確率変数である」と書くことにする。また $\{\omega \in \Omega \mid X(\omega) \leq x\}$ も単に $\{X \leq x\}$ と書き、さらに $P(\{X \leq x\})$ は単に $P(X \leq x)$ と書くことにする。

さて、確率変数 X に対して $F(x) = P(X \leq x)$ とおけば、 $F(x)$ は分布関数となる。このような $F(x)$ を確率変数 X の分布関数または分布という。さらに、確率変数 X は分布関数（あるいは分布） $F(x)$ をもつ、または、確率変数 X は分布関数（あるいは分布） $F(x)$ に従うという。以前と同様に、分布関数 $F(x)$ に対して密度関数 $f(x)$ も式 (3) 満たす非負関数として定義される。

問 1.12 非負の整数値をとる確率変数 X が、任意の非負の整数 m, n に対して $P(X > m + n \mid X > m) = P(X > n)$ を満たすとき、 $P(X = k)$ ($k = 0, 1, \dots$) はどのような形になるか論じよ。

問 1.13 非負の整数値をとる確率変数 X が、任意の非負の整数 m, n に対して $P(X \geq m + n \mid X \geq m) = P(X \geq n)$ を満たすとき、 $P(X = k)$ ($k = 0, 1, \dots$) はどのような形になるか論じよ。

1.4 確率ベクトルと結合分布，周辺分布

n 個の確率変数 X_1, X_2, \dots, X_n に対して、 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ を確率ベクトル (probability vector) という。また、 $B \in \mathcal{B}(\mathcal{R}^n)$ に対して、 $\{\mathbf{X} \in B\}$ は事象となり、確率空間 $(\mathcal{R}^n, \mathcal{B}(\mathcal{R}^n), P)$ を作る事が出来る。 $P(\mathbf{X} \in B)$ は \mathbf{X} の分布、あるいは確率変数 X_1, X_2, \dots, X_n の結合分布、あるいは同時分布 (joint distribution) と呼ばれる。確率変数 X_1, X_2, \dots, X_n の結合分布は n 次元分布関数 $F(x_1, x_2, \dots, x_n)$ を用いて一意に定める事が出来る。

$$F(x_1, x_2, \dots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n)$$

このとき、 $F(x_1, x_2, \dots, x_n)$ は X_1, X_2, \dots, X_n の結合分布関数 (joint distribution function) と呼ばれる。

n 次元分布関数 F に対して \mathcal{R} 上の関数 $F_1(x)$ を

$$F_1(x) = \lim_{x_2 \rightarrow \infty} \lim_{x_3 \rightarrow \infty} \cdots \lim_{x_n \rightarrow \infty} F(x, x_2, x_3, \dots, x_n) = P(X_1 \leq x)$$

とすると、 $F_1(x)$ は (1次元の) 分布関数である。このように多次元の分布関数 F において幾つかの引数に対して無限大の極限をとることで得られる分布関数で特徴づけられる分布を F の周辺分布 (marginal distribution) という。

定義 1.7 (確率変数の独立性) 確率変数 X_1, X_2, \dots, X_n が全ての $B_i \in \mathcal{B}(\mathcal{R})$ ($i = 1, \dots, n$) に対して

$$P(X_1 \in B_1, X_2 \in B_2, \dots, X_n \in B_n) = P(X_1 \in B_1)P(X_2 \in B_2) \cdots P(X_n \in B_n)$$

を満たすならば、確率変数 X_1, X_2, \dots, X_n は互いに独立であると呼ばれる。□

確率変数 X_1, X_2, \dots, X_n がそれぞれ周辺分布 $F_i(x)$ ($i = 1, \dots, n$) をもっているとき、これらが互いに独立であるということは結合分布が周辺分布の積に等しいことと等価である。

$$F(x_1, x_2, \dots, x_n) = F_1(x_1)F_2(x_2) \cdots F_n(x_n)$$

また、このとき、結合密度関数と周辺密度関数、あるいは結合確率関数と周辺確率関数に対しても同様の結果を得ることが出来る。

$$f(x_1, x_2, \dots, x_n) = f_1(x_1)f_2(x_2) \cdots f_n(x_n), \quad p(x_1, x_2, \dots, x_n) = p_1(x_1)p_2(x_2) \cdots p_n(x_n)$$

問 1.14 $P(X > x, Y > y)$ は (X, Y) の結合分布関数 $F(x, y)$ と周辺分布関数 $F_X(x), F_Y(y)$ を用いて $P(X > x, Y > y) = 1 - F_X(x) - F_Y(y) + F(x, y)$ と表されることを示せ。

問 1.15 X と Y を確率変数とする． X^2 と Y^2 は独立であるが， X と Y は独立でないような例を作れ．
 ヒント： $X \in \{-1, 0, 1\}$ ， $Y \in \{0, 1\}$ で考えてみよ．

問 1.16 ([3]) 二つのサイコロを振ったとき，大きい方の目を X ，小さい方の目を Y とする（同じ目なら $X = Y$ ）．
 X と Y の結合確率関数，および， X ， Y それぞれの周辺確率関数を求めよ．

問 1.17 (ポリヤの壺 (その 2)) 本文は問 1.5 の続きである． X_n ($n = 1, 2, \dots$) を n 回目の試行において取り出された玉が赤色なら 1，黒色なら 0 の値を取る確率変数とする．さらに Y_n ($n = 1, 2, \dots$) を次式で定義する．

$$Y_n = X_1 + X_2 + \dots + X_n$$

Y_n は最初の n 回の試行において赤色の玉が取り出された回数を表す確率変数である．

- (a) 確率 $P(Y_k = k, Y_n = k)$ ($n = 2, 3, \dots$ ， $k = 1, 2, \dots, n-1$) を求めよ．
 なお，「 $Y_k = k, Y_n = k$ 」は，最初の n 回 ($n = 1, 2, \dots$) の試行において， k 回目までの試行で取り出される玉が全て赤色であり，かつ，その後の $n-k$ 回の試行で取り出される玉が全て黒色である事象を表している．
- (b) 確率 $P(Y_n = k)$ ($n = 2, 3, \dots$ ， $k = 1, 2, \dots, n-1$) を求めよ．
- (c) 確率 $P(X_n = 1)$ ($n = 2, 3, \dots$) を求めよ．なお，任意の n ($n = 2, 3, \dots$) に対して

$$P(X_n = 1) = \sum_{i_1=0}^1 \sum_{i_2=0}^1 \dots \sum_{i_{n-1}=0}^1 P(X_1 = i_1, X_2 = i_2, \dots, X_{n-1} = i_{n-1}, X_n = 1)$$

が成立することを用いて良い．

- (d) $m < n$ なる任意の自然数 m, n に対して確率 $P(X_m = 1, X_n = 0)$ を求めよ．

問 1.18 X_1, X_2, \dots, X_N を互いに独立な N 個の確率変数とする．ただし $N \geq 3$ である．各確率変数 X_i ($i = 1, 2, \dots, N$) は 0, 1 のいずれかの値を取り，各 i ($i = 1, 2, \dots, N$) に対して， $p_i = P(X_i = 1)$ ， $q_i = 1 - p_i = P(X_i = 0)$ とする．以下では $p_i > 0$ ， $q_i > 0$ ($i = 1, 2, \dots, N$) と仮定し，これらの比を r_i で表す．

$$r_i = \frac{p_i}{q_i} \quad (i = 1, 2, \dots, N)$$

さらに

$$\sum_{i=k}^N r_i \neq 1 \quad (k = 2, 3, \dots, N)$$

であると仮定する．ここで K_N を次式で定義する．

$$K_N = \arg \max_{k \in \{1, 2, \dots, N\}} P(S_{k,N} = 1)$$

ただし K_N の定義式に現れる $S_{k,N}$ ($k = 1, 2, \dots, N$) は次式で与えられる．

$$S_{k,N} = \sum_{i=k}^N X_i$$

K_N は $P(S_{k,N} = 1)$ が最大となるような $k \in \{1, 2, \dots, N\}$ である．以下の (a) ~ (e) の設問に答えよ．

- (a) $P(S_{k,N} = 1)$ ($k = 1, 2, \dots, N$) を， q_i ($i = k, k+1, \dots, N$) と r_i ($i = k, k+1, \dots, N$) を用いて表せ．
- (b) $P(S_{k+1,N} = 1) - P(S_{k,N} = 1)$ ($k = 1, 2, \dots, N-1$) を， p_k, q_i ($i = k+1, k+2, \dots, N$) ならびに r_i ($i = k+1, k+2, \dots, N$) を用いて表せ．
- (c) $K_N = 1$ となるための必要十分条件を求めよ．
- (d) $K_N = k$ ($k = 2, 3, \dots, N$) となるための必要十分条件を求めよ．
- (e) $p_i = 1/(i+1)$ ($i = 1, 2, \dots, N$) であるとき，

$$\lim_{N \rightarrow \infty} \frac{K_N}{N}$$

を求めよ．ただし，3 以上の任意の自然数 N に対して， $r_k + r_{k+1} + \dots + r_N \neq 1$ ($k = 2, 3, \dots, N$) と仮定してよい．また，十分大きな N に対して $K_N \geq 2$ と仮定してよい．

問 1.19 (前問の続き: 古典的秘書問題) 研究室の秘書を募集したところ N 人の応募者があった. この N 人の中から最も良い人を採用したい. 採用に際しては, できるだけ順で一人ずつ面接を行い, その場で採用するか否かを判断する. 一旦, 採用する, あるいはしないと決めたら, その判断を変えることはできない. 判断を下す際, 過去に面接した応募者と比較し, 現在, 面接している応募者が過去に面接した応募者より良いか否かが確実に判別できるとする. 最良の応募者を採用することが目的なので, 過去に面接した応募者よりも劣るのであれば決して採用しない. よって, 誰も採用しないこともあり得る.

X_i ($i = 1, 2, \dots, N$) を 1 番目から i 番目までの応募者の中で i 番目の応募者が最良であるとき 1, そうでないとき 0 の値を取る確率変数とする. 定義より, $P(X_1 = 1) = 1$ である. 以下の (a) ~ (d) の設問に答えよ.

- (a) $P(X_i = 1)$ ($i = 2, 3, \dots, N$) を求めよ.
- (b) 任意の i ($i = 2, 3, \dots, N$) に対して, X_1, X_2, \dots, X_i は互いに独立であることを示せ.
- (c) 最初に面接する $K - 1$ 人の採用は無条件に見送り, 残りの $N - K$ 人の応募者の中から, 面接した応募者の中で最良であった場合, その応募者を採用する事にする. この採用戦略において, 最良の応募者を採用する確率が最大になる K^* は前問の K_N を用いて $K^* = K_N$ であることを示せ (ヒント) i 番目の応募者が最良であれば, $X_i = 1$ となり, それ以降の X_j ($j = i + 1, i + 2, \dots, N$) は全て 0 である. よって, 最良の応募者を選ぶということは, 最後に 1 の値をとるものを選ぶことと等価である.
- (d) この問題の設定で, 前問の (e) の極限を求めよ.

(コメント) 設問 (d) は, 選択肢がいくつかあり, それらを順に吟味して, どこかで一つを選択しなければならない場合 (例えば, お見合いなど), 最良のものを手にするには, 最初の $1/e$ を捨てて, 次に最も良いものが出てきたときにそれを選択する戦略が, 最良のものをを選択する確率を最大にする, ということを示している.

1.5 確率変数を引数とする関数と確率変数の畳み込み

密度関数 $f(x)$ をもつ確率変数 X と, 実数から実数への関数 $h(x)$ が与えられたとき, $Y = h(X)$ も確率変数となる. このとき, 確率変数 Y の分布関数 $G(x)$ は

$$G(x) = P(h(X) \leq x) = \int_{h(u) \leq x} f(u) du$$

で与えられる. もし, 確率変数 X が確率関数 $p(x)$ をもつならば

$$G(x) = \sum_{h(u) \leq x} p(u)$$

である.

問 1.20 確率変数 X の分布関数を $F(x)$ とする. このとき $Y = X^2$ の分布関数 $G(x)$ を $F(x)$ を用いて表せ.

問 1.21 A, B, C を独立で同一な $[0, 1]$ 上の一様分布に従う確率変数とする. このとき x の 2 次方程式 $Ax^2 + 2Bx + C = 0$ が実根をもつ確率を求めよ.

次に確率変数の和が従う分布を考える. 今, 確率変数 X_1, X_2 がそれぞれ分布関数 $F_1(x), F_2(x)$ ならびに密度関数 $f_1(x), f_2(x)$ をもっており, 独立であるとする. このとき, 新しい確率変数 $Y = X_1 + X_2$ の分布関数 $G(x) = P(Y \leq x)$ は

$$\begin{aligned} G(x) &= P(X_1 + X_2 \leq x) = \int \int_{u_1 + u_2 \leq x} f_1(u_1) f_2(u_2) du_1 du_2 \\ &= \int_{u_2 = -\infty}^{\infty} \int_{u_1 = -\infty}^{x - u_2} f_1(u_1) f_2(u_2) du_1 du_2 = \int_{-\infty}^{\infty} F_1(x - u_2) f_2(u_2) du_2 \end{aligned}$$

となる. 左辺の最後の積分は分布関数 $F_1(x)$ と $F_2(x)$ の畳み込み (convolution) と呼ばれ, $F_1 * F_2(x)$ と書かれることが多い.

$$F_1 * F_2(x) = \int_{-\infty}^{\infty} F_1(x - u) f_2(u) du$$

もし、上記において X_1 と X_2 の役割を交換して計算すると

$$\begin{aligned} G(x) &= P(X_1 + X_2 \leq x) = \int \int_{u_1+u_2 \leq x} f_2(u_2)f_1(u_1)du_2du_1 \\ &= \int_{u_1=-\infty}^{\infty} \int_{u_2=-\infty}^{x-u_1} f_2(u_2)f_1(u_1)du_2du_1 = \int_{-\infty}^{\infty} F_2(x-u_1)f_1(u_1)du_1 \end{aligned}$$

となる。すなわち、分布関数の畳み込みは計算の順序と無関係である。

$$G(x) = F_1 * F_2(x) = F_2 * F_1(x)$$

特に独立な非負の確率変数 X_1 と X_2 に対しては

$$G(x) = \int_0^x F_1(x-u_2)f_2(u_2)du_2 = \int_0^x F_2(x-u_1)f_1(u_1)du_1$$

となる。

さらに

$$\frac{d}{dx}G(x) = \frac{d}{dx} \int_{-\infty}^{\infty} F_1(x-u)f_2(u)du = \int_{-\infty}^{\infty} f_1(x-u)f_2(u)du$$

となるので、分布関数 $G(x)$ の密度関数 $g(x)$ は

$$g(x) = \int_{-\infty}^{\infty} f_1(x-u)f_2(u)du$$

で与えられる。左辺の積分は密度関数 $f_1(x)$ と $f_2(x)$ の畳み込みと呼ばれ、 $f_1 * f_2(x)$ と書かれる。上式で $v = x-u$ と変数変換すれば

$$g(x) = \int_{-\infty}^{\infty} f_1(v)f_2(x-v)dv = f_2 * f_1(x)$$

となるので、密度関数の畳み込みに関しても計算の順序とは無関係である。

$$g(x) = f_1 * f_2(x) = f_2 * f_1(x)$$

特に独立な非負の確率変数 X_1 と X_2 に対しては

$$g(x) = \int_0^x f_1(x-u_2)f_2(u_2)du_2 = \int_0^x f_1(u_1)f_2(x-u_1)du_1$$

となる ($F_1(0) = F_2(0) = 0$ と仮定している。問 1.25 参照)。

例 1.5 (正規分布の再現性) パラメタ m, σ (ただし $\sigma > 0$) をもつ密度関数

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$$

に従う分布を正規分布 (normal distribution) といい、 $N(m, \sigma^2)$ と書く。なお、 m は平均、 σ^2 は分散を表している。畳み込みの例として、独立な正規分布 $N(m_1, \sigma_1^2)$ 、 $N(m_2, \sigma_2^2)$ の密度関数の畳み込み $g(x)$ を考える。定義より

$$\begin{aligned} g(x) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{1}{2} \frac{(x-u-m_1)^2}{\sigma_1^2}\right) \frac{1}{\sqrt{2\pi}\sigma_2} \exp\left(-\frac{1}{2} \frac{(u-m_2)^2}{\sigma_2^2}\right) du \\ &= \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma_1^2\sigma_2^2} \left[(\sigma_1^2 + \sigma_2^2) \left(u - \frac{(x-m_1)\sigma_2^2 + m_2\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right)^2 + \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2} (x-m_1-m_2)^2 \right]\right) du \\ &= \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{(x-m_1-m_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right) \int_{-\infty}^{\infty} \exp\left(-\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \left(u - \frac{(x-m_1)\sigma_2^2 + m_2\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\right)^2\right) du \\ &= \frac{1}{2\pi\sigma_1\sigma_2} \exp\left(-\frac{(x-m_1-m_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right) \frac{\sigma_1\sigma_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \int_{-\infty}^{\infty} e^{-y^2/2} dy \end{aligned}$$

ここで $\int_{-\infty}^{\infty} e^{-y^2/a^2} dy = \sqrt{a\pi}$ を用いると

$$g(x) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_1^2 + \sigma_2^2}} \exp\left(-\frac{(x - m_1 - m_2)^2}{2(\sigma_1^2 + \sigma_2^2)}\right)$$

を得る．すなわち，独立な正規分布 $N(m_1, \sigma_1^2)$, $N(m_2, \sigma_2^2)$ の密度関数の畳み込み $g(x)$ は正規分布 $N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$ である．言い換えれば，正規分布 $N(m_1, \sigma_1^2)$, $N(m_2, \sigma_2^2)$ に従う独立な二つの確率変数 X_1, X_2 の和 $Y = X_1 + X_2$ は正規分布 $N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$ に従う．これは正規分布の再現性と呼ばれている． \square

問 1.22 X_i ($i = 1, 2, \dots$) を独立かつ同一なパラメタ μ ($\mu > 0$) をもつ指数分布に従う確率変数列とする．

$$P(X_i \leq x) = 1 - e^{-\mu x}, \quad x \geq 0$$

$F_k = X_1 + X_2 + \dots + X_k$ ($k = 1, 2, \dots$) としたとき，

$$P(F_k \leq x) = 1 - \sum_{i=1}^{k-1} e^{-\mu x} \frac{(\mu x)^i}{i!} \quad (11)$$

であることを帰納法を用いて証明せよ．なお， F_k が従う分布は k 次のアーラン分布 (Erlang distribution) と呼ばれる)

問 1.23 独立な確率変数 X, Y がともに $(0, 1]$ 上の一様分布に従うとする．このとき， $Z = X + Y$ の密度関数を求めよ．

問 1.24 X, Y を独立で同一の分布に従う確率変数とし，分布関数を $F(x) = P(X \leq x) = P(Y \leq x)$ とする．

(a) 正の実数 μ ($\mu > 0$) に対して

$$F(x) = \begin{cases} 0 & (x < 0) \\ 1 - \exp(-\mu x) & (x \geq 0) \end{cases} \quad (12)$$

であるとき， $Z_1 = X - Y$ の密度関数を求めよ．

(b) $F(x)$ が式 (12) で与えられるとき， $Z_2 = (X + Y)/\alpha$ の密度関数を求めよ．ただし α ($\alpha > 0$) は正の実数である．

(c) 正の実数 γ ($\gamma > 0$) に対して

$$F(x) = \begin{cases} 0 & (x < 0) \\ \frac{x}{\gamma} & (0 \leq x \leq \gamma) \\ 1 & (x > \gamma) \end{cases}$$

であるとき，新しい確率変数 U, V を $U = \min(X, Y)$, $V = \max(X, Y)$ と定義する． $Z_3 = V/U$ の分布関数を求めよ．

問 1.25 非負の値をとる独立な確率変数 X_1, X_2 がそれぞれ分布関数 $F_1(x), F_2(x)$ ならびに密度関数 $f_1(x), f_2(x)$ をもつとする．ただし $F_1(0) = F_2(0) = 0$ ．このとき， $Y = (X_1 + X_2)/2$ の分布関数ならびに密度関数を求めよ．

次に整数値をとる離散型分布の畳み込みを考える．独立な二つの確率変数 X_1, X_2 がそれぞれ分布 $\{p_1(k)\}, \{p_2(k)\}$ をもつとする．このとき確率変数 $Y = X_1 + X_2$ は整数値をとり，その確率関数 $q(k) = P(Y = k)$ は

$$q(k) = \sum_{i=-\infty}^{\infty} p_1(k-i)p_2(i) = \sum_{i=-\infty}^{\infty} p_1(i)p_2(k-i)$$

で与えられる．特に独立な非負の整数値を取る確率変数 X_1 と X_2 に対しては，確率変数 $Y = X_1 + X_2$ は非負の整数値をとり，

$$q(k) = \sum_{i=0}^k p_1(k-i)p_2(i) = \sum_{i=0}^k p_1(i)p_2(k-i)$$

となる .

問 1.26 (ポワソン分布の再現性) 正のパラメタ λ をもつ確率関数

$$p(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0, 1, \dots$$

に従う分布をポワソン分布 (Poisson distribution) という . 独立なパラメタ λ_1, λ_2 をもつポワソン分布に従う確率変数 X_1, X_2 の和 Y に対する確率関数 $q(k)$ ($k = 0, 1, \dots$) はパラメタ $\lambda_1 + \lambda_2$ をもつポワソン分布になることを示せ .

問 1.27 X_i ($i = 1, 2, \dots$) を独立かつ同一なパラメタ p ($0 < p < 1$) をもつ幾何分布に従う離散確率変数列とする .

$$P(X_i = m) = (1-p)p^{m-1}, \quad m = 1, 2, \dots$$

$F_k = X_1 + X_2 + \dots + X_k$ ($k = 1, 2, \dots$) としたとき ,

$$P(F_k = m) = \frac{(m-1)!}{(k-1)!(m-k)!} (1-p)^k p^{m-k}, \quad m = k, k+1, \dots \quad (13)$$

であることを帰納法を用いて証明せよ . ただし , 次の恒等式を用いて良い .

$$\frac{(n+m+1)!}{m!(n+1)!} = \sum_{k=0}^m \frac{(n+k)!}{k!n!}$$

問 1.28 以下の問いに答えよ .

- (a) 互いに独立な $n+1$ 個の連続確率変数 X_i ($i = 0, 1, \dots, n$) がある . X_i の分布関数と密度関数をそれぞれ $F_i(x), f_i(x)$ とする . $P(X_0 = \min(X_0, X_1, \dots, X_n))$ は分布関数 , 密度関数を用いてどのように表現できるか示せ .
- (b) X_i ($i = 0, 1, \dots$) を分布関数 $F(x)$, 密度関数 $f(x)$ をもつ , 独立で同一な分布に従う (independent and identically distributed, 略して i.i.d.) 連続確率変数の列とする . 問 (a) の結果を利用して , $P(X_0 = \min(X_0, X_1, \dots, X_n))$ と $P(X_0 = \min(X_0, X_1, \dots, X_{n-1}))$ の間に成り立つ関係を示せ . さらに , $P(X_0 = \min(X_0)) = 1$ に注意して , $P(X_0 = \min(X_0, X_1, \dots, X_n))$ を求めよ .

問 1.29 X_n ($n = 1, 2, \dots, N$) を独立かつ同一な分布に従う連続確率変数とする . このとき , 次式が成立することを示せ .

$$\lim_{x \rightarrow \infty} \frac{P(\max(X_1, X_2, \dots, X_N) > x)}{P(X_1 > x)} = N$$

問 1.30 X と Y が独立で同一なパラメタ 1 の指数分布に従う確率変数とする . すなわち $P(X \leq x, Y \leq y) = \{1 - \exp(-x)\} \cdot \{1 - \exp(-y)\}$. $Z_1 = X + Y$, $Z_2 = X/Y$ としたとき , Z_1 と Z_2 は独立であることを示せ .

問 1.31 非負の確率変数 X は微分可能な分布関数 $F(x) = P(X \leq x)$ ($x \geq 0$) をもち , かつ

$$P(X > x+y \mid X > x) = P(X > y)$$

が全ての x, y ($x, y \geq 0$) に対して成立するとき , $F(x)$ を求めよ .

2 期待値

2.1 期待値の定義

最初に非負の確率変数 X の期待値を定義する .

定義 2.1 (非負の確率変数の期待値) 密度関数 $f(x)$ をもつ非負の確率変数 X の期待値 (expectation) $E(X)$ は

$$E(X) = \int_0^{\infty} xf(x)dx \quad (14)$$

で与えられる。また、確率関数 $p(x)$ をもつ非負の確率変数 X の期待値 $E(X)$ は

$$E(X) = \sum_{i=1}^{\infty} x_i p(x_i) \quad (15)$$

で与えられる。 $E(X)$ は無限大となることもある。□

問 2.1 $\exp(\lambda) = \sum_{k=0}^{\infty} \lambda^k / k!$ に注意して、問 1.26 で与えたポワソン分布の平均を求めよ。

問 2.2 分布関数が $P(X \leq x) = F(x) = 1 - \exp(-\mu x)$ ($x \geq 0, \mu > 0$) で与えられるパラメタ μ をもつ指数分布の平均を求めよ。

問 2.3 正確な硬貨を初めて表が出るまで投げ続ける実験を行う。硬貨を投げた回数を N とする。実験の結果が $N = k$ ($k = 1, 2, \dots$) であるとき、 2^k 円の報酬を受け取ることが出来るとすると、報酬の期待値はいくらになるか。

式 (15) で与えられる確率関数が存在する場合の期待値は、日常でも普通に用いられている。例えば、サイコロを 1 度投げたときに出る目 X の平均は

$$\sum_{k=1}^6 kP(X = k) = \sum_{k=1}^6 k \cdot \frac{1}{6} = 3.5$$

と求められる。一方、式 (14) で与えられる密度関数をもつ非負の確率変数に対する期待値は積分を用いて与えられているが、これは本質的に確率関数をもつ場合と同じであることを以下で説明する。

まず、定義から

$$E(X) = \lim_{b \rightarrow \infty} \int_0^b xf(x)dx$$

に注意する。ここで右辺の積分は以下のように解釈される。まず、閉区間 $[0, b]$ を n 等分することにより、 n 個の小区間 $[x_{i-1}, x_i]$ ($i = 1, 2, \dots, n$) を作る。

$$x_0 = 0, \quad x_i = \frac{ib}{n} \quad (i = 1, 2, \dots, n)$$

このとき、小区間の長さは全て等しく $\Delta x_i = b/n$ であり、それらの最大値 $|\Delta| = |\Delta(n)|$ も b/n で与えられる。各小区間 $[x_{i-1}, x_i]$ の左端の点を $\xi_i = x_{i-1}$ とすると、Riemann 積分¹¹の定義から

$$\int_0^b xf(x)dx = \lim_{|\Delta(n)| \rightarrow 0} \sum_{i=1}^n \xi_i f(\xi_i) |\Delta(n)| = \lim_{n \rightarrow \infty} \sum_{i=1}^n x_{i-1} f(x_{i-1}) |\Delta(n)| = \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} x_i f(x_i) |\Delta(n)|$$

を得る。ここで、式 (4) より

$$f(x_i) |\Delta(n)| = P\left(x_i - \frac{|\Delta(n)|}{2} < X \leq x_i + \frac{|\Delta(n)|}{2}\right) + o(|\Delta(n)|)$$

と解釈されることに注意する。すなわち、式 (14) で与えられる密度関数 $f(x)$ をもつ非負の確率変数 X の期待値 $E(X)$ は

$$\lim_{n \rightarrow \infty} o(|\Delta(n)|) = \lim_{n \rightarrow \infty} o(b/n) = 0$$

に注意すると、

$$E(X) = \int_0^{\infty} xf(x)dx = \lim_{b \rightarrow \infty} \int_0^b xf(x)dx = \lim_{b \rightarrow \infty} \left[\lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} x_i P\left(x_i - \frac{b}{2n} < X \leq x_i + \frac{b}{2n}\right) \right] \quad (16)$$

¹¹ 区分求積法により定義される定積分。付録を参照。

と書き換えることができる。

式 (15) と式 (16) の類似性に注意する。式 (16) となる連続な確率変数の場合、標本空間を微少な小区間に分割し、確率変数の取る値が特定の小区間にある確率とその区間の中央値を掛け、全ての場合について和を取ったものの極限になっており、これは式 (15) の確率関数が存在する離散な確率変数の期待値と同形である。

次に任意の確率変数 X の期待値を考える。まず、 $X^+ = \max(0, X)$ 、 $X^- = -\min(0, X)$ によって新しい二つの確率変数を導入する。定義より $X^+ \geq 0$ 、 $X^- \geq 0$ であり、 $X = X^+ - X^-$ である。

定義 2.2 (確率変数の期待値) もし、 $E(X^+)$ と $E(X^-)$ の少なくとも一方が有限であるならば、 X の期待値は

$$E(X) = E(X^+) - E(X^-)$$

で与えられる。 $E(X^+) = E(X^-) = \infty$ の場合は期待値は存在しない。□

もし、期待値 $E(X)$ が存在するならば、定義 2.2 より、密度関数 $f(x)$ をもつ場合は

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

となり、離散的ならば

$$E(X) = \sum_{i=1}^{\infty} x_i p(x_i)$$

となる。これらの右辺は分布 $F(x)$ の平均 (mean) と呼ばれる。

期待値は以下の性質をもつ。

定理 2.1 (期待値の性質) 確率変数 X と Y の期待値 $E(X)$ 、 $E(Y)$ が共に有限であるとする。このとき

1. $X \leq Y$ ならば $E(X) \leq E(Y)$
2. a を定数としたとき $E(aX) = aE(X)$
3. $E(X + Y) = E(X) + E(Y)$
4. X と Y が互いに独立であるならば $E(XY) = E(X)E(Y)$

が成立する。□

問 2.4 箱の中に 1 等から 10 等までのくじがそれぞれ 1 本ずつ、10 本入っている。この箱からくじを引いて等級を確かめた後、箱に戻す操作を k 回行う。取り出されたくじの種類数の期待値を求めよ。

ヒント： X_i を少なくとも 1 回、 i 等のかくじが引かれたら 1、一度も引かれなければ 0 をとる確率変数としたとき、求めたい量はどのように書けるか。

問 2.5 1 から N ($N \geq 2$) の番号が書かれた N 枚のカードを良く切って並べたときカードの番号と順序が一致している (i 枚目のカードの番号が i) 枚数の期待値を求めよ。

ヒント： X_i を i 番目のカードが一致していれば 1、一致していなければ 0 をとる確率変数としたとき、求めたい量はどのように書けるか。

問 2.6 正数 r に対して、 $a_1 = 0$ 、 $a_2 = r$ とおき、数列 $\{a_n\}$ を次の漸化式で定める

$$a_{n+1} = a_n + r_n(a_n - a_{n-1}) \quad (n = 2, 3, \dots)$$

ただし a_n と a_{n-1} から漸化式を用いて a_{n+1} を決める際には硬貨を投げ、表がでたとき $r_n = r/2$ 、裏がでたとき $r_n = 1/(2r)$ とする。ここで表がでる確率と裏がでる確率は等しいとする。 a_n の期待値を p_n とするとき、以下の問いに答えよ。

- (1) $n \geq 3$ のときに p_n を、 n と r を用いて表せ。

(2) 数列 $\{p_n\}$ が収束するような正数 r の範囲を求めよ .

一般に , 確率変数の関数に対しては次の定理が成り立つ .

定理 2.2 密度関数 $f(x)$ をもつ確率変数 X と関数 $u(x)$ に対して期待値 $E(u(X))$ が存在するならば , それは

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

で与えられる . また , 確率関数 $p(x_i)$ ($i = 1, 2, \dots$) をもつ場合は

$$E(u(X)) = \sum_{i=1}^{\infty} u(x_i)p(x_i)$$

で与えられる . □

例 2.1 非負の確率変数 X の期待値を別の方法で求める .

$$\sum_{n=0}^{\infty} P(X > n) = \sum_{n=0}^{\infty} \sum_{k=n+1}^{\infty} P(X = k) = \sum_{k=1}^{\infty} \sum_{n=0}^{k-1} P(X = k) = \sum_{k=1}^{\infty} kP(X = k)$$

となるため ,

$$E(X) = \sum_{n=0}^{\infty} P(X > n)$$

で与えられる . □

なお , 例 2.1 と同様の結果が , 非負の確率変数 X が連続確率変数の場合にも以下のように成立する .

$$E(X) = \int_0^{\infty} (1 - F(x))dx \tag{17}$$

問 2.7 式 (17) が成立することを , $E(X) = \infty$ と $E(X) < \infty$ の場合に分けて示せ .

問 2.8 表が出る確率が p ($0 < p \leq 1$) , 裏が出る確率が $1 - p$ で与えられる硬貨がある . 今 , この硬貨を繰り返し投げ , 初めて k 回連続して表が出た時点で試行を停止する . ただし k は自然数である . 試行を停止するまでに硬貨を投げる回数の期待値を M_k としたとき , (i) $k = 1$, すなわち , 初めて表が出た時点で試行を停止する場合 , 停止するまでに硬貨を投げる回数の期待値 M_1 を求めよ . また , (ii) $k \geq 2$ の場合 , M_k を p と M_{k-1} を用いて表した上 , (iii) M_k ($k = 1, 2, \dots$) を求めよ .

問 2.9 (不幸のパラドックス (ill-luck paradox) [1]) 大勢の人がおみくじを引く場面を想定する . おみくじには , 連続な分布からの標本値が一つ書かれており , 大きい数字ほどより吉兆であるとする . この設定において , 最初におみくじを引いた人よりも , 悪いおみくじを引く人が出てくるまで , 平均 , 何人の人がさらにおみくじを引くことになるかを考える . $\{X_n; n = 0, 1, \dots\}$ を独立同一な分布に従う連続確率変数の列であるとする . このとき , $N = \min\{n; X_n < X_0\}$ と定義すると , この問いの答えは $E(N)$ である . 確率 $P(N > n)$ が $n + 1$ 個の独立同一な分布に従う確率変数 X_i ($i = 0, 1, \dots, n$) のうち , X_0 が最小の値を取る確率 (問 1.28 参照) と等価であることに注意して $E(N)$ を求めよ .

問 2.10 X_n ($n = 0, 1, 2, \dots$) を独立で同一の分布に従う連続な確率変数の列とする . X_n ($n = 0, 1, 2, \dots$) の分布関数を $F(x) = P(X_n \leq x)$ としたとき , 全ての x ($-\infty < x < \infty$) に対して $1 - F(x) = P(X_n > x) > 0$ であると仮定する . ここで確率変数 N を $N = \min\{k; X_k > X_0$ ($k = 1, 2, \dots\})$ と定義する . すなわち , 任意の自然数 k に対して

$$N = k \iff X_n \leq X_0 \text{ (} n = 1, 2, \dots, k-1 \text{) かつ } X_k > X_0$$

である . 以下の (a) ~ (d) の設問に答えよ .

(a) $N = k$ であり , かつ $X_N \leq x$ である結合確率 $P(N = k, X_N \leq x)$ を求めよ .

(b) N の期待値 $E[N]$ を求めよ .

(c) X_N の分布関数 $G(x) = P(X_N \leq x)$ を求めよ . ただし

$$\log(1-x) = -\sum_{n=1}^{\infty} \frac{x^n}{n} \quad (-1 < x < 1)$$

を用いてよい .

(d) ある正数 μ ($\mu > 0$) に対して

$$F(x) = \begin{cases} 0 & (x < 0) \\ 1 - \exp(-\mu x) & (x \geq 0) \end{cases}$$

であるとする . $N \geq 2$ という条件の下での X_1 の条件付き期待値 $E[X_1 | N \geq 2]$ を求めよ .

問 2.11 U_1, U_2, \dots を互いに独立かつ同一な $(0,1)$ 上の一様分布に従う確率変数列とする . $0 \leq x \leq 1$ なる x に対して $M(x)$ を $M(x) = \min(k \geq 1; U_1 + U_2 + \dots + U_k > x)$ と定義する . すなわち「 $M(x) = n$ 」は「 $U_1 + U_2 + \dots + U_{n-1} \leq x, U_1 + U_2 + \dots + U_n > x$ 」と等価である . 以下の三つの問いに答えよ .

(i) $P(M(x) > 1)$ を求めよ . (ii) $P(M(x) > n)$ を求めよ . (iii) $M(x)$ の期待値 $E[M(x)]$ を求めよ .

最後に分布と期待値の間に成立する不等式を導く . 密度関数 $f(x)$ をもつ非負の確率変数 X と $x > 0$ に対して

$$P(X \geq x) = \int_{x-}^{\infty} f(y)dy \leq \int_{x-}^{\infty} \frac{y}{x} f(y)dy = \frac{1}{x} \int_{x-}^{\infty} y f(y)dy \leq \frac{1}{x} \int_0^{\infty} y f(y)dy = \frac{E[X]}{x}$$

となるので , 次の定理を得る .

定理 2.3 (マルコフ不等式) 非負の確率変数 X に対して

$$P(X \geq x) \leq \frac{E(X)}{x}, \quad x > 0$$

が成立する . これはマルコフ不等式 (Markov's Inequality) と呼ばれている . □

もし , $x \leq E(X)$ ならば $E(X)/x \geq 1$ となるので , マルコフ不等式は $x > E(X)$ の場合にのみ意味がある .

2.2 積率 , 分散 , 共分散

X を分布関数 $F(x)$ をもつ確率変数とする . このとき , \mathcal{R} から \mathcal{R} への可測関数 $u(x)$ に対する期待値 $E(u(X))$ は , 分布関数 $F(x)$ の関数と見なすことが出来る¹² . それゆえ , 期待値 $E(u(X))$ を調べることで分布関数 $F(x)$ の特徴をある程度把握することが出来る . このような目的で良く用いられるものには $u(x) = x^n$ や $u(x) = (x - E(X))^n$ などがある .

定義 2.3 分布関数 $F(x)$ をもつ確率変数 X に対して , $E(X^n)$ を確率変数 X (あるいは分布 $F(x)$) の n 次積率 (n th moment) という . $n = 1$ のとき , これは確率変数 X の期待値 , あるいは分布 $F(x)$ の平均に等しい . また , $E(\{X - E(X)\}^n)$ を確率変数 X (あるいは分布 $F(x)$) の n 次中心積率 (n th central moment) という . 特に , $n = 2$ の場合 , $E(\{X - E(X)\}^2)$ を分散 (variance) と呼び , $V(X)$ と書く . □

分散 $V(X) = E(\{X - E(X)\}^2)$ は確率変数 X の値が期待値からどの程度離れた値を取るかという , バラツキを表す指標として用いられる . 分散には次のような性質がある .

定理 2.4 $E(X^2) < \infty$ ならば

1. $V(X) = E(X^2) - \{E(X)\}^2$
2. 定数 a, b に対して $V(aX + b) = a^2 V(X)$

¹²このような , 関数の関数を汎関数 (functional) という .

$$3. V(X) = 0 \Leftrightarrow P(X = E(X)) = 1$$

が成立する . □

問 2.12 X, Y を独立で同一の分布に従う確率変数とし, 分布関数を $F(x) = P(X \leq x) = P(Y \leq x)$ とする . 正の実数 β ($\beta > 0$) に対して

$$F(x) = \frac{\beta}{\sqrt{\pi}} \int_{-\infty}^x \exp[-\beta^2 y^2] dy \quad (-\infty < x < \infty)$$

であるとき, $Z = X^2 + Y^2$ の分散を求めよ .

問 2.13 独立な確率変数 X, Y に対して, $V(XY)$ を $E(X), V(X), E(Y), V(Y)$ を用いて表せ .

さて, 分散 $V(X)$ をもつ確率変数 X に対して $Y = (X - E(X))^2$ により, 新しい確率変数 Y を定義する . Y に定理 2.3 のマルコフ不等式を適用すると

$$P(Y \geq y) \leq \frac{E(Y)}{y}$$

を得る . ここで $y > 0$ に対して $y = x^2$ ($x > 0$) とおくと

$$P(Y \geq x^2) = P((X - E(X))^2 \geq x^2) = P(|X - E(X)| \geq x)$$

となり, $E(Y) = V(X)$ に注意すると

$$P(|X - E(X)| \geq x) \leq \frac{V(X)}{x^2} \tag{18}$$

を得る . これはチェビシェフの不等式 (Chebyshev's inequality) と呼ばれている . マルコフ不等式が分布の上側の裾野の上界値のみを与え, また, 非負の確率変数のみに適用できるのに対し, チェビシェフの不等式は分布の両側の裾野の上界値を与えていることに注意する . また, マルコフ不等式の上界値は 0 からの距離の逆数に比例する形で 0 へ向かうのに対し, チェビシェフの不等式の上界値は平均からの距離の二乗の逆数に比例する形で 0 へ向かう .

問 2.14 (標準化 (standardization)) 確率変数 X に対して $0 < V(X) < \infty$ のとき

$$Y = \frac{X - E(X)}{\sqrt{V(X)}} \tag{19}$$

によって新しい確率変数 Y を定義する . Y の期待値と分散を求めよ .

式 (19) の変換は標準化と呼ばれており, 分母に現れる $\sqrt{V(X)}$ は標準偏差 (standard deviation) と呼ばれる . また, 非負の確率変数 X に対しては, バラツキを表す無次元の量として変動係数 (coefficient of variation) $C(X)$ がある .

$$C(X) = \frac{\sqrt{V(X)}}{E(X)}$$

問 2.15 ある正数 $\mu > 0$ に対して, 分布関数 $F(x)$ ($x \geq 0$) が

$$F(x) = 1 - e^{-\mu x}$$

で与えられる分布をパラメタ μ をもつ指数分布という . パラメタ μ をもつ指数分布の平均, 分散, 変動係数を求めよ .

次に二つの確率変数の間の依存性を表す指標を導入する .

定義 2.4 (共分散) 平均をもつ二つの確率変数 X と Y に対して定義された

$$\text{Cov}(X, Y) = E(\{X - E(X)\}\{Y - E(Y)\}) \tag{20}$$

を確率変数 X と Y の共分散 (covariance) という . □

問 2.16 共分散 $\text{Cov}(X, Y)$ は期待値の性質を用いると

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

と書くことができることを示せ。

もし $X = Y$ ならば共分散は分散に等しい。また、 X と Y が独立ならば $E(XY) = E(X)E(Y)$ となるので、共分散は 0 である。確率変数 X と Y の共分散が 0 であるとき、 X と Y は無相関 (uncorrelated) であるという。しかし、 X と Y が無相関であっても独立であるとは限らない。

問 2.17 確率変数 X, Y の結合確率が

$$P(X = 1, Y = 1) = P(X = 2, Y = 0) = P(X = 3, Y = 1) = 1/3$$

で与えられているとする。このとき以下の問いに答えよ。

- (a) X, Y の周辺確率関数を求め、 X と Y が独立であるか否か、論じよ。
- (b) $E(X), E(Y), E(XY)$ を求め、 X と Y は無相関であるか否か、論じよ。

一般に n 個の確率変数 X_1, X_2, \dots, X_n の和の分散は共分散を用いて

$$V(X_1 + \dots + X_n) = \sum_{i=1}^n V(X_i) + 2 \sum_{i < j} \text{Cov}(X_i, X_j) \quad (21)$$

で与えられる。よって、 n 個の確率変数 X_1, X_2, \dots, X_n が互いに独立ならば、それらの和の分散はそれぞれの分散の和で与えられる。

問 2.18 式 (21) を導け。

問 2.19 X_1, X_2, \dots を独立同一な分布に従う平均 $E(X)$ 、分散 $V(X)$ をもつ確率変数とし、これらの最初の N 個の和 $S_N = X_1 + X_2 + \dots + X_N$ を考える。ただし N は X_i ($i = 1, 2, \dots$) とは独立な正の整数値をとる確率変数であり、その平均、分散はそれぞれ $E(N), V(N)$ で与えられる。このとき、 S_N の平均 $E(S_N)$ と分散 $V(S_N)$ を求めよ。

補題 2.1 (コーシー - シュワルツの不等式) $E(X^2), E(Y^2)$ が有限で正の値を取るならば

$$\{E(XY)\}^2 \leq E(X^2)E(Y^2)$$

が成立する。等号成立はある定数 a に対して $P(Y = aX) = 1$ となるときのみである。これはコーシー - シュワルツの不等式 (Cauchy-Schwarz inequality) と呼ばれている。□

任意の実数 t に対して $E((X - tY)^2) \geq 0$ であるので、 $E((X - tY)^2) = E(X^2) - 2E(XY)t + E(Y^2)t^2 \geq 0$ が成立する。これを t の 2 次関数とみれば判別式は非正でなければならない。よって $\{E(XY)\}^2 - E(X^2)E(Y^2) \leq 0$ を得る。また、等号が成り立つのは $E((X - tY)^2) = 0$ 、すなわち $P(X - tY = 0) = 1$ の場合に限る。

共分散の $\text{Cov}(X, Y)$ の取る値は、例えば X を a 倍すると a 倍となる。このように共分散は単位の取り方に依存する。そこでこのような単位の取り方に依存しないように正規化した指標が相関係数である。

定義 2.5 (相関係数) 確率変数 X, Y が有限でかつ 0 でない分散を持つとする。このとき

$$\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)}\sqrt{V(Y)}} \quad (22)$$

で定義される $\rho_{X,Y}$ を確率変数 X と Y の相関係数 (correlation coefficient) という。□

補題 2.1 より $-1 \leq \rho_{X,Y} \leq 1$ である¹³。特にある正数 a に対して $Y = aX$ ならば $\rho_{X,Y} = 1$ となり、 $Y = -aX$ ならば $\rho_{X,Y} = -1$ となる。

¹³補題 2.1 において、 X の代わりに $X - E(X)$ 、 Y の代わりに $Y - E(Y)$ とすると $\text{Cov}(X, Y)^2 \leq V(X)V(Y)$ を得る。

問 2.20 長さ x の間に到着する客数は平均 λx のポワソン分布に従うとする．すなわち，長さ X の間に到着する客数を N としたとき

$$P(N = k | X = x) = e^{-\lambda x} \frac{(\lambda x)^k}{k!}$$

である．以下の問いに答えよ．

- (1) $P(X = a) = 1$ であるとき， $E(N)$ ， $V(N)$ を求めよ．
- (2) ある実数 p ($0 < p < 1$) に対して $P(X = a) = p$ ， $P(X = b) = 1 - p$ であるとき， $P(N = k)$ を求めよ．
- (3) (2) の場合における $E(N)$ を求めよ．
- (4) $P(X = n) = (1 - p)p^{n-1}$ ($n = 1, 2, \dots$) であるとき， $E(N)$ を求めよ．
- (5) $P(X \leq x) = 1 - e^{-\mu x}$ ($x \geq 0$) であるとき， $E(X)$ ， $V(X)$ ， $P(N = k)$ を求めよ．
- (6) (5) の場合における $E(N)$ ， $V(N)$ を求めよ．ただし下記を用いて良い． $\sum_{k=0}^{\infty} k(k-1)(1-x)x^k = 2x^2/[(1-x)^2]$ ．
- (7) (5) の場合における $\text{Cov}(X, N)$ を求めよ．

問 2.21 確率変数 Z の密度関数 $f_Z(x)$ は 0 に対して対称であるとする．すなわち $f_Z(x) = f_Z(-x)$ ($x \geq 0$)．さらに Z の期待値が存在すると仮定する．また， X を正の値をとる確率変数とし， Z とは独立であると仮定する． $Y = Z/X$ としたとき $E(Y | X = x) = E(Y)$ であることを示せ．

問 2.22 ある工場で生産される機械は，稼働させてから故障するまでの時間が未知の平均 $1/\lambda$ ($\lambda > 0$) をもつ指数分布に従うことが知られている．故障するまでの平均時間を推定するために，この工場で生産された n 台 ($n \geq 2$) の機械を無作為に選び，それらを同時に稼働させ， i 番目 ($i = 1, 2, \dots, m$) に故障した機械が稼働していた時間 $X_{n,i}$ を計測する．ただし $2 \leq m \leq n$ である．定義より，確率 1 で $0 < X_{n,1} < X_{n,2} < \dots < X_{n,m}$ が成立し， m 台の機械が故障した時点において，残りの $n - m$ 台の機械はまだ故障せず，稼働中である．

上記の手順によって計測された故障を起こすまでの稼働時間 $X_{n,i}$ ($i = 1, 2, \dots, m$) を用いて，平均 $1/\lambda$ に対する推定量 $W_{n,m}$ を

$$W_{n,m} = a_{n,1}X_{n,1} + a_{n,2}X_{n,2} + \dots + a_{n,m}X_{n,m} \quad (23)$$

と定義する．ただし， $a_{n,i}$ ($i = 1, 2, \dots, m$) は実数である． $W_{n,m}$ の平均 $E[W_{n,m}]$ が $1/\lambda$ となり，かつ $W_{n,m}$ の分散 $V[W_{n,m}]$ が最小となるように $a_{n,i}$ ($i = 1, 2, \dots, m$) を定めたい．以下の設問 (a) ~ (e) に答えよ．

- (a) 平均 $1/\lambda$ をもつ指数分布の分散を求めよ．ただし，平均 $1/\lambda$ をもつ指数分布の確率分布関数 $F(x)$ ($x \geq 0$) は $F(x) = 1 - \exp(-\lambda x)$ で与えられる．
- (b) $Y_{n,i}$ ($i = 1, 2, \dots, m$) を次式で定義する．

$$Y_{n,1} = X_{n,1}, \quad Y_{n,i} = X_{n,i} - X_{n,i-1} \quad (i = 2, 3, \dots, m)$$

$Y_{n,i}$ ($i = 1, 2, \dots, m$) の確率分布関数 $F_{n,i}(x) = \Pr(Y_{n,i} \leq x)$ ($x \geq 0$) を求めよ．なお，非負の値を取る連続な確率変数の確率分布関数 $F(x)$ と確率密度関数 $f(x)$ に対して成立する以下の式を用いても良い．

$$\frac{(n-1)!}{(i-1)!(n-i)!} \int_0^{\infty} F^{i-1}(y)f(y)(1-F(y))^{n-i} dy = \frac{1}{n} \quad (i = 1, 2, \dots, n)$$

- (c) 設問 (b) で定義した $Y_{n,i}$ ($i = 1, 2, \dots, m$) は互いに独立な確率変数である．この事実を用いて， $W_{n,m}$ の分散 $V[W_{n,m}]$ を $Y_{n,i}$ ($i = 1, 2, \dots, m$) の分散 $V[Y_{n,i}]$ と $a_{n,i}$ ($i = 1, 2, \dots, m$) を用いて表せ．
- (d) 式 (23) で定義された $W_{n,m}$ の平均 $E[W_{n,m}]$ が $1/\lambda$ となり，かつ $W_{n,m}$ の分散 $V[W_{n,m}]$ が最小となるように $a_{n,i}$ ($i = 1, 2, \dots, m$) を定めよ．なお，以下の不等式を用いても良い．

$$\left(\sum_{i=1}^m x_i^2 \right) \left(\sum_{i=1}^m y_i^2 \right) \geq \left(\sum_{i=1}^m x_i y_i \right)^2, \quad x_i, y_i > 0 \quad (i = 1, 2, \dots, m)$$

ただし，この不等式における等号が成立するための必要十分条件は， $x_1/y_1 = x_2/y_2 = \dots = x_m/y_m$ である．

- (e) 平均 $1/\lambda$ を推定する際， $n = m$ とはせず， $n > m$ とする利点と欠点を述べよ．

3 確率変数と分布の収束

この章では確率変数ならびに分布の様々な収束の概念について解説する．これらの収束概念に対して正確な感覚を身につけることは，確率・統計における最も重要な定理である大数の法則や中心極限定理の意味するところを正しく理解するために必須である．

3.1 確率変数列の概収束と確率収束

1.3 節で述べたように，確率変数は標本空間 Ω から実数 \mathcal{R} への関数である．よって確率変数の収束は，数列の収束とは異なり，様々な定義がある．以下では $X_1(\omega), X_2(\omega), \dots$ を確率空間 (Ω, \mathcal{F}, P) で定義された確率変数として，代表的な確率変数の収束の定義を与える．

定義 3.1 (概収束) 事象 Ω_1 が $P(\Omega_1) = 1$ であるとする．このとき，任意の $\omega \in \Omega_1$ に対して

$$\lim_{n \rightarrow \infty} X_n(\omega) = X(\omega) \quad (24)$$

となるような確率変数 $X(\omega)$ が存在するとき，確率変数列 $\{X_n; n = 1, 2, \dots\}$ は確率変数 X へ概収束 (almost sure convergence) する，あるいは確率 1 で収束 (convergence with probability 1) するといい， $X_n \xrightarrow{a.s.} X$ または $\lim_{n \rightarrow \infty} X_n = X$ w.p.1 と書く． \square

概収束するということは，全ての $\omega \in \Omega_1$ に対して，各標本 ω 毎に収束するということであり，上で与えた定義は $P(\lim_{n \rightarrow \infty} X_n = X) = 1$ と等価である．すなわち，式 (24) を満たす ω 全体が形成する事象 $\{\omega; \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\}$ の確率は 1 である (このような事象を Ω_1 とした)．言い換えれば，もし $\lim_{n \rightarrow \infty} X_n(\omega) \neq X(\omega)$ であるような ω が存在するならば，そのような ω は $\omega \in \Omega_0 = \Omega - \Omega_1$ (すなわち ω は $P(\Omega_0) = 0$ であるような事象 Ω_0 の要素) である．

系 3.1 $X_n \xrightarrow{a.s.} X$ であるということは，全ての $\epsilon > 0$ に対して

$$\lim_{n \rightarrow \infty} P \left(\sup_{m \geq n} |X_m - X| \geq \epsilon \right) = 0$$

となることと等価である． \square

$A_n^{(\epsilon)} = \{\omega \mid |X_n(\omega) - X(\omega)| \geq \epsilon\}$, $A^{(\epsilon)} = \limsup_{n \rightarrow \infty} A_n^{(\epsilon)} = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k^{(\epsilon)}$ とする．このとき， $A = \{\omega \mid \lim_{n \rightarrow \infty} X_n(\omega) \neq X(\omega)\}$ とすると $A = \bigcup_{\epsilon > 0} A^{(\epsilon)} = \bigcup_{\epsilon > 0} [\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k^{(\epsilon)}] = \bigcup_{m=1}^{\infty} [\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k^{(1/m)}] = \bigcup_{m=1}^{\infty} A^{(1/m)}$ である．よって $X_n \xrightarrow{a.s.} X \Leftrightarrow P(A) = P(\bigcup_{m=1}^{\infty} A^{(1/m)}) = 0 \Leftrightarrow \forall m \geq 1, P(A^{(1/m)}) = 0 \Leftrightarrow \forall \epsilon > 0, P(A^{(\epsilon)}) = 0$.

事象 $\sup_{m \geq n} |X_m - X| \geq \epsilon$ は $|X_m - X| \geq \epsilon$ となる m ($m \geq n$) が少なくとも一つ存在するということである．

定義 3.2 (確率収束) 任意の正数 ϵ に対して

$$\lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0$$

が成立するとき，確率変数列 $\{X_n; n = 1, 2, \dots\}$ は確率変数 X へ確率収束 (convergence in probability) するといいい， $X_n \xrightarrow{P} X$ と書く． \square

定義より $X_n \xrightarrow{P} X$ は $\lim_{n \rightarrow \infty} P(|X_n - X| \leq \epsilon) = 1$ と等価である．確率収束は，概収束のような関数の収束を意味しているのではなく，実数列 $P(|X_n - X| > \epsilon)$ ($n = 1, 2, \dots$) の 0 への収束を意味している．

例 3.1 確率空間 $(R, \mathcal{B}(R), P)$ において， P は区間 $[0, 1]$ 上の一様分布とする．この確率空間上で確率変数列

$\{X_n; n = 1, 2, \dots\}, \{Y_n; n = 1, 2, \dots\}$ を以下のように定義する .

$$X_n(\omega) = \begin{cases} 1, & \omega \in [0, \frac{1}{n}] \\ 0, & \text{その他} \end{cases}$$

$$Y_n(\omega) = \begin{cases} 1, & \omega \in \left[\frac{n \bmod 2^i}{2^i}, \frac{n \bmod 2^i + 1}{2^i} \right] \\ 0, & \text{その他} \end{cases} \quad (\text{ただし } 2^i \leq n < 2^{i+1})$$

ここで $n \bmod m$ は n を m で割ったときの余りを表す . 確率変数列 $\{X_n; n = 1, 2, \dots\}$ は任意の選ばれたある $\omega \in [0, 1]$ に対して

$$\lim_{n \rightarrow \infty} X_n(\omega) = \begin{cases} 1, & \omega = 0, \\ 0, & \text{その他,} \end{cases} \quad P(|X_n| \geq \epsilon) = \frac{1}{n} \quad (0 < \epsilon < 1)$$

である . $P(\{0\}) = 0$ であるので確率変数列 $\{X_n; n = 1, 2, \dots\}$ は

$$X(\omega) = 0, \quad \omega \in \mathcal{R}$$

へ概収束する . また $\lim_{n \rightarrow \infty} P(|X_n| \geq \epsilon) = 0$ なので $\lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0$ となり , 確率変数列 $\{X_n; n = 1, 2, \dots\}$ は X へ確率収束する .

一方 , 任意の選ばれたある $\omega \in [0, 1]$ に対して確率変数列 $\{Y_n; n = 1, 2, \dots\}$ は

$$\limsup_{n \rightarrow \infty} Y_n(\omega) = 1, \quad \liminf_{n \rightarrow \infty} Y_n(\omega) = 0$$

となるので極限は存在せず , 概収束しない . しかし , 任意の $0 < \epsilon < 1$ に対して , $2^i \leq n < 2^{i+1}$ のとき

$$P(|Y_n - X| > \epsilon) = P(Y_n \neq 0) = \frac{1}{2^i}$$

となるので , $n \rightarrow \infty$ のとき $2^i \rightarrow \infty$ となることに注意すれば , $\lim_{n \rightarrow \infty} P(|Y_n - X| > \epsilon) = 0$ となり , 確率変数列 $\{Y_n; n = 1, 2, \dots\}$ は X へ確率収束する . □

この例から分かるように概収束は確率収束よりも強い収束である . 実際 , 次の定理が成立する .

定理 3.1 (概収束 \Rightarrow 確率収束) $X_n \xrightarrow{a.s.} X$ ならば $X_n \xrightarrow{P} X$ である . □

概収束 , 確率収束以外の重要な確率変数収束概念として次の法則収束がある .

定義 3.3 (法則収束) 確率変数列 $\{X_n; n = 1, 2, \dots\}$ と確率変数 X があるとする . X_n, X の分布関数をそれぞれ $F_n(x), F(x)$ としたとき , $F(x)$ が連続である全ての x で

$$\lim_{n \rightarrow \infty} F_n(x) = F(x)$$

が成立するならば , 確率変数列 $\{X_n; n = 1, 2, \dots\}$ は確率変数 X へ法則収束 (convergence in law, convergence in distribution) するといい , $X_n \xrightarrow{D} X$ と書く . また $X_n \xrightarrow{D} X$ であるとき , X の分布 $F(x)$ を X_n の漸近分布 (asymptotic distribution) または極限分布 (limiting distribution) という . □

例 3.2 (例 3.1 の続き) 例 3.1 で与えた非負確率変数 $X_n(\omega)$ の分布関数 $F_n(x)$ は

$$F_n(x) = \begin{cases} 0, & x < 0 \\ \frac{n-1}{n}, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

であるので , 極限分布

$$F(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

をもつ。すなわち $X_n \xrightarrow{D} X$ である。また、例 3.1 で与えた非負確率変数 $Y_n(\omega)$ の分布関数 $G_n(x)$ は $2^i \leq n < 2^{i+1}$ なる n に対して

$$G_n(x) = \begin{cases} 0, & x < 0 \\ 1 - \frac{1}{2^i}, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

なので、極限分布

$$F(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

をもつ。すなわち $Y_n \xrightarrow{D} X$ である。 □

一方、法則収束するが確率収束しない例は簡単に作る事が出来る。

例 3.3 確率空間 $(R, \mathcal{B}(R), P)$ において、 P は区間 $[0, 1]$ 上の一様分布とする。この確率空間上で確率変数列 $\{Z_n; n = 1, 2, \dots\}$ を以下のように定義する。

$$n \text{ が奇数のとき} : Z_n(\omega) = \begin{cases} 1 & \omega \in [0, 1/2] \\ 0, & \text{その他} \end{cases} \quad n \text{ が偶数のとき} : Z_n(\omega) = \begin{cases} 1 & \omega \in (1/2, 1] \\ 0, & \text{その他} \end{cases}$$

定義より $P(Z_n = 0) = P(Z_n = 1) = 1/2$ であるので、全ての n に対して Z_n は同じ分布をもつ。よって $Z_n \xrightarrow{D} Z_1$ である。しかし n が偶数のとき $P(|Z_n - Z_1| > \epsilon) = 1$ となり、確率収束の意味では収束しない。 □

このように法則収束は確率収束よりも弱い収束であり、実際、確率収束するならば法則収束することが知られている。

定理 3.2 (確率収束 \Rightarrow 法則収束) $X_n \xrightarrow{P} X$ ならば $X_n \xrightarrow{D} X$ である。 □

x を $F(x)$ の連続点とする。このとき任意の正数 ϵ に対して $F_n(x) - F(x) = P(X_n \leq x) - P(X \leq x) = P(X_n \leq x, X > x) + P(X_n \leq x, X \leq x) - P(X \leq x) \leq P(X_n \leq x, X > x) \leq P(X_n \leq x, X > x + \epsilon) + P(x < X \leq x + \epsilon) \leq P(|X_n - X| > \epsilon) + F(x + \epsilon) - F(x)$ となり、 $\epsilon \rightarrow 0$ とすると $\lim_{n \rightarrow \infty} F_n(x) - F(x) \leq 0$ を得る。同様に $F(x) - F_n(x) = P(X \leq x) - P(X_n \leq x) = P(X \leq x, X_n > x) + P(X \leq x, X_n \leq x) - P(X_n \leq x) \leq P(X \leq x, X_n > x) \leq P(X \leq x - \epsilon, X_n > x) + P(x - \epsilon < X \leq x) \leq P(|X_n - X| > \epsilon) + F(x) - F(x - \epsilon)$ となり、 $\epsilon \rightarrow 0$ とすると $\lim_{n \rightarrow \infty} F(x) - F_n(x) \leq 0$ を得る。

補題 3.1 ある定数 c に対して

$$X_n \xrightarrow{D} c \Leftrightarrow X_n \xrightarrow{P} c$$

である。 □

定理 3.2 より $X_n \xrightarrow{P} c$ ならば $X_n \xrightarrow{D} c$ である。そこで $X_n \xrightarrow{D} c \Rightarrow X_n \xrightarrow{P} c$ を示す。 X_n の分布関数を $F_n(x)$ 、 c の分布関数を $F(x)$ とする。すなわち $F(x) = 0 (x < c)$ 、 $F(x) = 1 (x \geq c)$ である。もし $X_n \xrightarrow{D} c$ ならば、全ての正数 ϵ に対して $\lim_{n \rightarrow \infty} F_n(c - \epsilon) = F(c - \epsilon) = 0$ であり、 $\lim_{n \rightarrow \infty} F_n(c + \epsilon) = F(c + \epsilon) = 1$ である。よって $\lim_{n \rightarrow \infty} P(|X_n - c| \geq \epsilon) = \lim_{n \rightarrow \infty} P(X_n \geq c + \epsilon) + \lim_{n \rightarrow \infty} P(X_n \leq c - \epsilon) = 1 - \lim_{n \rightarrow \infty} P(X_n < c + \epsilon) + \lim_{n \rightarrow \infty} F(c - \epsilon) = 0$ 。

3.2 大数の法則

多くの回数、同じ条件で確率的実験を繰り返したときに得られる数値の平均は、ある値に近付いていくことが知られている。一般にこのような性質は大数の法則 (law of large numbers) と呼ばれている。大数の法則が成り立つ条件は色々知られているが、ここでは確率変数列 $\{X_n; n = 1, 2, \dots\}$ が互いに独立な場合の結果について紹介する。以下では確率変数列 $\{X_n; n = 1, 2, \dots\}$ に対して最初の n 個の確率変数の和を

$$S_n = \sum_{k=1}^n X_k$$

で表す。定義より $E(S_n) = \sum_{k=1}^n E(X_k)$ であり, $\{X_n; n = 1, 2, \dots\}$ が互いに独立な場合には $V(S_n) = \sum_{k=1}^n V(X_k)$ である。

定理 3.3 (大数の弱法則 (1)) $\{X_n; n = 1, 2, \dots\}$ を独立な確率変数列とし, $V(X_n) = \sigma_n^2 < \infty$ であるとする。もし, $\lim_{n \rightarrow \infty} \sum_{k=1}^n \sigma_k^2/n^2 = 0$ ならば,

$$\frac{S_n - E(S_n)}{n} \xrightarrow{P} 0$$

が成立する。これをチェビシェフの大数の弱法則 (weak law of large numbers) という。□

式 (18) で与えたチェビシェフの不等式を $(S_n - E(S_n))/n$ へ適用すると $P(|S_n - E(S_n)| \geq n\epsilon) \leq \frac{1}{n^2\epsilon^2} \sum_{k=1}^n \sigma_k^2$ を得る。よって仮定より, $\lim_{n \rightarrow \infty} P(|S_n - E(S_n)|/n \geq \epsilon) = 0$ 。

分散に関する条件 $\lim_{n \rightarrow \infty} \sum_{k=1}^n \sigma_k^2/n^2 = 0$ が成立する例として, 全ての k に対して $\sigma_k^2 \leq K$ となるような定数 K が存在する場合などがある。

大数の弱法則 (1) は分散が有限であるという制約があった。しかし, この制約は外すことができる。

定理 3.4 (大数の弱法則 (2)) $\{X_n; n = 1, 2, \dots\}$ を独立で同一な分布に従う確率変数列とする。このとき, $E(X_1) < \infty$ ならば

$$\frac{S_n}{n} \xrightarrow{P} E(X_1)$$

が成立する。□

この定理の証明は切断 (truncation) と呼ばれる技法を用いて, 無限の分散をもつ確率変数を有限な分散をもつ確率変数へ変換することで行われる。詳細は付録 B.1 を参照すること。

上記の二つの大数の弱法則はいずれも確率収束を与えていた。一方, この収束が概収束の意味で成り立つとき, 大数の強法則 (strong law of large numbers) という。

定理 3.5 (大数の強法則) $\{X_n; n = 1, 2, \dots\}$ を独立で同一な分布に従う確率変数列とする。もし, $E(X_1) < \infty$ ならば

$$\frac{S_n}{n} \xrightarrow{a.s.} E(X_1)$$

が成立する。□

証明するには多くの準備を必要とするため省略する (例えば [2] を参照)。

大数の弱法則と大数の強法則の違い, すなわち確率収束と概収束の違いをしっかりと理解する必要がある。系 3.1 より大数の強法則は

$$\lim_{n \rightarrow \infty} P \left(\sup_{m \geq n} \left| \frac{S_m}{m} - E(X_1) \right| \geq \epsilon \right) = 0$$

と書くことが出来る。ここで事象 $\{\sup_{m \geq n} |S_m/m - E(X_1)| > \epsilon\}$ は $|S_m/m - E(X_1)|$ の値が ϵ より大きくなるような m ($m \geq n$) が少なくとも一つあるということである。よって大数の強法則が示していることは, 各標本において, n が十分に大きければ, n 以上の m で $|S_m/m - E(X_1)| > \epsilon$ となってしまうことはほとんどないということである。言い換えれば, このことは $S_1/1, S_2/2, \dots, S_n/n, \dots$ という列が n が増加するにつれて, 徐々に, それより大きな $m > n$ のいずれの項も平均との差が ϵ より大きくなるならないようになるということである。一方, 大数の弱法則がいうことは n が増加するにつれて, 徐々に, S_n/n と平均との差が ϵ より大きくなるならないようになるということである。このように大数の強法則は列 $\{S_n/n; n = 1, 2, \dots\}$ の最初の $n-1$ 個を除いた残り全てに関する性質に言及しているのに対し, 大数の弱法則は項 S_n/n それ自身に対する性質を述べている。

例 3.4 例 1.1 を拡張して, コインを無限に投げ続ける実験を考える。 X_n ($n = 1, 2, \dots$) を n 回目に表が出れば 1, 裏が出れば 0 の値を取る確率変数とする。このとき S_n/n は確率 1 で $1/2$ へ収束する。一方, S_n/n が $1/2$ でないような標本は簡単に作る事が出来る。例えば全て表が出る, 全て裏が出る, 全体の $1/3$ だけ表が出る, などである。大数の強法則は, このような S_n/n が $1/2$ に収束しないような標本全体からなる事象は確率 0 をもつ

ということを主張している．一方，大数の弱法則は n を大きくすれば事象 $\{|S_n/n - 1/2| \geq \epsilon\}$ が起こる確率をいくらでも小さくできるということを主張している． \square

3.3 分布の弱収束と中心極限定理

確率変数列の法則収束は，対応する分布関数列の収束と考えることが出来る．このような意味で定義された分布関数の収束を弱収束 (weak convergence) といい， $F_n(x) \xrightarrow{D} F(x)$ と書く．法則収束の定義から明らかなように，弱収束では $F(x)$ が連続であるような点でのみ収束が要求されており， $F(x)$ の不連続点での収束は要求されていない．

例 3.5 以下で定義される分布関数 $F_n(x)$ ， $F(x)$ を考える．

$$F_n(x) = \begin{cases} 0, & x < 0 \\ nx, & x \in [0, 1/n] \\ 1, & x > 1/n \end{cases} \quad F(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

明らかに $F_n(x)$ は $x = 0$ を除いて $F(x)$ へ収束するが， $x = 0$ では収束していない．しかし， $x = 0$ は $F(x)$ の不連続点であるので，定義より $F_n(x) \xrightarrow{D} F(x)$ である． \square

分布関数が弱収束することと等価な条件は幾つか知られており，例えば次のようなものがある．

定理 3.6 $F_n(x) \xrightarrow{D} F(x)$ は，任意の有界で連続な関数 $g(x)$ に対して

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} g(x) dF_n(x) = \int_{-\infty}^{\infty} g(x) dF(x)$$

が成立することと等価である． \square

弱収束を用いて表現された非常に重要な定理に中心極限定理がある．

定理 3.7 (中心極限定理) $\{X_n; n = 1, 2, \dots\}$ を独立で同一な分布に従う確率変数列とする．もし， $\mu = E(X_1) < \infty$ かつ $\sigma^2 = V(X_1)$ が $0 < \sigma^2 < \infty$ を満たすならば， $S_n = \sum_{k=1}^n X_k$ としたとき

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq x\right) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du \quad (25)$$

が成立する．これを中心極限定理 (central limit theorem) という． \square

定理 3.7 の証明は割愛する．

中心極限定理 (25) の左辺は平均 0，分散 1 をもつ標準正規分布 $N(0, 1)$ の分布関数である．よって

$$\lim_{n \rightarrow \infty} P\left(a \leq \frac{1}{\sqrt{n}\sigma} \sum_{k=1}^n (X_k - \mu) \leq b\right) = \int_a^b \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) du$$

と等価である．ガウス (Gauss) は実験から得られたデータの期待値からのバラツキに一定の法則があることを見出した．それゆえ正規分布はガウス分布 (Gaussian distribution) とも呼ばれる．

この節を終えるにあたり，中心極限定理の意味するところについて述べる．まず， S_n が平均 $n\mu$ ，分散 $n\sigma^2$ をもつことに注意する．中心極限定理 (25) の左辺に現れる

$$Z_n = (S_n - n\mu)/(\sqrt{n}\sigma)$$

は式 (19) に従って S_n を標準化したものであり，全ての自然数 n に対して平均 0，分散 1 をもつ確率変数である．中心極限定理はこの確率変数 Z_n の従う分布が，確率変数 S_n の分布，すなわち，確率変数 X_1 の分布が如何

なるものであろうとも、有限の平均と分散をもつならば、 n を増加させると標準正規分布 $N(0, 1)$ に弱収束するということを主張している。すなわち、 n を増加させると $(S_n - n\mu)/\sqrt{n}$ は平均 0、分散 σ^2 をもつ正規分布に弱収束するということである。

中心極限定理は確率変数 S_n が必ずしも（標準正規分布 $N(0, 1)$ の密度関数で近似できるような）密度関数をもつことを主張しているわけではない。実際、 X_1 が離散の確率変数ならば全ての n に対して確率変数 S_n は密度関数をもたない。このような場合、大きな n に対して、その確率関数は非常に細かな階段状の変化をする形状をもち、その形状は積分する（すなわち分布関数を考える）ことで平滑化されることになる。

絶対値が非常に大きい負の値 x に対しては、式 (25) の右辺は非常に 0 に近い値となる。よって n がそれほど大きくなければ、右辺と（左辺に現れる）確率 $P(Z_n \leq x)$ の差の絶対値は小さいが、その比は 1 から懸け離れたものになっていると思われる。特に非負の確率変数の和を考えると、全ての標本に対して $S_n \geq 0$ であるが、正規分布の定義域は $(-\infty, \infty)$ であるので、有限の n に対して $x < -n\mu/(\sqrt{n}\sigma) = -\sqrt{n}\mu/\sigma$ では明らかに近似的評価として採用できない。また同様に、非常に大きい正数 x に対しては右辺と $P(Z_n \leq x)$ の両方とも 1 に非常に近く、この場合も右辺は左辺に現れる確率の良い近似とはなっていないと思われる。すなわち左辺に現れる確率は適当な大きさの n に対しては、 x が 0 に比較的近い場合にのみ良い近似と見なせると考えられ、この定理が極限定理ではなく、中心極限定理と呼ばれている理由もそこにある。

前節で紹介した大数の法則や中心極限定理では独立な確率変数列を仮定していた。実は、大数の法則は独立でない確率変数列に対しても、一定の条件下で成立する。この意味で、大数の法則の適用範囲は（中心極限定理に比べて）相当広い。一方、独立でないような確率変数列に対する中心極限定理はほとんど知られていない。すなわち、中心極限定理においては独立性が非常に重要な仮定となっており、中心極限定理を応用で用いる際には、この点に十分注意を払う必要がある。

4 高校までの確率の復習用問題

4.1 基本問題

問 4.1 袋の中にくじが m 本 ($m \geq 3$) あり、そのうち当たりくじは n 本 ($2 \leq n \leq m - 1$) である。この袋から同時にくじを 3 本引くとき、2 本が当たりくじである確率を求めよ。

問 4.2 二つのサイコロ A, B を投げ、A の目を a , B の目を b としたとき、(i) b/a が整数になる確率、(ii) $a + b = 6$ となる確率、(iii) $ab = 12$ となる確率、(iv) $a + b$ の期待値、(v) ab の期待値を求めよ

問 4.3 サイコロを N 回投げたとき、(i) 出る目の最大値 M が 5 以下である確率と (ii) 出る目の最大値 M が 5 である確率を求めよ。

問 4.4 サイコロを 3 回投げたとき、出る目の和が 16 以下である確率を求めよ。

問 4.5 x 軸上を移動する点 P は、最初、原点 O にあり、サイコロを投げるたびに、5 以上の目が出れば +1、4 以下の目が出れば -2 だけ x 軸上を動くとき、(i) サイコロを 3 回投げたとき、点 P が原点 O にある確率と、(ii) サイコロを 5 回投げたとき、点 P が $x = -1$ の位置にある確率を求めよ。

問 4.6 男 M 人 ($M \geq 2$)、女 N 人がくじ引きで順序を決めて一列に並ぶとき、両端が男である確率を求めよ。

問 4.7 M 本のくじがあり、そのうち N 本は当たりである。1 本ずつ、くじを合計 K 本 ($K \leq M - N$) 引くとき、(i) 最後に引いたくじが当たる確率と (ii) K 本のうち、少なくとも 1 本が当たる確率を求めよ。

問 4.8 N 人 ($N \geq 3$) でじゃんけんを 1 回だけするとき、(i) 一人が勝つ確率と (ii) 二人が勝つ確率を求めよ。

問 4.9 正確なコインを繰り返し投げるとき、裏が M 回出るまでに表が N 回出る確率を求めよ。

問 4.10 箱の中に 1 等から N 等 ($N \geq 2$) までのくじがそれぞれ 1 枚ずつ合計 N 枚入っている。この箱から無作為にくじを 1 枚取り出し、等級を確かめて箱に戻す操作を N 回繰り返す。(i) 1 等から N 等までのくじがそれぞれ 1 回ずつ取り出される確率と (ii) 1 等が k 回 ($1 \leq k \leq N-1$)、2 等が $N-k$ 回取り出される確率を求めよ。

4.2 練習問題

問 4.11 箱の中に 1 等から 3 等までそれぞれ 1 本ずつ、3 本入っている。この袋からくじを引いて等級を確かめた後、箱に戻す操作を 3 回行う。取り出されたくじの種類期待値を、全ての場合を書き出さずに、計算のみで求めよ。

問 4.12 袋の中に 1 から m ($m \geq 2$) までの番号が書かれたカードがそれぞれ 1 枚ずつ、 m 枚入っている。この袋からカードを引いて等級を確かめた後、箱に戻す試行を繰り返し、以下のいずれかの条件が成立すれば停止する。

a. 試行の回数が $n-1$ 以下で、かつ、引いたカードに書かれている数字の合計が m の倍数になったとき

b. (a. で終了せず) 試行の回数が n 回になったとき

ただし、 n は 2 以上の整数である。試行回数の期待値を求めよ。

問 4.13 A, B, C の 3 人が $A \rightarrow B \rightarrow C \rightarrow A \rightarrow B \dots$ の順に正確なコインを投げ、最初に表を出した者を勝ちとする。(i) A が勝つ確率と (ii) B が勝つ確率を求めよ。

問 4.14 正確なコインを繰り返し n 回投げたとき、表が 2 回以上続けて出ることがなく、かつ、最後が表(裏)である確率を a_n (b_n) とする。(i) a_n , b_n に関する漸化式を考え、(ii) n 回投げたとき表が 2 回以上続けて出ることがない確率 $p_n = a_n + b_n$ ($n \geq 3$) を p_{n-1}, p_{n-2} を用いて表せ。

問 4.15 数直線上に 3 点 L, M, R があり、最初、点 M に玉が置かれている。サイコロを繰り返し投げ、以下の規則で玉の位置を動かす。

- ・玉が点 M にあるとき、1, 2 が出れば点 L へ、5, 6 が出れば点 R へ動かし、3, 4 が出れば点 M のまま動かさない
 - ・玉が点 L にあるとき、奇数の目が出たら点 M へ動かし、偶数の目が出たら点 L のまま動かさない
 - ・玉が点 R にあるとき、奇数の目が出たら点 M へ動かし、偶数の目が出たら点 R のまま動かさない
- n 回の試行の後、玉が点 L, M, R にある確率をそれぞれ l_n, m_n, r_n とする。これらの確率の間に成立する漸化式を考えることにより、 m_n を n を用いて表せ。

問 4.16 当たりくじ M 本を含む N 本のくじがある ($M \leq N$)。くじを 1 本引いて当たりか外れかを確認したのちに元に戻す試行を当たりくじが K 回出るまで繰り返すとき、ちょうど n 回目の試行で終了する確率 p_n とする。(i) p_n ($n \geq K$) を求めたのち、(ii) M と $N(K-1)$ が互いに素であるという仮定の下で、 p_n を最大にする n を求めよ。

問 4.17 サイコロを投げて 5 以上の目が出れば 2 点、4 以下の目が出れば 1 点を得る。サイコロを n 回投げたときまでに得た得点の合計が偶数である確率 p_n を求めよ。

問 4.18 表が出る確率が $2/3$ であるコインを、2 回連続して表が出るまで投げ続けるとき、 n 回 ($n \geq 2$) で終了する確率 p_n を求めよ。

問 4.19 正方形の頂点を A, B, C, D とし、この順を正の向きとし、逆を負の向きとする。動点 P は常に頂点にあり、1 秒毎に、次の頂点に写っていく。このとき、正の向きに次の頂点に移る確率は $2/3$ であり、負の向きに次の頂点へ移る確率は $1/3$ である。動点 P は最初、頂点 A にある。 a_n, b_n, c_n, d_n をそれぞれ n 秒後に動点 P が頂点 A, B, C, D にある確率とする。 n が奇数か偶数かで場合分けをし、 a_n, b_n, c_n, d_n を求めよ。

問 4.20 n を 3 以上の整数, k を $1 \leq k \leq n$ を満たす整数とする. 箱の中に k 個の白玉と $n - k$ 個の赤玉が入っている. 箱から玉を一つ取り出し

- ・白玉であれば, この白玉を箱に戻し, さらに箱の中の赤玉を一つ白玉に取り替え,
- ・赤玉であれば, この赤玉を箱に戻し, さらに箱の中の白玉を一つ赤玉に取り替える

という操作を, 箱の中の玉が全て白玉になるか, あるいは全て赤玉になるまで繰り返す. 箱の中の玉が全て白玉になる確率を p_k としたとき, (i) p_k を p_{k-1} と p_{k+1} で表せ. ただし $p_0 = 0, p_n = 1$ とする. (ii) さらに p_k を p_1, n, k で表し, p_1 を求めよ.

問 4.21 箱の中に 1 等から N 等 ($N \geq 2$) までのくじがそれぞれ 1 枚ずつ合計 N 枚入っている. この箱から無作為にくじを 1 枚取り出し, 等級を確かめて箱に戻す操作を繰り返す. この操作を n 回繰り返したとき, 取り出したくじの等級が k 種類するとき $X_n = k$ と定める. たとえば $n = 4$ において 1,2,1,1 ならば $X_n = 2$ である. $p_n(k)$ ($n = 1, 2, \dots, N, k = 1, 2, \dots, n$) を $X_n = k$ となる確率とする. $p_{n+1}(k), p_n(k), p_n(k-1)$ の間に成立する関係式を示せ.

問 4.22 n を自然数とする. n 個の箱すべてに, 1, 2, 3, 4, 5 の五種類のカードがそれぞれ 1 枚ずつ, 計 5 枚入っている. 各々の箱から 1 枚ずつカードを取り出し, 取り出した順に右から並べて n 桁の数 X_n を作る. 自然数 n に対して, X_n が 3 の倍数となる確率を p_n とする. (a) p_{n+1} を p_n で表せ. (b) p_n を求めよ.

A Riemann 積分と Riemann-Stieltjes 積分について

A.1 Riemann 積分

ここでは [4] に沿って Riemann 積分を説明する. これは, いわゆる区分求積法と呼ばれているものであり, 定積分が面積に対応することを明確に表している.

まず, 閉区間 $[a, b]$ に対して

$$a = x_0 < x_1 < \dots < x_{n-1} < x_n = b$$

となるような有限個の点 x_i ($i = 0, 1, \dots, n$) を取ることをこの区間の分割といい, 点 x_i を分点, 閉区間 $[x_i, x_{i-1}]$ ($i = 1, \dots, n$) を小区間とよぶ. さらにこの分割を Δ で表し, 小区間の長さを Δx_i と記す.

$$\Delta x_i = x_i - x_{i-1}, \quad i = 1, 2, \dots, n$$

明らかに

$$\sum_{i=1}^n \Delta x_i = b - a$$

である. 閉区間 $[a, b]$ の分割 Δ に対して, さらにいくつかの分点を追加して得られる分割を Δ の細分という. 閉区間 $[a, b]$ の任意の二通りの分割 Δ, Δ' に対して, 常にこれらの共通の細分 Δ'' を作ることができる. 実際, Δ, Δ' の分点を合わせたものを Δ'' の分点とすればよい.

閉区間 $[a, b]$ で定義された関数 $u(x)$ が与えられたとき, 閉区間 $[a, b]$ の分割 Δ および各小区間 $[x_{i-1}, x_i]$ 上の点 ξ_i ($i = 1, 2, \dots, n$) をとり, 和

$$S(\Delta) = \sum_{i=1}^n u(\xi_i) \Delta x_i$$

を作る. これを Riemann 和という. 小区間の長さ Δx_i ($i = 1, 2, \dots, n$) のうち最大のを $|\Delta|$ としたとき, 分割 Δ および点 ξ_i の取り方に関係なく, 極限值

$$\lim_{|\Delta| \rightarrow 0} S(\Delta) = A$$

が定まるとき，関数 $u(x)$ は $[a, b]$ において積分可能であるといい，この値を $f(x)$ の $[a, b]$ における定積分，または a から b までの定積分という．そして，これを

$$A = \int_a^b u(x) dx$$

で表す．

関数 $u(x)$ が区間 $[a, b]$ において積分可能であるか否かは，以下のようにして検証できる． m_i と l_i ($i = 1, \dots, n$) をそれぞれ小区間 $[x_i, x_{i-1}]$ における関数 $u(x)$ の上界値，下界値を与える点とし，これらの点を分点とした Riemann 和

$$S(\Delta) = \sum_{i=1}^n u(\xi_i) \Delta x_i, \quad M(\Delta) = \sum_{i=1}^n u(m_i) \Delta x_i, \quad L(\Delta) = \sum_{i=1}^n u(l_i) \Delta x_i$$

を作れば， $M(\Delta)$ ， $L(\Delta)$ の値は分割 Δ によって定まり

$$L(\Delta) \leq S(\Delta) \leq M(\Delta)$$

となる．さらに，分割 Δ の細分 Δ'' をとれば， $M(\Delta)$ ， $L(\Delta)$ の作り方から

$$L(\Delta) \leq L(\Delta''). \quad M(\Delta'') \leq M(\Delta)$$

となる．よって

$$\lim_{|\Delta| \rightarrow 0} |M(\Delta) - L(\Delta)| = 0$$

となることを示せばよい．

そこで，関数 $u(x)$ は閉区間 $[a, b]$ で連続¹⁴であると仮定する．有界閉集合において連続な関数は，その集合内で必ず最大値と最小値をもち，かつ，一様連続¹⁵であることが知られている．よって次式が成立する．

$$\min_{x \in [a, b]} u(x) \cdot (b - a) \leq L(\Delta), \quad M(\Delta) \leq \max_{x \in [a, b]} u(x) \cdot (b - a)$$

閉区間 $[a, b]$ の任意の二通りの分割 Δ ， Δ' に対して，これらの共通の細分 Δ'' をとれば

$$L(\Delta) \leq L(\Delta'') \leq M(\Delta'') \leq M(\Delta')$$

である．すなわち，閉区間 $[a, b]$ のあらゆる分割 Δ に対して， $L(\Delta)$ は上に有界であり，その上限 L が存在する．同様に $M(\Delta)$ は下に有界であり，その下限 M が存在する．

$$L(\Delta) \leq L \leq M \leq M(\Delta)$$

$u(x)$ は閉区間 $[a, b]$ において一様連続であるので，任意の正数 ϵ に対して正数 δ が存在し

$$x, x' \in [a, b], |x - x'| < \delta \Rightarrow |u(x) - u(x')| < \epsilon$$

となる．よって， $|\Delta| < \delta$ となる任意の分割 Δ をとれば， Δ の小区間 $[x_{i-1}, x_i]$ において， $u(x)$ を最大，最小とする点 m_i, l_i に対しても

$$0 \leq u(m_i) - u(l_i) < \epsilon$$

となる．この不等式の各辺に小区間の長さ Δx_i を掛けて，全ての i について加えれば

$$0 \leq M(\Delta) - L(\Delta) < (b - a)\epsilon$$

¹⁴点集合 D を定義域にもつ関数 $u(x)$ が与えられたとし， D 上の 1 点 c を考える．任意の正数 ϵ に対して，正数 $\delta = \delta(c, \epsilon)$ が存在し， $|x - c| < \delta$ ならば $|u(x) - u(c)| < \epsilon$ となるとき，関数 $u(x)$ は点 c において連続であるという．さらに，関数 $u(x)$ が定義域 D の全ての点において連続であるとき， $u(x)$ は D において連続であるという．

¹⁵連続の定義において， δ を ϵ のみの関数として定めることができる場合，一様連続という．正確には下記の通りである．点集合 D を定義域にもつ関数 $u(x)$ が与えられたとき，任意の正数 ϵ に対して，正数 $\delta = \delta(\epsilon)$ が存在し， $x, x' \in D$ かつ $|x - x'| < \delta$ ならば $|u(x) - u(x')| < \epsilon$ となるとき，関数 $u(x)$ は D において一様連続であるという．

となる．よって $\epsilon \rightarrow 0$ の極限を取れば

$$\lim_{|\Delta| \rightarrow 0} |M(\Delta) - L(\Delta)| = 0, \quad L = M = \lim_{|\Delta| \rightarrow 0} S(\Delta)$$

となり，閉区間 $[a, b]$ において連続な関数 $u(x)$ の定積分は存在する．

上記の証明から以下のことが分かる．関数 $u(x)$ が开区間 (a, b) で一様連続であれば， $u(x)$ は $[a, b]$ において積分可能である．この場合，区間の両端 a, b における $u(x)$ の値は，この定積分とは全く無関係である．一般に，関数 $u(x)$ は，閉区間 $[a, b]$ において，有限個の不連続点をもつものとし，隣り合う二つの不連続点の間では一様連続となっていれば， $u(x)$ は $[a, b]$ において積分可能である．

A.2 Riemann-Stieltjes 積分

次に，[5] に沿って，簡単に Riemann-Stieltjes 積分の概略を紹介する．

閉区間 $[a, b]$ で定義された関数 $u(x)$ と区間 $[a, b]$ で定義された右連続単調非減少関数 $F(x)$ を考える．閉区間 $[a, b]$ の分割は Riemann 積分の場合と同様に定義されているものとする．さらに， $\Delta F_i = F(x_i) - F(x_{i-1})$ とする．このとき，上界，下界の重み付き和

$$L(\Delta) = \sum_{i=1}^n u(l_i) \Delta F_i, \quad M(\Delta) = \sum_{i=1}^n u(m_i) \Delta F_i$$

が， $|\Delta| \rightarrow 0$ の極限において一致する時， $u(x)$ の $F(x)$ に関する Riemann-Stieltjes 積分が以下のように定義される．

$$\int_a^b u(x) dF(x) = \lim_{|\Delta| \rightarrow 0} L(\Delta) = \lim_{|\Delta| \rightarrow 0} M(\Delta)$$

以下のことに注意する．ある点 $\tau \in [a, b]$ で $u(x)$ は連続であるが $F(x)$ が不連続な場合， $n \rightarrow \infty$ の極限において点 τ における上界あるいは下界の積和への貢献は $u(\tau)(F(\tau) - F(\tau-))$ となる．ただし， $F(\tau-)$ は $F(x)$ の $x = \tau$ における左極限 (x を左から τ に近付けた時の極限) である．よって， $F(x)$ が x_i で不連続となる階段関数ならば， $x = x_i$ において定義可能な $u(x)$ に対して

$$\int_a^b u(x) dF(x) = \sum_{i=1}^n u(x_i) \Delta F_i$$

が成立する．

Riemann 積分は $F(x) = x$ の場合に対応する．もし，区間 $[a, b]$ で $F(x)$ が微分可能であれば，Riemann-Stieltjes 積分は Riemann 積分で表現できる．

$$\int_a^b u(x) dF(x) = \int_a^b u(x) \frac{dF(x)}{dx} dx$$

なお，異常積分は Riemann 積分と同様に極限として定義される．

$$\int_a^\infty u(x) dF(x) = \lim_{b \rightarrow \infty} \int_a^b u(x) dF(x)$$

B 定理の証明

B.1 大数の弱法則の証明

この節では定理 3.4 の証明を行う．互いに独立で同じ分布に従い，有限の平均 μ_X をもつ確率変数数列 $\{X_n; n = 1, 2, \dots\}$ と任意の正数 $\epsilon > 0$ が与えられたとき，ある正数 $b > 0$ を用いて新しい確率変数数列 $\{Y_n; n = 1, 2, \dots\}$ を

以下で定義する .

$$Y_n = \begin{cases} X_n, & |X_n - \mu_X| \leq b \\ \mu + b, & X_n - \mu_X > b \\ \mu - b, & X_n - \mu_X < -b \end{cases}$$

確率変数列 $\{Y_n; n = 1, 2, \dots\}$ は互いに独立で同一の分布に従っており , その平均を μ_Y とする . Y_1 の分散 σ_Y^2 は

$$E((Y_1 - \mu_X)^2) = E(\{(Y_1 - \mu_Y) + (\mu_Y - \mu_X)\}^2) = \sigma_Y^2 + (\mu_Y - \mu_X)^2$$

であるので , $\sigma_Y^2 \leq E((Y_1 - \mu_X)^2)$ である . さらに

$$\sigma_Y^2 \leq E((Y_1 - \mu_X)^2) = \int_{-\infty}^{\infty} (x - \mu_X)^2 dF_Y(x) \leq b \int_{-\infty}^{\infty} |x - \mu_X| dF_Y(x)$$

を得る . ただし $F_Y(x)$ は Y_1 の分布関数である . 最後の不等号は $|Y - \mu_X| \leq b$ を用いた . さらに $\mu_X - b \leq x < \mu_X + b$ では $F_Y(x)$ は X_1 の分布関数に等しいことに注意すると ,

$$\sigma_Y^2 \leq b \int_{-\infty}^{\infty} |x - \mu_X| dF_Y(x) \leq b \int_{-\infty}^{\infty} |x - \mu_X| dF_X(x)$$

である .

さて , $T_n = Y_1 + \dots + Y_n$ とし , 式 (18) のチェビシエフの不等式を ϵ を $\epsilon/2$ に変更して適用すると

$$P\left(\left|\frac{T_n}{n} - \mu_Y\right| \geq \frac{\epsilon}{2}\right) \leq \frac{4\sigma_Y^2}{n\epsilon^2} \leq \frac{4b\alpha}{n\epsilon^2}$$

となる . ただし $\alpha = \int_{-\infty}^{\infty} |x - \mu_X| dF_X(x) < \infty$ とした . $|\mu_Y - \mu_X|$ は $b \rightarrow \infty$ としたとき 0 へ近付くので , 任意の $\epsilon > 0$ に対して , 十分大きな b が存在し $|\mu_Y - \mu_X| < \epsilon/2$ とすることができる . よって , そのような b に対して

$$P\left(\left|\frac{T_n}{n} - \mu_X\right| \geq \epsilon\right) \leq \frac{4b\alpha}{n\epsilon^2} \quad (26)$$

を得る¹⁶ .

S_n と T_n は全ての k ($k = 1, \dots, n$) で $|X_k - \mu_X| \leq b$ となるような標本に対しては同じ値を取る . よって A_k で事象 $|X_k - \mu_X| > b$ を表すと

$$P(T_n \neq S_n) \leq P(\cup_{k=1}^n A_k) = nP(A_1) = nP(|X_1 - \mu_X| > b)$$

を得る . さらに $G(x) = P(|X_1 - \mu_X| \leq x)$ とすると

$$P(|X_1 - \mu_X| > b) = \int_b^{\infty} dG(x) \leq \frac{1}{b} \int_b^{\infty} xdG(x)$$

である . α が有限であるので最後の積分は $b \rightarrow \infty$ の極限で 0 へ収束する . よって , 任意の $\delta > 0$ に対して十分大きな b が存在し $P(|X_1 - \mu_X| > b) \leq \delta^2/b$ とすることが出来る . すなわち , 十分大きな b に対して

$$P(T_n \neq S_n) \leq nP(|X_1 - \mu_X| > b) \leq \frac{n\delta^2}{b} \quad (27)$$

事象 $|S_n/n - \mu_X| \geq \epsilon$ は事象 $|T_n/n - \mu_X| \geq \epsilon$ あるいは $S_n \neq T_n$ のいずれかが起こったときのみ起こり得るので , 式 (26) と式 (27) を組合せて , $\delta = b/n$ とすると

$$P\left(\left|\frac{S_n}{n} - \mu_X\right| \geq \epsilon\right) \leq \frac{4\delta\alpha}{\epsilon^2} + \delta \quad (28)$$

を得る . 式 (26) と式 (27) は任意の n と十分大きな b について成立するので , 式 (28) は任意の正数 $\delta > 0$ と十分大きな n について成立する . よって , 任意の正数 $\epsilon > 0$ に対して , 適当な δ を選ぶことで式 (28) の右辺第 1 項 , 第 2 項共に , いくらでも小さくすることが出来る . よって定理が成立する .

¹⁶ $|\mu_Y - \mu_X| < \epsilon/2$ のとき , $\{|T_n/n - \mu_X| \geq \epsilon\} = \{|T_n/n - \mu_Y + \mu_X - \mu_Y| \geq \epsilon\} \Rightarrow \{|T_n/n - \mu_Y| + |\mu_X - \mu_Y| \geq \epsilon\} \Rightarrow \{|T_n/n - \mu_Y| \geq \epsilon/2\}$ であるので , $A \Rightarrow B$ ならば $P(A) \leq P(B)$ に注意すると式 (26) を得る .

C 練習問題の略解

問 1.1.

$$F(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x < 1, \\ 1, & x \geq 1 \end{cases}, \quad f(x) = \begin{cases} 0, & x < 0 \text{ (あるいは } x \leq 0), \\ 1, & 0 \leq x < 1 \text{ (あるいは } 0 < x < 0, 0 \leq x < 0), \\ 0, & x \geq 1 \text{ (あるいは } x > 1), \end{cases}$$

(密度関数に関しては, 区間の境界をいずれに含めるかは任意, 全区間で積分して 1 になればよい)

問 1.2.

$$F(x) = \begin{cases} 0, & x < x_1 \\ 1/3, & x_1 \leq x < x_2, \\ 2/3, & x_2 \leq x < x_3, \\ 1, & x \geq x_3 \end{cases},$$

($x = x_1, x_2, x_3$ での取り扱いに注意)

問 1.3. (a) $A \cap B, A^c \cap B, A \cap B^c, A^c \cap B^c$ (b) $P(A^c \cap B) = b - c, P(A \cap B^c) = a - c, P(A^c \cap B^c) = 1 - a - b + c$
(c) $c \geq 0, a - c \geq 0, b - c \geq 0, 1 - a - b + c \geq 0$

問 1.4. (a) $3/4 \leq P(A \cup B) \leq 1$ (b) $1/12, 1/3$

問 1.5. $P(1 \text{ 回目} \text{ が赤}) = r/(r+b), P(2 \text{ 回目} \text{ が赤}) = r/(r+b), P(1 \text{ 回目} \text{ が赤} \mid 2 \text{ 回目} \text{ が赤}) = (r+c)/(r+b+c)$

問 1.6. A, B, C をランダムに選ばれた製品が生産された機械が A, B, C である事象とし, F を不良品である事象とすると

$$P(F) = P(A)P(F \mid A) + P(B)P(F \mid B) + P(C)P(F \mid C) = \frac{4}{10} \cdot \frac{1}{100} + \frac{3}{10} \cdot \frac{2}{100} + \frac{3}{10} \cdot \frac{3}{100} = \frac{19}{1000}$$

である. 一方, ランダムに一つを選んだとき, その製品が機械 C で生産され, かつ, 不良品である確率は

$$P(C \cap F) = P(C)P(F \mid C) = \frac{3}{10} \cdot \frac{3}{100} = \frac{9}{1000}$$

である. よって求める答えは $P(C \mid F) = P(C \cap F)/P(F) = 9/19$.

問 1.7. 検査を受けた人がウイルスに感染しているという事象を D , 検査の結果が陽性となる事象を A とすると

$$P(A) = P(D)P(A \mid D) + P(D^c)P(A \mid D^c) = 0.001 \times 0.9 + 0.999 \times 0.03 = 0.03087$$

となり,

$$P(D \mid A) = 0.001 \times 0.9 / 0.03087 \approx 0.0292$$

である. よって陽性と判断されても実際に感染しているとすぐに判断するわけにはいかない.

問 1.8. A, B, C をそれぞれ A, B, C が釈放されるという事象とする. また, 看守が B, C の名を告げる事象を b, c とする. このとき $p_b = P(b \mid A), p_c = P(c \mid A)$ を用いて

$$P(A \cap b) = P(A)P(b \mid A) = \frac{1}{3} \times p_b, \quad P(A \cap c) = P(A)P(c \mid A) = \frac{1}{3} \times p_c$$

と考えることができ, また

$$P(B \cap b) = P(B)P(b \mid B) = 1/3 \times 0, \quad P(B \cap c) = P(B)P(c \mid B) = 1/3 \times 1,$$

$$P(C \cap b) = P(C)P(b \mid C) = 1/3 \times 1 \quad P(C \cap c) = P(C)P(c \mid C) = 1/3 \times 0$$

である. よって $P(b) = P(A \cap b) + P(B \cap b) + P(C \cap b) = (1 + p_b)/3$ なので

$$P(A \mid b) = \frac{P(A \cap b)}{P(b)} = \frac{p_b}{1 + p_b}$$

を得る．もし， A が釈放される（ $= B, C$ の死刑が執行される）とき，看守が B, C の名をランダムに告げるならば $p_b = 1/2$ であり，この場合 $P(A | b) = 1/3$ となるので，看守の考えは正しい．一方，ランダムではなく，偏りがある（たとえば口にしやすい方の名前を言う等）場合， $P(A | b) \neq 1/3$ となり，看守は何らかの情報を A に与えたことになる．なお， $P(A | b)$ は p_b の増加関数なので， B の方が告げやすい名前であった（ $p_b > 1/2$ ）場合， A が釈放される確率は増えたことになり，特に， $p_b = 1$ ならば $P(A | b) = 1/2$ となる．

問 1.9. 解答者が最初に腕時計が背後にある衝立を選択する事象を S とし，腕時計を獲得できるという事象を W とする．選択し直さない場合，腕時計を獲得できる確率は

$$P(W) = P(S)P(W | S) + P(S^c)P(W | S^c) = \frac{1}{3} \times 1 + \frac{2}{3} \times 0 = \frac{1}{3}$$

となる．一方，選択し直す場合，腕時計を獲得できる確率は

$$P(W) = P(S)P(W | S) + P(S^c)P(W | S^c) = \frac{1}{3} \times 0 + \frac{2}{3} \times 1 = \frac{2}{3}$$

となるので，選択し直すことが合理的．

問 1.10. $P(A_1) = P(A_2) = P(A_3) = 1/2$, $P(A_1 \cap A_2) = 1/4$, $P(A_2 \cap A_3) = 1/4$, $P(A_1 \cap A_3) = 1/4$ より A_i と A_j ($i \neq j$) は互いに独立． $P(A_1 \cap A_2 \cap A_3) = 0 \neq 1/8 = P(A_1)P(A_2)P(A_3)$ より A_1, A_2, A_3 は互いに独立ではない．

問 1.11. 各対戦の勝者を並べたものを標本空間 Ω とする．すなわち，例えば $ACBACBACC$ を $A(CBA)^2CC$ と書くとする． $\Omega = \{A(CBA)^n A, A(CBA)^n CC, A(CBA)^n CBB, B(CAB)^n B, B(CAB)^n CC, B(CAB)^n CAA, n = 0, 1, \dots\}$ である． A が優勝する場合は $A(CBA)^n A$ と $B(CAB)^n CAA$ なので，その確率は $\sum_{n=0}^{\infty} (1/2)(p^3)^n(1-p) + \sum_{n=0}^{\infty} (1/2)(p^3)^n p p(1-p) = (1+p^2)/[2(1+p+p^2)]$ ， B が優勝する確率は A と同じ． C が優勝する確率は $p/(1+p+p^2)$ ． C が常に不利となる．

問 1.12. $n = 0$ のとき $P(X > 0) = P(X > m | X > m) = 1$ より

$$P(X = 0) = 1 - P(X > 0) = 0$$

さらに，

$$P(X > m+n | X > m) = \frac{P(X > m+n, X > m)}{P(X > m)}$$

なので， $P(X > m+n | X > m) = P(X > n)$ ならば

$$P(X > m+n) = P(X > m)P(X > n)$$

である． $p = P(X > 1)$ とおくと， $m = n = 1$ のとき

$$P(X > 2) = p^2$$

さらに， $n = 1$ とおくと

$$P(X > m+1) = P(X > m) \cdot p$$

よって帰納法より $P(X > m) = p^m$ を得る．なお，これは $m = 0, 1$ でも成立．以上より

$$P(X = k) = P(X > k-1) - P(X > k) = (1-p)p^{k-1}, \quad (k = 1, 2, \dots)$$

問 1.13. $m = n = 1$ とすると

$$P(X \geq 2) = P(X \geq 1)P(X \geq 1)$$

となる．すなわち， $p = P(X \geq 1)$ ($0 \leq p \leq 1$) とおくと $P(X \geq 2) = p^2$ ．そこで，ある n ($n = 1, 2, 3, \dots$) に対して $P(X \geq n) = p^n$ と仮定すると

$$P(X \geq n+1) = P(X \geq n)P(X \geq 1) = p^n \cdot p = p^{n+1}$$

を得る . また , $P(X \geq 0) = 1 = p^0$ なので , 全ての n ($n = 0, 1, \dots$) に対して $P(X \geq n) = p^n$ が成立する . よって , $k = 1, 2, \dots$ に対して ,

$$P(X = k) = P(X \geq k) - P(X \geq k + 1) = p^k - p^{k+1} = (1 - p)p^k$$

となり , $P(X = 0) = 1 - \sum_{k=1}^{\infty} P(X = k) = 1 - p$ である . すなわち , ある p ($0 \leq p \leq 1$) を用いて , $P(X = k) = (1 - p)p^k$ ($k = 0, 1, \dots$) と書くことができる .

問 1.14.

$$P(X \leq x, Y \leq y) + P(X \leq x, Y > y) + P(X > x, Y \leq y) + P(X > x, Y > y) = 1$$

$$F_X(x) = P(X \leq x, Y \leq y) + P(X \leq x, Y > y)$$

$$F_Y(y) = P(X \leq x, Y \leq y) + P(X > x, Y \leq y)$$

の両辺をそれぞれ足し合わせて整理すれば題意が示される .

問 1.15. たとえば ,

$$P(X = -1, Y = 0) = P(X = 0, Y = 0) = P(X = 0, Y = 1) = P(X = 1, Y = 1) = 1/4$$

$$P(X = -1, Y = 1) = P(X = 1, Y = 0) = 0$$

このとき ,

$$P(X = -1) = P(X = -1, Y = 0) + P(X = -1, Y = 1) = 1/4,$$

$$P(Y = 0) = P(X = -1, Y = 0) + P(X = 0, Y = 0) + P(X = 1, Y = 0) = 1/2$$

なので

$$P(X = -1, Y = 0) \neq P(X = -1)P(Y = 0)$$

となり X と Y は独立ではない . しかし ,

$$P(X^2 = 0, Y^2 = 0) = P(X^2 = 0, Y^2 = 1) = P(X^2 = 1, Y^2 = 0) = P(X^2 = 1, Y^2 = 1) = 1/4$$

であり ,

$$P(X^2 = 0) = P(X^2 = 1) = 1/2, \quad P(Y^2 = 0) = P(Y^2 = 1) = 1/2$$

なので $i \in \{-1, 0, 1\}$, $j \in \{0, 1\}$ に対して $P(X^2 = i, Y^2 = j) = P(X^2 = i)P(Y^2 = j)$ が成立し , X^2 と Y^2 は独立である .

問 1.16. $p_{i,j} = P(X = i, Y = j)$ ($i \geq j$) とする . $1 \leq j < i \leq 6$ (全 15 通り) に対して $p_{i,j} = 1/18$, $1 \leq j = i \leq 6$ に対して $p_{i,i} = 1/36$. $P(X = i) = (i-1)/18 + 1/36 = (2i-1)/36$ ($i = 1, 2, \dots, 6$) , $P(Y = j) = (6-j)/18 + 1/36 = (13-2j)/36$.

問 1.17. (a)

$$\begin{aligned} P(Y_k = k, Y_n = k) &= P(Y_1 = 1)P(Y_2 = 2 | Y_1 = 1) \cdots P(Y_k = k | Y_{k-1} = k-1) \\ &\quad \times P(Y_{k+1} = k | Y_k = k)P(Y_{k+2} = k | Y_{k+1} = k) \cdots P(Y_n = k | Y_{n-1} = k) \\ &= \frac{r}{r+b} \cdot \frac{r+c}{r+b+c} \cdots \frac{r+(k-1)c}{r+b+(k-1)c} \\ &\quad \times \frac{b}{r+b+kc} \cdot \frac{b+c}{r+b+(k+1)c} \cdots \frac{b+(n-k-1)c}{r+b+(n-1)c} \end{aligned}$$

(b) i_j ($j = 1, 2, \dots, k$) 回目の試行で赤色の玉が取り出され , l_j ($j = 1, 2, \dots, n-k$) 回目の試行で黒色の玉が取り出される確率を考える . ただし $(i_1, i_2, \dots, i_k, l_1, l_2, \dots, l_{n-k})$ は $(1, 2, \dots, n)$ の置換であり , 前者を昇順に並べ替

えると後者となる．上記が起こる確率は

$$\frac{r}{r+b+(i_1-1)c} \cdot \frac{r+c}{r+b+(i_2-1)c} \cdots \frac{r+(k-1)c}{r+b+(i_k-1)c} \\ \times \frac{b}{r+b+(l_1-1)c} \cdot \frac{b+c}{r+b+(l_2-1)c} \cdots \frac{b+(n-k-1)c}{r+b+(l_{n-k}-1)c}$$

であり, $(i_1, i_2, \dots, i_k, l_1, l_2, \dots, l_{n-k})$ は $(1, 2, \dots, n)$ の置換であることから, 上式の分母に現れる項の積

$$[r+b+(i_1-1)c][r+b+(i_2-1)c] \cdots [r+b+(i_k-1)c] \times [r+b+(l_1-1)c][r+b+(l_2-1)c] \cdots [r+b+(l_{n-k}-1)c]$$

は (a) の答えの分母に現れる項の積 $[r+b][r+b+c] \cdots [r+b+(n-1)c]$ に等しいので, この確率は (a) で求めたものに等しい．すなわち, $Y_n = k$ となるような取り出し方は $n!/ [k!(n-k)!]$ 通りあり, それらは同様に確からしい．以上より

$$P(Y_n = k) = \binom{n}{k} P(Y_k = k, Y_n = k)$$

(c) (b) における考察より

$$P(X_1 = i_1, X_2 = i_2, \dots, X_{n-1} = i_{n-1}, X_n = 1) \\ = P(X_1 = 1, X_2 = i_1, X_3 = i_2, \dots, X_{n-1} = i_{n-2}, X_n = i_{n-1})$$

である．よって $n = 2, 3, \dots$ に対して

$$P(X_n = 1) = \sum_{i_1=0}^1 \sum_{i_2=0}^1 \cdots \sum_{i_{n-1}=0}^1 P(X_1 = i_1, X_2 = i_2, \dots, X_{n-1} = i_{n-1}, X_n = 1) \\ = \sum_{i_1=0}^1 \sum_{i_2=0}^1 \cdots \sum_{i_{n-1}=0}^1 P(X_1 = 1, X_2 = i_1, X_3 = i_2, \dots, X_{n-1} = i_{n-2}, X_n = i_{n-1}) \\ = P(X_1 = 1) = \frac{r}{r+b}$$

(d)

$$P(X_m = 1, X_n = 0) \\ = \sum_{i_1=0}^1 \cdots \sum_{i_{m-1}=0}^1 \sum_{i_{m+1}=0}^1 \cdots \sum_{i_{n-1}=0}^1 P(X_m = 1, X_n = 0, X_j = i_j \ (j = 1, 2, \dots, n-1, j \neq m)) \\ = \sum_{i_1=0}^1 \cdots \sum_{i_{m-1}=0}^1 \sum_{i_{m+1}=0}^1 \cdots \sum_{i_{n-1}=0}^1 P(X_1 = 1, X_2 = 0, X_m = i_1, X_n = i_2, X_j = i_j \ (j = 3, 4, \dots, n-1, j \neq m)) \\ = P(X_1 = 1, X_2 = 0) = P(X_1 = 1)P(X_2 = 0 | X_1 = 1) = \frac{r}{r+b} \cdot \frac{b}{r+b+c}$$

問 1.18. (a) $S_{k,N} = 1$ は X_k, X_{k+1}, \dots, X_N の内, 一つが 1 で残り全てが 0 である事象に等しい． $X_i = 1$ となる i で場合分けし, X_k が互いに独立であることに注意すると

$$P(S_{k,N} = 1) = \sum_{i=k}^N P(X_i = 1)P(S_{k,N} = 1 | X_i = 1) \\ = \sum_{i=k}^N p_i P(X_k = 0, \dots, X_{i-1} = 0, X_{i+1} = 0, \dots, X_N = 0) = \sum_{i=k}^N p_i \prod_{\substack{j=k \\ j \neq i}}^N P(X_j = 0) \\ = \sum_{i=k}^N r_i q_i \prod_{\substack{j=k \\ j \neq i}}^N q_j = \left(\prod_{j=k}^N q_j \right) \cdot \sum_{i=k}^N r_i$$

(b) 任意の $k \in \{1, 2, \dots, N-1\}$ に対して

$$\begin{aligned} P(S_{k+1,N} = 1) - P(S_{k,N} = 1) &= \left(\prod_{j=k+1}^N q_j \right) \left(\sum_{i=k+1}^N r_i - q_k \sum_{i=k}^N r_i \right) = \left(\prod_{j=k+1}^N q_j \right) \left((1 - q_k) \sum_{i=k+1}^N r_i - q_k r_k \right) \\ &= \left(\prod_{j=k+1}^N q_j \right) \left(p_k \sum_{i=k+1}^N r_i - p_k \right) = p_k \left(\prod_{j=k+1}^N q_j \right) \left(\sum_{i=k+1}^N r_i - 1 \right) \end{aligned}$$

(c) もし $K_N = 1$ ならば $P(S_{2,N} = 1) - P(S_{1,N} = 1) < 0$ である．すなわち， $q_j > 0$ ， $p_k > 0$ なので，前問の結果より

$$\sum_{i=2}^N r_i < 1 \quad (29)$$

が必要である．逆に，式 (29) を仮定すると $r_i = p_i/q_i > 0$ ($i = 1, 2, \dots, N$) なので，任意の $k \in \{3, 4, \dots, N-1\}$ に対して

$$\sum_{i=k}^N r_i < 1, \quad k = 3, 4, \dots, N$$

が成立し，式 (29) より，

$$P(S_{1,N} = 1) > P(S_{2,N} = 1) > P(S_{3,N} = 1) > \dots > P(S_{N,N} = 1)$$

を得る．以上より，式 (29) が $K_N = 1$ であるための必要十分条件である．

(d) 与条件より， $P(S_{k+1,N} = 1) = P(S_{k,N} = 1)$ となるような k は存在しない．よって， $K_N = k$ ($k = 2, 3, \dots, N$) ならば

$$P(S_{k,N} = 1) - P(S_{k-1,N} = 1) > 0, \quad P(S_{k+1,N} = 1) - P(S_{k,N} = 1) < 0$$

が成立し，これは

$$\sum_{i=k}^N r_i > 1, \quad \sum_{i=k+1}^N r_i < 1 \quad (30)$$

と等価である．逆に式 (30) が成立すれば， $r_i > 0$ なので

$$\sum_{i=j}^N r_j > 1 \quad (j = 1, 2, \dots, k), \quad \sum_{i=j}^N r_j < 1 \quad (j = k+1, k+2, \dots, N)$$

が成立する．すなわち

$$P(S_{1,N} = 1) < P(S_{2,N} = 1) < \dots < P(S_{k-1,N} = 1) < P(S_{k,N} = 1) > P(S_{k+1,N} = 1) > \dots > P(S_{N,N} = 1)$$

となり， $P(S_{k,N} = 1)$ が最大である．よって $K_N = k \in \{2, 3, \dots, N\}$ となるための必要十分条件は式 (30) で与えられる．

(e) 題意より $q_i = i/(i+1)$ であるので $r_i = p_i/q_i = 1/i$ ($i = 2, 3, \dots$) である．よって，前問の結果より $K_N \geq 2$ は次式を満たす．

$$\sum_{i=K_N}^N \frac{1}{i} > 1, \quad \sum_{i=K_N+1}^N \frac{1}{i} < 1,$$

一方 ($k \geq 2$ に対して)

$$\log \left(\frac{N+1}{k} \right) = \int_k^{N+1} \frac{dx}{x} < \sum_{i=k}^N \frac{1}{i} < \int_{k-1}^N \frac{dx}{x} = \log \left(\frac{N}{k-1} \right)$$

であるので (ζ 十分に大きな N に対して)

$$\begin{aligned} 1 &< \sum_{i=K_N}^N \frac{1}{i} < \log \left(\frac{N}{K_N-1} \right) \\ \log \left(\frac{N+1}{K_N+1} \right) &< \sum_{i=K_N+1}^N \frac{1}{i} < 1 \end{aligned}$$

が成立する．さらに

$$1 < \log\left(\frac{N}{K_N - 1}\right) \Leftrightarrow K_N < Ne^{-1} + 1$$

$$\log\left(\frac{N+1}{K_N + 1}\right) < 1 \Leftrightarrow K_N > (N+1)e^{-1} - 1$$

なので

$$(N+1)e^{-1} - 1 < K_N < Ne^{-1} + 1$$

すなわち

$$\frac{(N+1)e^{-1} - 1}{N} < \frac{K_N}{N} < \frac{Ne^{-1} + 1}{N}$$

を得る．ここで

$$\lim_{N \rightarrow \infty} \frac{(N+1)e^{-1} - 1}{N} = e^{-1}, \quad \lim_{N \rightarrow \infty} \frac{Ne^{-1} + 1}{N} = e^{-1}$$

となるため，

$$\lim_{N \rightarrow \infty} \frac{K_N}{N} = e^{-1}$$

問 1.20. $x \geq 0$ に対して $G(x) = P(X^2 \leq x) = P(-\sqrt{x} \leq X \leq \sqrt{x}) = P(X \leq \sqrt{x}) - P(X < -\sqrt{x}) = F(\sqrt{x}) - \lim_{u \rightarrow x-} F(-\sqrt{u})$.

問 1.21. 実根を持つための必要十分条件は $B^2 \geq AC$. X ($X = A, B, C$) の分布関数を $F_X(x) = P(X \leq x)$, 密度関数を $f_X(x)$ とすると，一般には

$$P(B^2 \geq AC) = P(B \geq \sqrt{AC}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(B \geq \sqrt{AC} \mid A = x, C = y) f_A(x) f_C(y) dx dy$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(B \geq \sqrt{xy}) f_A(x) f_C(y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1 - F_B(\sqrt{xy})) f_A(x) f_C(y) dx dy$$

を計算すればよい．題意より

$$F_X(x) = \begin{cases} 0, & x \leq 0 \\ x, & 0 \leq x \leq 1 \\ 1, & x \geq 1 \end{cases} \quad f_X(x) = \begin{cases} 0, & x \leq 0 \\ 1, & 0 < x \leq 1 \\ 0, & x > 1 \end{cases}$$

なので (注意: $x = 0, 1$ における密度関数の値は $0, 1$ のいずれでも良い) , $0 \leq x, y \leq 1$ のとき , $0 \leq \sqrt{xy} \leq 1$ に注意すると

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1 - F_B(\sqrt{xy})) f_A(x) f_C(y) dx dy = \int_0^1 \int_0^1 (1 - \sqrt{xy}) dx dy = \int_0^1 \int_0^1 (1 - \sqrt{xz}) dx \cdot 2z dz = \dots = 5/9$$

問 1.22. $k = 1$ のとき , 定義より $P(F_1 \leq x) = P(X_1 \leq x) = 1 - \exp(-\mu x)$ なので , 式 (11) は $k = 1$ に対して成立している . そこで , ある $k = n$ ($n = 1, 2, \dots$) に対して式 (11) が成立すると仮定する . このとき

$$P(F_{n+1} \leq x) = \int_0^x P(F_n \leq y) \mu e^{-\mu(x-y)} dy = \int_0^x \left(1 - \sum_{i=0}^{n-1} e^{-\mu y} \frac{(\mu y)^i}{i!}\right) \mu e^{-\mu(x-y)} dy$$

$$= \int_0^x \mu e^{-\mu(x-y)} dy - \sum_{i=0}^{n-1} \mu e^{-\mu x} \int_0^x \frac{(\mu y)^i}{i!} dy = 1 - e^{-\mu x} - \sum_{i=0}^{n-1} \mu e^{-\mu x} \left[\frac{\mu^i y^{i+1}}{(i+1)!}\right]_0^x$$

$$= 1 - e^{-\mu x} - \sum_{i=0}^{n-1} \mu e^{-\mu x} \frac{\mu^i x^{i+1}}{(i+1)!} = 1 - \sum_{i=0}^n e^{-\mu x} \frac{(\mu x)^i}{i!}$$

となり , 式 (11) は $k = n + 1$ に対しても成立する . よって全ての k ($k = 1, 2, \dots$) に対して式 (11) は成立する .

問 1.23. X の密度関数は $x \in (0, 1]$ のとき 1 , それ以外は 0 となる関数 . Z の密度関数を $f(x)$ とすると $x \leq 0$ ならびに $x > 2$ では $f(x) = 0$, $0 < x \leq 1$ では $f(x) = \int_0^x 1 \cdot 1 dx = x$, $1 < x \leq 2$ では $\int_{x-1}^1 1 \cdot 1 dx = -x + 2$,

問 1.24. (a) $z < 0$ のとき

$$\begin{aligned} P(Z_1 \leq z) &= \int_{x=0}^{\infty} \mu e^{-\mu x} \int_{y=x-z}^{\infty} \mu e^{-\mu y} dy dx = \int_0^{\infty} \mu e^{-\mu x} \cdot e^{-\mu(x-z)} dx \\ &= \frac{1}{2} e^{\mu z} \int_0^{\infty} 2\mu e^{-2\mu x} dx = \frac{1}{2} e^{\mu z} \end{aligned}$$

なので $z < 0$ の場合の密度関数は

$$\frac{d}{dz} P(Z_1 \leq z) = \frac{\mu}{2} e^{\mu z}$$

一方, $z \geq 0$ のとき

$$\begin{aligned} P(Z_1 \leq z) &= \int_{y=0}^{\infty} \mu e^{-\mu y} \int_{x=0}^{y+z} \mu e^{-\mu x} dx dy = \int_{y=0}^{\infty} \mu e^{-\mu y} (1 - e^{-\mu(y+z)}) dy \\ &= 1 - e^{-\mu z} \int_{y=0}^{\infty} \mu e^{-2\mu y} dy = 1 - \frac{1}{2} e^{-\mu z} \end{aligned}$$

なので $z \geq 0$ の場合の密度関数は

$$\frac{d}{dz} P(Z_1 \leq z) = \frac{\mu}{2} e^{-\mu z}$$

これら合わせて

$$\frac{\mu}{2} e^{-\mu|z|} \quad (-\infty < z < \infty)$$

(b)

$$\begin{aligned} P(Z_2 \leq z) &= P(X + Y \leq \alpha z) = \int_{x=0}^{\alpha z} \mu e^{-\mu x} \int_{y=0}^{\alpha z - x} \mu e^{-\mu y} dy dx \\ &= \int_{x=0}^{\alpha z} \mu e^{-\mu x} (1 - e^{-\mu(\alpha z - x)}) dx = 1 - e^{-\alpha \mu z} - \alpha \mu z e^{-\alpha \mu z} \end{aligned}$$

なので, 密度関数は

$$\frac{d}{dz} P(Z_2 \leq z) = \alpha \mu e^{-\alpha \mu z} - \alpha \mu e^{-\alpha \mu z} + \alpha^2 \mu^2 z e^{-\alpha \mu z} = \alpha^2 \mu^2 z e^{-\alpha \mu z} \quad (z \geq 0)$$

(c) 題意より

$$P(Z_3 \leq z) = 0 \quad (z < 1)$$

である. 以下では $z \geq 1$ に対する $P(Z_3 \leq z)$ を考える. X と Y の結合密度関数は $1/\gamma^2$ ($0 \leq x, y \leq \gamma$) であり $P(X \leq Y) = P(X > Y) = 1/2$ であるので, U と V の結合密度関数は $2/\gamma^2$ ($0 \leq u \leq v \leq \gamma$) で与えられる. また, $Z_3 \leq z \Leftrightarrow V/z < U$ かつ $0 \leq U \leq V \leq \gamma$ なので

$$Z_3 \leq z \Leftrightarrow 0 \leq V/z \leq U \leq V \leq \gamma$$

である. 以上より

$$\begin{aligned} P(Z_3 \leq z) &= \int_{v=0}^{\gamma} \int_{u=v/z}^v \frac{2}{\gamma^2} du dv = \frac{2}{\gamma^2} \int_{v=0}^{\gamma} \left(v - \frac{v}{z}\right) dv = \frac{2}{\gamma^2} \cdot \frac{z-1}{z} \int_{v=0}^{\gamma} v dv = \frac{2}{\gamma^2} \cdot \frac{z-1}{z} \cdot \frac{\gamma^2}{2} \\ &= \frac{z-1}{z} \quad (z \geq 1) \end{aligned}$$

問 1.25.

$$\begin{aligned} P(Y \leq x) &= P((X_1 + X_2)/2 \leq x) = P(X_1 + X_2 \leq 2x) = \int \int_{u_1+u_2 \leq 2x} f_1(u_1) f_2(u_2) du_1 du_2 \\ &= \int_{u_2=0}^{2x} \int_{u_1=0}^{2x-u_2} f_1(u_1) f_2(u_2) du_1 du_2 = \int_0^{2x} F_1(2x-u_2) f_2(u_2) du_2 \end{aligned}$$

密度関数は分布関数 $G(x) = P(Y \leq x)$ の一階微分で与えられるが, 積分の区間にも x があるため注意が必要. 一般に, このような関数の微分は合成関数の微分としてとらえることができる. 偏微分可能な 2 変数関数 $H(u, v)$ と $u = h_1(x)$, $v = h_2(x)$ で構成される合成関数 $G(x) = H(h_1(x), h_2(x))$ に対して, 合成関数の微分公式より

$$\frac{d}{dx} G(x) = \frac{\partial H}{\partial u} \frac{du}{dx} + \frac{\partial H}{\partial v} \frac{dv}{dx}$$

である．これを応用して密度関数を求める．まず，対応関係は

$$H(u, v) = \int_0^u F_1(v - y) f_2(y) dy, \quad u = h_1(x) = 2x, \quad v = h_2(x) = 2x$$

このとき

$$\frac{\partial H}{\partial u} = F_1(v - u) f_2(u), \quad \frac{du}{dx} = 2, \quad \frac{\partial H}{\partial v} = \int_0^u \frac{\partial}{\partial v} F_1(v - y) f_2(y) dy = \int_0^u f_1(v - y) f_2(y) dy, \quad \frac{dv}{dx} = 2$$

これらを微分公式へ代入し， $u = v = x$ とおき， $F_1(0) = 0$ に注意すると

$$\frac{d}{dx} G(x) = F_1(0) f_2(x) \cdot 2 + \int_0^{2x} f_1(2x - y) f_2(y) dy \cdot 2 = 2 \int_0^{2x} f_1(2x - y) f_2(y) dy$$

問 1.26. $p_i(k) = \exp(-\lambda_i) \lambda_i^k / k!$ としたとき， $q(k) = \sum_{n=0}^k p_1(n) p_2(k - n)$ である．

$$\begin{aligned} q(k) &= \sum_{n=0}^k e^{-\lambda_1} \frac{\lambda_1^n}{n!} e^{-\lambda_2} \frac{\lambda_2^{k-n}}{(k-n)!} = e^{-(\lambda_1 + \lambda_2)} \frac{(\lambda_1 + \lambda_2)^k}{k!} \sum_{n=0}^k \frac{k!}{n!(k-n)!} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} \right)^n \left(\frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^{k-n} \\ &= e^{-(\lambda_1 + \lambda_2)} \frac{(\lambda_1 + \lambda_2)^k}{k!} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2} + \frac{\lambda_2}{\lambda_1 + \lambda_2} \right)^k = e^{-(\lambda_1 + \lambda_2)} \frac{(\lambda_1 + \lambda_2)^k}{k!} \end{aligned}$$

問 1.27. $k = 1$ のとき，定義より $P(F_1 = m) = P(X_1 = m) = (1 - p)p^{m-1}$ なので，式 (13) は $k = 1$ に対して成立している．そこで，ある $k = n$ ($n = 1, 2, \dots$) に対して式 (13) が成立すると仮定する．このとき $m = n + 1, n + 2, \dots$ に対して

$$\begin{aligned} P(F_{n+1} = m) &= \sum_{i=n}^{m-1} P(F_n = i) P(X_{n+1} = m - i) = \sum_{i=n}^{m-1} \frac{(i-1)!}{(n-1)!(i-n)!} (1-p)^n p^{i-n} \cdot (1-p)p^{m-i-1} \\ &= (1-p)^{n+1} p^{m-n-1} \sum_{i=n}^{m-1} \frac{(i-1)!}{(i-n)!(n-1)!} = (1-p)^{n+1} p^{m-n-1} \sum_{i=0}^{m-n-1} \frac{(i+n-1)!}{i!(n-1)!} \\ &= (1-p)^{n+1} p^{m-n-1} \frac{((m-n-1) + (n-1) + 1)!}{(m-n-1)!(n-1)!} = (1-p)^{n+1} p^{m-n-1} \frac{(m-1)!}{n!(m-n-1)!} \end{aligned}$$

となり，式 (13) は $k = n + 1$ に対しても成立する．よって全ての k ($k = 1, 2, \dots$) に対して式 (13) は成立する．

問 1.28. (a) $P(X_0 = \min(X_0, X_1, \dots, X_n)) = \int_{-\infty}^{\infty} [\prod_{i=1}^n (1 - F_i(x))] f_0(x) dx$

(b) $p(n) = P(X_0 = \min(X_0, X_1, \dots, X_n))$ とおく．

$$\begin{aligned} p(n) &= \int_{-\infty}^{\infty} (1 - F(x))^n f(x) dx = [(1 - F(x))^n F(x)]_{-\infty}^{\infty} + n \int_{-\infty}^{\infty} (1 - F(x))^{n-1} F(x) f(x) dx \\ &= n \int_{-\infty}^{\infty} (1 - F(x))^{n-1} f(x) dx - n \int_{-\infty}^{\infty} (1 - F(x))^n f(x) dx = np(n-1) - np(n) \end{aligned}$$

より $(n+1)p(n) = np(n-1)$ ．よって $(k+1)p(k) = kp(k-1)$ の両辺をそれぞれ全ての $k = 1, 2, \dots, n$ について足し合わせると $(n+1)p(n) = p(0)$ より， $p(n) = 1/(n+1)$ ．

問 1.29. X_i ($i = 1, 2, \dots, N$) は互いに独立なので， $F(x) = P(X_i \leq x)$ とすると，

$$\begin{aligned} P(\max(X_1, X_2, \dots, X_N) > x) &= 1 - P(\max(X_1, X_2, \dots, X_N) \leq x) = 1 - P(X_1 \leq x, X_2 \leq x, \dots, X_N \leq x) \\ &= 1 - P(X_1 \leq x) P(X_2 \leq x) \cdots P(X_N \leq x) = 1 - F^N(x) \end{aligned}$$

となるので

$$\frac{P(\max(X_1, X_2, \dots, X_N) > x)}{P(X_1 > x)} = 1 + F(x) + F^2(x) + \cdots + F^{N-1}(x)$$

である． $\lim_{x \rightarrow \infty} F(x) = 1$ なので題意が成立．

(別解) 題意の極限值を L_N とする． X_N が x より大きいかわりかで場合分けをし，

$$P(\max(X_1, X_2, \dots, X_N) > x \mid X_N > x) = 1$$

$$P(\max(X_1, X_2, \dots, X_N) > x \mid X_N \leq x) = P(\max(X_1, X_2, \dots, X_{N-1}) > x)$$

に注意すると

$$\begin{aligned} P(\max(X_1, X_2, \dots, X_N) > x) &= P(\max(X_1, X_2, \dots, X_N) > x \mid X_N \leq x)P(X_N \leq x) \\ &\quad + P(\max(X_1, X_2, \dots, X_N) > x \mid X_N > x)P(X_N > x) \\ &= P(\max(X_1, X_2, \dots, X_{N-1}) > x)P(X_N \leq x) + P(X_N > x) \end{aligned}$$

この両辺を $P(X_1 > x)$ で割って, $P(X_N > x) = P(X_1 > x)$ に注意すると

$$\frac{P(\max(X_1, X_2, \dots, X_N) > x)}{P(X_1 > x)} = P(X_N \leq x) \cdot \frac{P(\max(X_1, X_2, \dots, X_{N-1}) > x)}{P(X_1 > x)} + 1$$

さらに $x \rightarrow \infty$ の極限を考えると $N \geq 2$ に対して

$$L_N = 1 \cdot L_{N-1} + 1 = L_{N-1} + 1$$

を得る. $L_1 = 1$ は自明. よって $L_N = N$.

問 1.30. X と Y が独立であることに注意. 周辺分布関数 $F(x)$ は共に $F(x) = 1 - \exp(-x)$ なので, 周辺密度関数は $\frac{d}{dx}F(x) = \exp(-x)$. よって

$$\begin{aligned} P(Z_1 \leq z_1) &= P(X + Y \leq z_1) = \int_{y=0}^{z_1} \int_{x=0}^{z_1-y} e^{-x} e^{-y} dy = \int_0^{z_1} (1 - e^{-(z_1-y)}) e^{-y} dy \\ &= \int_0^{z_1} (e^{-y} - e^{-z_1}) dy = 1 - (1 + z_1)e^{-z_1} \\ P(Z_2 \leq z_2) &= P(X/Y \leq z_2) = P(X \leq z_2 Y) = \int_0^\infty P(X \leq z_2 y) e^{-y} dy = \int_0^\infty (1 - e^{-z_2 y}) \exp(-y) dy \\ &= z_2 / (z_2 + 1) \end{aligned}$$

であり

$$\begin{aligned} P(Z_1 \leq z_1, Z_2 \leq z_2) &= P(X + Y \leq z_1, X \leq z_2 Y) = \int_0^{z_1} P(X \leq z_1 - y, X \leq z_2 y) e^{-y} dy \\ &= \int_0^{z_1} P(X \leq \min(z_1 - y, z_2 y)) e^{-y} dy = \int_0^{z_1} (1 - \exp(-\min(z_1 - y, z_2 y))) e^{-y} dy \\ &= 1 - e^{-z_1} - \int_0^{z_1} \exp(-\min(z_1, (z_2 + 1)y)) dy \\ &= 1 - e^{-z_1} - \int_0^{z_1/(z_2+1)} e^{-(z_2+1)y} dy - \int_{z_1/(z_2+1)}^{z_1} e^{-y} dy \\ &= z_2 / (z_2 + 1) \cdot (1 - (1 + z_1)e^{-z_1}) \\ &= P(Z_2 \leq z_2) P(Z_1 \leq z_1) \end{aligned}$$

となり, 題意が示された.

問 1.31. $P(X > x + y \mid X > x) = P(X > x + y) / P(X > x)$ であるので $P(X > x + y) = P(X > x) P(X > y)$ が成立する. $\bar{F}(x) = P(X > x) = 1 - F(x)$ とすると, $\bar{F}(x)$ は微分可能であり, かつ

$$\bar{F}(x + y) = \bar{F}(x) \bar{F}(y)$$

である. 特に $x = y = 0$ とすると $\bar{F}(0) = \bar{F}^2(0)$ となり, $\bar{F}(0) = 0$ あるいは $\bar{F}(0) = 1$ のいずれかが成立する. もし $\bar{F}(0) = P(X > 0) = 0$ ならば, 題意より $P(X = 0) = 1$ である. すなわち, $F(x) = 1 (x \geq 0)$. 次に, $\bar{F}(0) = 1$ であると仮定する. $\bar{F}(x + y) = \bar{F}(x) \bar{F}(y)$ において $y = \Delta x$ とし, $\bar{F}(x)$ の導関数 $\bar{f}(x)$ を考えると

$$\begin{aligned} \bar{f}(x) &= \frac{d}{dx} \bar{F}(x) = \lim_{\Delta x \rightarrow 0} \frac{\bar{F}(x + \Delta x) - \bar{F}(x)}{\Delta x} = \bar{F}(x) \cdot \lim_{\Delta x \rightarrow 0} \frac{\bar{F}(\Delta x) - 1}{\Delta x} = \bar{F}(x) \cdot \lim_{\Delta x \rightarrow 0} \frac{\bar{F}(\Delta x) - \bar{F}(0)}{\Delta x} \\ &= \bar{F}(x) \cdot \bar{f}(0) \end{aligned}$$

となる．よって

$$\bar{F}(x) = F(0)e^{\bar{f}(0)x} = e^{\bar{f}(0)x}$$

となり， $F(x) = 1 - e^{\bar{f}(0)x}$ ($x \geq 0$) である．なお， X の密度関数を $f(x)$ とすると

$$f(x) = \frac{d}{dx}F(x) = -\frac{d}{dx}(1 - F(x)) = -\frac{d}{dx}\bar{F}(x) = -\bar{f}(x)$$

であるので次式を得る．

$$F(x) = 1 - e^{-f(0)x}, \quad x \geq 0$$

問 2.1. $\sum_{k=0}^{\infty} kp(k) = \sum_{k=1}^{\infty} kp(k) = \sum_{k=1}^{\infty} \exp(-\lambda)\lambda^k/(k-1)! = \lambda \exp(-\lambda) \sum_{k=1}^{\infty} \lambda^{k-1}/(k-1)! = \lambda \exp(-\lambda) \sum_{k=0}^{\infty} \lambda^k/k! = \lambda$.

問 2.2. 密度関数 $f(x)$ は $f(x) = \mu \exp(-\mu x)$. $E(X) = \int_0^{\infty} x \cdot \mu \exp(-\mu x) dx = [-x \exp(-\mu x)]_0^{\infty} + \int_0^{\infty} \exp(-\mu x) dx = 1/\mu$.

問 2.3. $P(N = k) = (1/2)^k$ なので $\sum_{k=1}^{\infty} 2^k P(N = k) = \sum_{k=1}^{\infty} 1 = \infty$ となり，報酬の期待値は無限大 .

問 2.4. 取り出されたくじの種類を K とすると

$$E(K) = E(X_1 + X_2 + \cdots + X_{10}) = E(X_1) + E(X_2) + \cdots + E(X_{10})$$

一方， $P(X_i = 0) = 1 - P(X_i = 1) = (9/10)^k$ より

$$E(X_i) = 0 \times P(X_i = 0) + 1 \times P(X_i = 1) = 1 - (9/10)^k$$

よって $E(K) = 10 \cdot (1 - (9/10)^k)$.

問 2.5. 一致する枚数を J とおくと

$$E(J) = E(X_1 + X_2 + \cdots + X_N) = E(X_1) + E(X_2) + \cdots + E(X_N)$$

一方， $P(X_i = 1) = 1/N$ より $E(X_i) = 1/N$. よって $E(J) = N \times 1/N = 1$.

問 2.6. (1) $b = E[r_n] = (r^2 + 1)/(4r)$. $a_{n+1} = (1 + r_n)a_n - r_n a_{n-1}$ の両辺の期待値をとると $p_{n+1} = (1 + b)p_n - bp_{n-1}$. これを変形して $p_{n+1} - p_n = b(p_n - p_{n-1}) = b^{n-1}(p_2 - p_1) = b^{n-1}r$. よって $p_n = (p_n - p_{n-1}) + (p_{n-1} - p_{n-2}) + \cdots + (p_2 - p_1) + p_1 = r(1 - b^{n-1})/(1 - b)$. (2) $b < 1$ であればよい . $2 - \sqrt{3} < r < 2 + \sqrt{3}$.

問 2.7. 分布関数を $F(x) = P(X \leq x)$, 密度関数を $f(x) = dF(x)/dx$ とする . このとき

$$\int_0^{\infty} P(X > x) dx = \int_0^{\infty} (1 - F(x)) dx = [x(1 - F(x))]_0^{\infty} + \int_0^{\infty} x f(x) dx = [x(1 - F(x))]_0^{\infty} + E[X]$$

もし $E[X] = \infty$ ならば

$$[x(1 - F(x))]_0^{\infty} = \lim_{x \rightarrow \infty} x(1 - F(x)) \geq 0$$

より題意は成立 . 一方， $E[X]$ が有限ならば右辺第 1 項は

$$[x(1 - F(x))]_0^{\infty} = \lim_{x \rightarrow \infty} x(1 - F(x)) \geq 0$$

かつ

$$\lim_{x \rightarrow \infty} x(1 - F(x)) = \lim_{x \rightarrow \infty} x \int_x^{\infty} f(y) dy = \lim_{x \rightarrow \infty} \int_x^{\infty} x f(y) dy \leq \lim_{x \rightarrow \infty} \int_x^{\infty} y f(y) dy$$

である . ここで

$$\int_0^{\infty} y f(y) dy = \int_0^x y f(y) dy + \int_x^{\infty} y f(y) dy = E[X] < \infty$$

かつ，

$$\int_0^{\infty} y f(y) dy = \lim_{x \rightarrow \infty} \int_0^x y f(y) dy = E[X] < \infty$$

より,

$$\lim_{x \rightarrow \infty} \int_x^{\infty} y f(y) dy = 0$$

を得る. よって $[x(1 - F(x))]_0^{\infty} = \lim_{x \rightarrow \infty} x(1 - F(x)) = 0$ となり題意が成立.

問 2.8. (i) 投げる回数を N_1 としたとき, $N_1 = n$ ($n = 1, 2, \dots$) は, 裏が連続して $n-1$ 回出た後, 表が出ることに等価なので $P(N_1 = n) = (1-p)^{n-1}p$. よって $M_1 = \sum_{n=1}^{\infty} n(1-p)^{n-1}p = 1/p$. (ii) k 回 ($k = 1, 2, \dots$) 連続して表が出たとき停止する場合に硬貨を投げる回数を N_k とする. 初めて表が $k-1$ 回続いて出た後, 次に表が出れば (確率 p) $N_k = N_{k-1} + 1$ となる. もし, 次に裏がでれば (確率 $1-p$), $N_{k-1} + 1$ 回投げた後, 改めて最初からやり直すことになるので, この場合は, N_k と同じ分布に従う確率変数 N'_k を用いて, $N_k = N_{k-1} + 1 + N'_k$ と書くことができる. $M_{k-1} = E[N_{k-1}]$, $M_k = E[N_k] = E[N'_k]$ に注意すると

$$M_k = pE[N_{k-1} + 1] + (1-p)E[N_{k-1} + 1 + N'_k] = p(M_{k-1} + 1) + (1-p)(M_{k-1} + 1 + M_k) = M_{k-1} + 1 + (1-p)M_k$$

となり, $M_k = M_{k-1}/p + 1/p$ を得る. (iii) 問 (ii) の結果から, 任意の k ($k = 1, 2, \dots$) に対して

$$M_k = \frac{1}{p} + \frac{1}{p^2} + \dots + \frac{1}{p^k} \quad (31)$$

となることが予想される.

(等比級数の和の公式を使う場合は $p = 1$ と $p < 1$ の区別が必要)

$k = 1$ のとき, 問 (a) の結果から (31) は成立. ある $k = n$ に対して (31) が成立すると仮定すると,

$$M_{n+1} = \frac{1}{p} \left[\frac{1}{p} + \frac{1}{p^2} + \dots + \frac{1}{p^n} \right] + \frac{1}{p} = \frac{1}{p} + \frac{1}{p^2} + \dots + \frac{1}{p^n} + \frac{1}{p^{n+1}}$$

となり, $k = n + 1$ でも (31) が成立. よって (31) が全ての k ($k = 1, 2, \dots$) で成立.

問 2.9. 例 2.1 ならびに問 1.28 の結果を用いると $E(N) = \sum_{n=0}^{\infty} P(N > n) = \sum_{n=0}^{\infty} (n+1)^{-1} > \int_0^{\infty} (x+1)^{-1} dx = \infty$ より $E(N)$ は無限大.

問 2.10. (a)

$$\begin{aligned} P(N = k, X_N \leq x | X_0 = y) &= P(X_n \leq X_0 \ (n = 1, 2, \dots, k-1), X_0 < X_N \leq x | X_0 = y) \\ &= \prod_{n=1}^{k-1} P(X_n \leq y) \cdot P(y < X_N \leq x) = F^{k-1}(y)(F(x) - F(y)) \end{aligned}$$

なので

$$P(N = k, X_N \leq x) = \int_{-\infty}^x F^{k-1}(y)(F(x) - F(y))dF(y) = F(x) \int_{-\infty}^x F^{k-1}(y)dF(y) - \int_{-\infty}^x F^k(y)dF(y)$$

となる. ここで部分積分を用いると

$$\int_{-\infty}^x F^k(y)dF(y) = [F^k(y)F(y)]_{-\infty}^x - k \int_{-\infty}^x dF(y)F^{k-1}(y)F(y)$$

より $\int_{-\infty}^x F^k(y)dF(y) = F^{k+1}(x)/(k+1)$ を得る. よって

$$P(N = k, X_N \leq x) = F(x) \cdot \frac{F^k(x)}{k} - \frac{F^{k+1}(x)}{k+1} = \frac{F^{k+1}(x)}{k(k+1)}$$

(a) の別解: $X_0 \leq x, X_1 \leq x, \dots, X_k \leq x$ となる場合において, さらに $X_1, X_2, \dots, X_{k-1} \leq X_0 \leq X_k$ のなるものを考えればよい. まず

$$P(X_0 \leq x, X_1 \leq x, \dots, X_k \leq x) = P(X_1 \leq x)P(X_2 \leq x) \cdots P(X_k \leq x) = F(x)^{k+1}(x)$$

である, さらに X_i ($i = 0, 1, \dots, K$ の昇順の並べ方は $(k+1)!$ 通りあり, 同様に確からしい. さらに, そのうちで $X_1, X_2, \dots, X_{k-1} \leq X_0 \leq X_k$ となる並べ方は $(k-1)!$ 通りあるので, 求める解は

$$F^{k+1}(x) \cdot \frac{(k-1)!}{(k+1)!} = \frac{F^{k+1}(x)}{k(k+1)}$$

(b)

$$P(N = k) = \lim_{x \rightarrow \infty} P(N = k, X_N \leq x) = \frac{1}{k(k+1)}$$

なので

$$E[N] = \sum_{k=1}^{\infty} kP(N = k) = \sum_{k=1}^{\infty} \frac{1}{k+1} \geq \int_1^{\infty} \frac{1}{x+1} dx = [\log(x+1)]_1^{\infty} = \infty$$

より, $E[N] = \infty$.

(c)

$$\begin{aligned} G(x) &= \sum_{k=1}^{\infty} P(N = k, X_N \leq x) = \sum_{k=1}^{\infty} \frac{F^{k+1}(x)}{k(k+1)} = \sum_{k=1}^{\infty} \frac{F^{k+1}(x)}{k} - \sum_{k=1}^{\infty} \frac{F^{k+1}(x)}{k+1} = F(x) \sum_{k=1}^{\infty} \frac{F^k(x)}{k} - \sum_{k=2}^{\infty} \frac{F^k(x)}{k} \\ &= F(x) \sum_{k=1}^{\infty} \frac{F^k(x)}{k} - \sum_{k=1}^{\infty} \frac{F^k(x)}{k} + F(x) = F(x) + (F(x) - 1) \sum_{k=1}^{\infty} \frac{F^k(x)}{k} = F(x) + (1 - F(x)) \log(1 - F(x)) \end{aligned}$$

(d) X_0 と X_1 の結合密度は $f(x, y) = \mu \exp(-\mu x) \cdot \mu \exp(-\mu y)$ であり,

$$P(N \geq 2) = 1 - P(N = 1) = 1 - \frac{1}{1 \cdot 2} = \frac{1}{2} \quad \left(= P(X_1 \leq X_0) = \int_0^{\infty} \mu e^{-\mu x} \cdot e^{-\mu x} dx \right)$$

に注意すると

$$\begin{aligned} E[X_1 | N \geq 2] &= \frac{1}{P(N \geq 2)} \int_{x=0}^{\infty} \int_{y=0}^x y \cdot \mu e^{-\mu x} \cdot \mu e^{-\mu y} dx dy = 2 \int_{y=0}^{\infty} y \cdot \mu e^{-\mu y} \int_{x=y}^{\infty} \mu e^{-\mu x} dx dy \\ &= 2 \int_{y=0}^{\infty} y \cdot \mu e^{-\mu y} e^{-\mu y} dy = \int_{y=0}^{\infty} y \cdot 2\mu e^{-2\mu y} dy = \frac{1}{2\mu} \end{aligned}$$

問 2.11. (i) $P(M(x) > 1) = P(U_1 \leq x) = \frac{x}{1} = x$ (ii)

$$P(M(x) > n) = P(U_1 + U_2 + \dots + U_n \leq x) = \int_0^x P(U_1 + U_2 + \dots + U_{n-1} \leq u | U_n = u) \frac{du}{1} = \int_0^x P(M(x-u) > n-1) du$$

なので, $n = 2$ のとき $P(M(x) > 2) = x^2/2$, さらにこれを再度用いて, $P(M(x) > 3) = x^3/6$. よって $P(M(x) > n) = x^n/n!$ と仮定すると

$$P(M(x) > n+1) = \int_0^x P(M(x-u) > n) du = \int_0^x \frac{(x-u)^n}{n!} du = \int_0^x \frac{u^n}{n!} du = \frac{x^{n+1}}{(n+1)!}$$

となり, $n+1$ でも成立. よって

$$P(M(x) > n) = \frac{x^n}{n!}$$

が任意の自然数 n で成立 (iii) $P(M(x) > 0) = 1$ に注意して

$$E(M(x)) = \sum_{n=0}^{\infty} P(M(x) > n) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x$$

問 2.12.

$$P(Z \leq z) = \iint_{x^2+y^2 \leq z} \frac{\beta}{\sqrt{\pi}} e^{-\beta^2 x^2} \cdot \frac{\beta}{\sqrt{\pi}} e^{-\beta^2 y^2} dx dy = \frac{\beta^2}{\pi} \iint_{x^2+y^2 \leq z} e^{-\beta^2(x^2+y^2)} dx dy$$

ここで変数変換を行う.

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

ヤコビアンは

$$J = \begin{vmatrix} \frac{dx}{dr} & \frac{dy}{dr} \\ \frac{dx}{d\theta} & \frac{dy}{d\theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & \sin \theta \\ -r \sin \theta & r \cos \theta \end{vmatrix} = r(\cos^2 \theta + \sin^2 \theta) = r$$

なので

$$\begin{aligned} P(Z \leq z) &= \frac{\beta^2}{\pi} \iint_{x^2+y^2 \leq z} e^{-\beta^2(x^2+y^2)} dx dy = \frac{\beta^2}{\pi} \int_{r=0}^{\sqrt{z}} \int_{\theta=0}^{2\pi} r e^{-\beta^2 r^2} dr d\theta \\ &= \int_{r=0}^{\sqrt{z}} 2\beta^2 r e^{-\beta^2 r^2} dr = \left[-e^{-\beta^2 r^2} \right]_0^{\sqrt{z}} = 1 - e^{-\beta^2 z} \quad (z \geq 0) \end{aligned}$$

これはパラメタ β^2 の指数分布．平均と 2 次積率は

$$\int_0^\infty z \cdot \beta^2 e^{-\beta^2 z} dz = \frac{1}{\beta^2}, \quad \int_0^\infty z^2 \cdot \beta^2 e^{-\beta^2 z} dz = \frac{2}{\beta^4},$$

となるので分散は $2/\beta^4 - (1/\beta^2)^2 = 1/\beta^4$.

問 2.13.

$$\begin{aligned} V(XY) &= E(\{XY - E(XY)\}^2) = E(\{XY - E(X)E(Y)\}^2) \\ &= E(\{X - E(X)\}\{Y - E(Y)\} + E(X)\{Y - E(Y)\} + E(Y)\{X - E(X)\})^2 \\ &= E(\{X - E(X)\}^2\{Y - E(Y)\}^2) + \{E(X)\}^2 E(\{Y - E(Y)\}^2) + \{E(Y)\}^2 E(\{X - E(X)\}^2) \\ &= E(\{X - E(X)\}^2) E(\{Y - E(Y)\}^2) + \{E(X)\}^2 V(Y) + \{E(Y)\}^2 V(X) \\ &= V(X)V(Y) + \{E(X)\}^2 V(Y) + \{E(Y)\}^2 V(X) \end{aligned}$$

問 2.14. $E(Y) = E((X - E(X))/\sqrt{V(X)}) = E((X - E(X)))/\sqrt{V(X)} = [E(X) - E(X)]/\sqrt{V(X)} = 0$. $V(Y) = V(X - E(X))/V(X) = V(X)/V(X) = 1$.

問 2.15. 平均 $1/\mu$, 分散 $1/\mu^2$, 変動係数 1 .

問 2.16. $\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y))) = E(XY - E(X)Y - XE(Y) + E(X)E(Y)) = E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y)$ から導かれる .

問 2.17. (a) $P(X = i) = 1/3$ ($i = 1, 2, 3$) , $P(Y = 0) = 1/3$, $P(Y = 1) = 2/3$ なので , 例えば $P(X = 1, Y = 1) = 1/3 \neq 2/9 = P(X = 1)P(Y = 1)$ より , X と Y は独立でない . (b) $E(X) = 2$, $E(Y) = 2/3$, $E(XY) = 1 \times 1/3 + 0 \times 1/3 + 3 \times 1/3 = 4/3$ なので $E(X)E(Y) = E(XY)$ が成立 . よって無相関 .

問 2.18. $Y_k = X_n - E(X_n)$ とおくと $V(X_1 + \cdots + X_n) = E(\{X_1 + \cdots + X_n - E(X_1 + \cdots + X_n)\}^2) = E(\{X_1 + \cdots + X_n - E(X_1) - E(X_2) - \cdots - E(X_n)\}^2) = E((Y_1 + \cdots + Y_n)^2) = E(Y_1^2 + \cdots + Y_n^2 + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n Y_i Y_j)$ となり , 期待値の線形性から式 (21) を得る .

問 2.19.

$$\begin{aligned}
 E(S_N) &= \sum_{n=0}^{\infty} E(X_1 + \cdots + X_n | N = n)P(N = n) \\
 &= \sum_{n=0}^{\infty} [E(X_1 | N = n) + \cdots + E(X_n | N = n)]P(N = n) \\
 &= \sum_{n=0}^{\infty} [E(X_1) + \cdots + E(X_n)]P(N = n) = \sum_{n=0}^{\infty} nE(X)P(N = n) = E(N)E(X) \\
 V(S_N) &= E(S_N^2) - E(S_N)^2 \\
 &= \sum_{n=0}^{\infty} E(\{X_1 + \cdots + X_n\}^2 | N = n)P(N = n) - E(N)^2E(X)^2 \\
 &= \sum_{n=0}^{\infty} E(X_1^2 + \cdots + X_n^2 + 2X_1X_2 + \cdots + 2X_{n-1}X_n | N = n)P(N = n) - E(N)^2E(X)^2 \\
 &= \sum_{n=0}^{\infty} [n(V(X) + E(X)^2) + n(n-1)E(X)^2]P(N = n) - m^2E(N)^2 \\
 &= V(X)E(N) + E(N^2)E(X)^2 - E(N)^2E(X)^2 = V(X)E(N) + V(N)E(X)^2
 \end{aligned}$$

問 2.20. (1) 題意より

$$P(N = k) = \exp[-\lambda a](\lambda a)^k/k!$$

さらに

$$E(N) = \sum_{k=0}^{\infty} kP(N = k) = \lambda a, \quad E(N(N-1)) = \sum_{k=0}^{\infty} k(k-1)P(N = k) = (\lambda a)^2,$$

なので

$$V(N) = E(N^2) - E(N)^2 = E(N(N-1)) + E(N) - E(N)^2 = \lambda a$$

(2)

$$P(N = k) = P(X = a)P(N = k | X = a) + P(X = b)P(N = k | X = b)$$

(以下略, 代入せよ)

(3)

$$E(N) = \sum_{k=0}^{\infty} kP(N = k) = P(X = a) \sum_{k=0}^{\infty} kP(N = k | X = a) + P(X = b) \sum_{k=0}^{\infty} kP(N = k | X = b) = p\lambda a + (1-p)\lambda b$$

(4)

$$E(N) = \sum_{n=1}^{\infty} P(X = n) \sum_{k=0}^{\infty} kP(N = k | X = n) = \sum_{n=1}^{\infty} (1-p)p^{n-1} \cdot \lambda n = \frac{\lambda}{1-p}$$

(5) 密度関数は分布関数を微分して $f(x) = \mu \exp[-\mu x]$ ($x \geq 0$). $E(X) = \int_0^{\infty} xf(x)dx = 1/\mu$, $E(X^2) = \int_0^{\infty} x^2f(x)dx = 2/\mu^2$ より $V(X) = E(X^2) - E(X)^2 = 1/\mu^2$, $P(N = k) = \int_0^{\infty} P(N = k | X = x)f(x)dx$ に注意.

$$P(N = 0) = \int_0^{\infty} e^{-\lambda x} \mu e^{-\mu x} dx = \mu \int_0^{\infty} e^{-(\lambda+\mu)x} dx = \frac{\mu}{\lambda + \mu}$$

$k \geq 1$ に対しては部分積分することで

$$\begin{aligned}
 P(N = k) &= \int_0^{\infty} e^{-\lambda x} \frac{(\lambda x)^k}{k!} \mu e^{-\mu x} dx = \frac{\lambda^k}{k!} \mu \int_0^{\infty} x^k e^{-(\lambda+\mu)x} dx \\
 &= \frac{\lambda^k}{k!} \mu \left(\left[-\frac{x^k}{\lambda + \mu} e^{-(\lambda+\mu)x} \right]_0^{\infty} + \frac{k}{\lambda + \mu} \int_0^{\infty} x^{k-1} e^{-(\lambda+\mu)x} dx \right) = \frac{\lambda}{\lambda + \mu} \int_0^{\infty} \frac{(\lambda x)^{k-1}}{(k-1)!} \mu e^{-(\lambda+\mu)x} dx \\
 &= \frac{\lambda}{\lambda + \mu} P(N = k-1) = \left(\frac{\lambda}{\lambda + \mu} \right)^2 P(N = k-2) = \cdots = \left(\frac{\lambda}{\lambda + \mu} \right)^k P(N = 0)
 \end{aligned}$$

となり，最終的に

$$P(N = k) = \left(\frac{\lambda}{\lambda + \mu} \right)^k \frac{\mu}{\lambda + \mu}, \quad k = 0, 1, \dots$$

(6) $E(N) = \lambda/\mu, V(N) = \lambda(\lambda + \mu)/[\mu^2]$.

(7)

$$E(XN) = \int_0^\infty x\mu e^{-\mu x} \sum_{k=0}^\infty k e^{-\lambda x} \frac{(\lambda x)^k}{k!} dx = \int_0^\infty x\mu e^{-\mu x} \cdot \lambda x dx = \lambda \int_0^\infty x^2 \mu e^{-\mu x} dx = \frac{2\lambda}{\mu^2}$$

よって $\text{Cov}(X, N) = E(XN) - E(X)E(N) = \lambda/\mu^2 (= E(X)E(N))$.

問 2.21. まず対称性を用いて $E(Z) = 0$ を示す .

$$\begin{aligned} E(Z) &= \int_{-\infty}^\infty x f_Z(x) dx = \int_{-\infty}^0 x f_Z(x) dx + \int_0^\infty x f_Z(x) dx = \int_\infty^0 (-y) f_Z(y) (-1) dy + \int_0^\infty x f_Z(x) dx \\ &= - \int_0^\infty y f_Z(y) dy + \int_0^\infty x f_Z(x) dx = 0 \end{aligned}$$

このとき X の密度関数を $f_X(x)$ とすると

$$\begin{aligned} E(Y | X = x) &= E(Z/X | X = x) = E(Z/x) = (1/x)E(Z) = 0 \\ E(Y) &= \int_{-\infty}^\infty E(Y | X = x) f_X(x) dx = 0 \end{aligned}$$

問 2.22. (a) X の確率密度関数 $f(x)$ は $f(x) = \lambda \exp(-\lambda x)$ である . $E[X] = \int_0^\infty x f(x) dx = 1/\lambda$ なので

$$\begin{aligned} V[X] &= \int_0^\infty (x - E[X])^2 f(x) dx = \int_0^\infty x^2 f(x) dx - (E[X])^2 = \int_0^\infty x^2 \lambda e^{-\lambda x} dx - (E[X])^2 \\ &= [-x^2 e^{-\lambda x}]_0^\infty + \int_0^\infty 2x e^{-\lambda x} dx - (E[X])^2 = 2/\lambda^2 - 1/\lambda^2 = (1/\lambda)^2 \end{aligned}$$

(b) $Z_i (i = 1, 2, \dots, n)$ を互いに独立，かつ，同じ平均 $1/\lambda$ をもつ指数分布に従う確率変数とする . $i = 1$ のとき，

$$\begin{aligned} \Pr(Y_{n,1} \leq x) &= n \Pr(Z_1 \leq x, Z_1 < Z_j (j = 2, 3, \dots, n)) = n \int_0^x f(y) (1 - F(y))^{n-1} dy \\ &= n \int_0^x \lambda e^{-\lambda y} dy = 1 - e^{-n\lambda x} \end{aligned}$$

$i = 2, 3, \dots, m$ のとき，指数分布の無記憶性より

$$\Pr(Y_{n,i} \leq x) = (n + 1 - i) \Pr(Z_i \leq x, Z_i < Z_j (j = i + 1, i + 2, \dots, n)) = 1 - e^{-(n-i+1)\lambda x}$$

以上より， $Y_{n,i} (i = 1, 2, \dots, m)$ は平均 $1/[(n + 1 - i)\lambda]$ をもつ指数分布に従う .

(b) の別解 . $i = 2, 3, \dots, m$ のとき，

$$\begin{aligned} \Pr(Y_{n,i} \leq x) &= n(n-1) \binom{n-2}{i-2} \\ &\quad \cdot \Pr(Z_j < Z_{i-1} (j = 1, 2, \dots, i-1), Z_i - Z_{i-1} \leq x, Z_i < Z_j (j = i+1, i+2, \dots, n)) \\ &= n(n-1) \binom{n-2}{i-2} \int_{y=0}^\infty F^{i-2}(y) f(y) \int_{z=0}^x f(y+z) (1 - F(y+z))^{n-i} dy dz \\ &= n(n-1) \binom{n-2}{i-2} \int_{y=0}^\infty F^{i-2}(y) f(y) \int_{z=0}^x \lambda e^{-\lambda(y+z)} e^{-(n-i)\lambda(y+z)} dy dz \\ &= n(n-1) \binom{n-2}{i-2} \int_{y=0}^\infty F^{i-2}(y) f(y) e^{-(n-i+1)\lambda y} \int_{z=0}^x \lambda e^{-(n-i+1)\lambda z} dy dz \\ &= \frac{n(n-1)}{n-i+1} \binom{n-2}{i-2} \int_{y=0}^\infty F^{i-2}(y) f(y) (1 - F(y))^{n-i+1} dy \times (1 - e^{-(n-i+1)\lambda x}) \\ &= n \binom{n-1}{i-2} \int_{y=0}^\infty F^{i-2}(y) f(y) (1 - F(y))^{n-i+1} dy \times (1 - e^{-(n-i+1)\lambda x}) \\ &= 1 - e^{-(n-i+1)\lambda x} \end{aligned}$$

(c) 定義より

$$X_{n,i} = Y_{n,1} + Y_{n,2} + \cdots + Y_{n,i} \quad (i = 1, 2, \dots, m)$$

よって

$$W_{n,m} = a_{n,1}Y_{n,1} + a_{n,2}(Y_{n,1} + Y_{n,2}) + \cdots + a_{n,m}(Y_{n,1} + Y_{n,2} + \cdots + Y_{n,m}) = \sum_{i=1}^m \left(\sum_{j=1}^i a_{n,j} \right) Y_{n,i}$$

である． $Y_{n,i}$ ($i = 1, 2, \dots, m$) は互いに独立な確率変数であり，任意の実数 a に対して $V[aY_{n,i}] = a^2V[Y_{n,i}]$ が成立することに注意すると

$$V[W_{n,m}] = \sum_{i=1}^m V \left[\left(\sum_{j=1}^i a_{n,j} \right) Y_{n,i} \right] = \sum_{i=1}^m \left(\sum_{j=1}^i a_{n,j} \right)^2 V[Y_{n,i}]$$

(d) 表記を簡略化するため

$$b_i = \sum_{j=i}^m a_{n,j}$$

定義する．このとき

$$E[W_{n,m}] = \sum_{i=1}^m b_i E[Y_{n,i}] \quad V[W_{n,m}] = \sum_{i=1}^m b_i^2 V[Y_{n,i}]$$

となる．さらに設問 (a), (b) の結果より， $i = 1, 2, \dots, m$ に対して

$$E[Y_{n,i}] = \frac{1}{(n-i+1)\lambda}, \quad V[Y_{n,i}] = \left(\frac{1}{(n-i+1)\lambda} \right)^2 = (E[Y_{n,i}])^2$$

である．すなわち，

$$E[W_{n,m}] = \sum_{i=1}^m b_i E[Y_{n,i}] = \frac{1}{\lambda}, \tag{A}$$

という条件下で

$$V[W_{n,m}] = \sum_{i=1}^m b_i^2 V[Y_{n,i}] = \sum_{i=1}^m (b_i E[Y_{n,i}])^2$$

を最小化することになる．ここで，問題文で与えられた不等式において，

$$x_i = b_i E[Y_{n,i}], \quad y_i = 1 \quad (i = 1, 2, \dots, m)$$

とすると

$$mV[W_{n,m}] = m \sum_{i=1}^m (b_i E[Y_{n,i}])^2 \geq \left(\sum_{i=1}^m b_i E[Y_{n,i}] \right)^2 = \left(\frac{1}{\lambda} \right)^2$$

が成立し， $b_i E[Y_{n,i}]$ が i に依らず，定数であるとき， $V[W_{n,m}]$ が最小となる．よって，不偏推定量であるという条件 (A) を勘案すると，

$$b_i E[Y_{n,i}] = \frac{b_i}{(n-i+1)\lambda} = \frac{1}{m\lambda} \quad (i = 1, 2, \dots, m)$$

となり，

$$b_i = \frac{n-i+1}{m}, \quad i = 1, 2, \dots, m$$

である．よって

$$a_{n,i} = b_i - b_{i+1} = \frac{1}{m} \quad (i = 1, 2, \dots, m-1), \quad a_{n,m} = b_m = \frac{n-m+1}{m} \tag{B}$$

を得る．逆に (B) に従って $a_{n,i}$ を定めると，導出手順より $E[W_{n,m}] = 1/\lambda$ かつ $V[W_{n,m}]$ が最小値 $V[W_{n,m}] = (1/\lambda)^2/m$ を取ることが分かる．

(e) 利点は以下の通りである．一般に， n 台用意した際， m 台目が故障する時間の平均は

$$E[X_{n,m}] = \sum_{i=1}^m E[Y_{n,i}] = \left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{n-m+1} \right) \frac{1}{\lambda}$$

であり, $n > m$ とすることで, 故障までの時間のサンプルを m 個集める時間が短縮できる. 特に, 故障が起こりにくいと想定される場合, $n > m$ とすることは有効である. 一方, 欠点は以下の通りである. $n = m$ の場合と比較して, $n > m$ のときには, 故障するまでの時間を調べるために, $n - m$ 台, 多くの機械を用意する必要がある. なお, 推定量の分散 $V[W_{n,m}] = (1/\lambda)^2/m$ は n に依存しないため, 推定量の善し悪しという点では差がない. また, 故障までの時間のサンプルを m 個集めるまでに必要な延べ稼働時間の平均は

$$\sum_{i=1}^m (n+1-i)E[Y_{n,i}] = \left(n \frac{1}{n\lambda} + (n-1) \frac{1}{(n-1)\lambda} + \cdots + (n-m+1) \frac{1}{(n-m+1)\lambda} \right) = \frac{m}{\lambda}$$

であり, n に依存しないため, 機械を動かすための平均コスト (電気代あるいは燃料代) は (止まっている機械を稼働させるときに必要な初期エネルギーを無視すると) 同じである.

問 4.1. m 本のくじを区別すると引き方の場合の数は ${}_m C_3 = m(m-1)(m-2)/6$ 通り. 一方, 当たりが 2 本の場合の数は ${}_n C_2 = n(n-1)/2$ であり, 外れの場合の数は ${}_{m-n} C_1 = m-n$. よって, ${}_{m-n} C_1 \cdot {}_n C_2 / {}_m C_3 = 3n(n-1)(m-n)/[m(m-1)(m-2)]$

問 4.2. (i) $7/18$, (ii) $5/36$, (iii) $1/9$, (iv) 7 , (v) $49/4$

問 4.3. (i) $(5/6)^N$, (ii) $P(M=5) = P(M \leq 5) - P(M \leq 4) = (5/6)^N - (2/3)^N$

問 4.4. 和の最大値 18 . 余事象を考える. $53/54$

問 4.5. (i) 原点に戻ってくるためには 3 回中 1 回が 4 以下. ${}_3 C_1 (2/3)^1 (1/3)^2 = 2/9$, (ii) 4 以下がでる回数を n とすると $(5-n) \times 1 + n \times (-2) = -1$ より, $n = 2$. ${}_5 C_2 (2/3)^2 (1/3)^3 = 40/243$

問 4.6. たとえば ${}_{M+N-2} C_{M-2} / {}_{M+N} C_M = [M(M-1)] / [(M+N)(M+N-1)]$. 他にも色々な考え方あり.

問 4.7. (i) たとえば M 本を 1 番から M 番まで並べて, そこから最初の K 番目までを順に引くと考える. ${}_{M-1} C_{N-1} / {}_M C_N = N/M$, (ii) 余事象を考える. $1 - {}_{M-K} C_N / {}_M C_N = 1 - [(M-N)(M-N-1) \cdots (M-N-K+1)] / [M(M-1) \cdots (M-K+1)]$

問 4.8. (i) 勝ち方はグー, チョキ, パーで 3 通り. 一人勝ちするのは N 通り. $(3 \cdot N)/3^N$, (ii) $[{}_3 C_2] / 3^N = N(N-1)/(2 \cdot 3^{N-1})$

問 4.9. 表が N 回でた時点で試行をとめるとする. 試行が終わったとき裏が k 回出ている確率は, $N+k-1$ 回の間に k 回裏がでて, かつ, $N+k$ 回目に表が出ればよいので ${}_{N+k-1} C_k (1/2)^{N+k-1} \cdot (1/2)$. $k \leq M-1$ なら題意を満たす. $(1/2)^N [1 + {}_{N+1} C_1 (1/2) + {}_{N+2} C_2 (1/2)^2 + \cdots + {}_{N+M-1} C_{M-1} (1/2)^{M-1}]$

問 4.10. (i) $(N)!/N^N$, (ii) ${}_N C_k / N^N$

問 4.11. 順番も含めるとくじの引き方は $3^3 = 27$ 通りあり, 同様に確からしい. 1 種類ずつ引く場合の数は $3! = 6$ 通り. 一方, 1 種類しかない場合の数は 3 通り. よって 2 種類の場合の数は $27 - 6 - 3 = 18$ 通り. 以上より $1 \times 3/27 + 2 \times 18/27 + 3 \times 6/27 = 19/9$. なお, 2 種類の場合の数は, 2 種類の選び方が ${}_3 C_2 = 3$ 通り. それぞれに対して, どちらが 1 回かで 2 通り. さらにそれらの並べ方が ${}_3 C_1 = 3$ 通り. よって $3 \times 2 \times 3 = 18$ 通り, としても得られる.

問 4.12. いかなる時点においても, その時点での合計が m の倍数でない場合, 次に引いた時点で m の倍数になる確率は $1/m$. よって, k 回 ($k = 1, 2, \dots, n-1$) で試行が停止する確率 p_k は $p_k = (1 - 1/m)^{k-1} (1/m)$ ($k = 1, 2, \dots, n-1$) であり, n 回試行を繰り返すということは $n-1$ 回目までに m の倍数とならなかったことと等価なので $p_n = (1 - 1/m)^{n-1}$. よって $\sum_{k=1}^{n-1} k p_k + n p_n = m[1 - (1 - 1/m)^n]$.

問 4.13. (i) A が n 回目にコインを投げたとき勝つ確率は $[(1/2)(1/2)(1/2)]^{n-1} (1/2)$ なので, A が勝つ確率は $\sum_{n=1}^{\infty} (1/8)^{n-1} (1/2) = 4/7$, (ii) 同様に B が勝つ確率は $\sum_{n=1}^{\infty} (1/8)^{n-1} (1/2)(1/2) = 2/7$

問 4.14. (i) $a_n = (1/2)b_{n-1}$, $b_n = (1/2)a_{n-1} + (1/2)b_{n-1}$ (ii) $p_n = (1/2)p_{n-1} + (1/4)p_{n-2}$

問 4.15. $l_0 = r_0 = 0, m_0 = 1 \cdot l_{n+1} = (1/2)l_n + (1/3)m_n, m_{n+1} = (1/2)l_n + (1/3)m_n + (1/2)r_n, r_{n+1} = (1/3)m_n + (1/2)r_n \cdot l_n + r_n = 1 - m_n$ なので, $m_{n+1} = (1/2)[1 - m_n] + (1/3)m_n = 1/2 - (1/6)m_n$. これを解いて $m_n = 3/7 - (2/21)(-1/6)^{n-1}$ ($n = 1, 2, \dots$).

問 4.16. (i) n 回目で終了するのは最初の $n-1$ 回で $K-1$ 回当たりが出て, かつ, n 回目で当たりが出ればよい. よって $p_n = {}_{n-1}C_{K-1}(M/N)^{K-1}(1-M/N)^{(n-1)-(K-1)} \cdot (M/N)$ (ii) $p_{n+1}/p_n = n(1-M/N)/(n-K+1) \geq 1 \Leftrightarrow n \leq N(K-1)/M$ なので, $n = \lfloor N(K-1)/M \rfloor + 1$ で最大 ($\lfloor x \rfloor$ は x の整数部分 (小数点以下切り捨て)).

問 4.17. n 回の試行の内, 4 以下が偶数回でる確率が p_n である. n 回の試行の内, 4 以下が奇数回でる確率を $q_n = 1-p_n$ とすると, $p_1 = 1/3, q_1 = 2/3$ であり, $n \geq 2$ に対して $p_n = p_{n-1}(2/3) + q_{n-1}(1/3) = (-1/3)p_{n-1} + 2/3$. を得る. よって $p_n = 1/2 + (1/2)(-1/3)^n$ ($n = 1, 2, \dots$)

問 4.18. n 回コインを投げた時点でまだ表が連続して 2 回出ておらず, かつ, n 回目に表 (裏) が出る確率を a_n (b_n) とする. $a_1 = 2/3, b_1 = 1/3, p_1 = 0$ であり, $n \geq 2$ に対して $a_n = b_{n-1}(2/3), b_n = (a_{n-1} + b_{n-1})(1/3), p_n = a_{n-1}(2/3)$ である. 最後の式から $a_{n-1} = (3/2)p_n$, さらに最初の式から $b_{n-1} = (3/2)a_n = (9/4)p_{n+1}$ なので $(9/4)p_{n+2} = (1/3)[(3/2)p_n + (9/4)p_{n+1}]$, すなわち $p_{n+2} = (1/3)p_{n+1} + (2/9)p_n$ が $n \geq 2$ で成立. $p_2 = a_1(2/3) = 4/9, p_3 = a_2(2/3) = b_1(4/9) = 4/27$ に注意して解くと, $p_n = (2/3)^{n+1} - 4(-1/3)^{n+1}$ ($n = 2, 3, \dots$)

問 4.19. 定義より $a_1 = c_1 = 0, b_1 = 2/3, d_1 = 1/3$. よって $n \geq 2$ を考える. n が偶数のとき $a_n = b_{n-1}(1/3) + d_{n-1}(2/3), c_n = b_{n-1}(2/3) + d_{n-1}(1/3), b_n = c_n$ であり, n が奇数のとき $a_n = c_n = 0, b_n = a_{n-1}(2/3) + c_{n-1}(1/3), d_n = a_{n-1}(1/3) + c_{n-1}(2/3)$ である. よって, $n = 2k$ に対して $b_{2k} = d_{2k} = 0, c_{2k} = 1 - a_{2k}$ であり, $a_{2k+2} = a_{2k}(4/9) + c_{2k}(5/9) = a_{2k}(4/9) + (1 - a_{2k})(5/9)$ ならびに $a_0 = 1$ より $a_{2k} = (1/2) + (1/2)(-1/9)^k$. 一方, $n = 2k+1$ に対して $a_{2k+1} = c_{2k+1} = 0, d_{2k+1} = 1 - b_{2k+1}$ であり, $b_{2k+1} = b_{2k-1}(4/9) + d_{2k+1}(5/9) = b_{2k-1}(4/9) + (1 - a_{2k-1})(5/9)$ ならびに $b_1 = 2/3$ より $b_{2k+1} = 1/2 + (1/6)(-1/9)^k$. 以上をまとめて
 n が偶数: $a_n = (1/2) + (1/2)(-1/9)^{n/2}, c_n = (1/2) - (1/2)(-1/9)^{n/2}, b_n = d_n = 0$
 n が奇数: $b_n = (1/2) + (1/6)(-1/9)^{(n-1)/2}, d_n = (1/2) - (1/6)(-1/9)^{(n-1)/2}, a_n = c_n = 0$

問 4.20. (i) 最初に取り出す玉の色で場合分けをする. もし, 最初に取り出す玉が白玉ならば, 2 回目の試行をする前の箱の中の状態は白玉が $k+1$ 個であり, この局面から白玉だけになる確率は p_{k+1} である. 一方, 最初に取り出す玉が赤玉ならば, 2 回目の試行をする前の箱の中の状態は白玉が $k-1$ 個であり, この局面から白玉だけになる確率は p_{k-1} である. よって $p_k = (k/n)p_{k+1} + (1-k/n)p_{k-1}$, (ii) 階差を $q_k = p_{k+1} - p_k$ とすると, (i) の結果から $q_k = [(n-k)/k]q_{k-1}$ を得る. これを右辺に繰り返し適用すると $q_k = q_0(n-1)!/[k!(n-k-1)!] = q_0 \cdot {}_{n-1}C_k$ を得る. $q_0 = p_1 - p_0 = p_1$ の注意して, $p_k = p_1 + \sum_{j=1}^{k-1} (p_{j+1} - p_j) = p_1 + \sum_{j=1}^{k-1} q_0 \cdot {}_{n-1}C_j = \sum_{j=0}^{k-1} p_1 \cdot {}_{n-1}C_j$ を得る. $k = n$ のとき $p_n = 1$ より $1 = p_1 \sum_{j=0}^{n-1} {}_{n-1}C_j = p_1(1+1)^{n-1} = p_1 2^{n-1}$ となる. よって $p_k = p_1 \sum_{j=0}^{k-1} {}_{n-1}C_j, p_1 = 1/2^{n-1}$

問 4.21. $n+1$ 回目の操作終了後に k 種類のくじが出ているという事は, n 回目の操作終了後に k 種類のくじが出ていて, かつ, $n+1$ 回目に既に取り出したくじと同じ等級のくじを引くか, あるいは, n 回目の操作終了後に $k-1$ 種類のくじが出ていて, かつ, $n+1$ 回目に初出の等級のくじを引くか, のいずれかである. よって, $p_{n+1}(k) = p_n(k)(k/N) + p_n(k-1)[(N-k+1)/N]$

問 4.22. (a) X_n が 3 の倍数ならば, $10X_n$ も 3 の倍数なので X_{n+1} は確率 $1/5$ で 3 の倍数である. X_n が 3 の倍数でなければ, $X_n = 3Y_n + i$ ($i = 1, 2$) と書くことができ, $10X_n = 30Y_n + 10i = 3(10Y_n + 3i) + i$ である. よって X_n が 3 の倍数でなければ, X_{n+1} は確率 $2/5$ で 3 の倍数となる. よって $p_{n+1} = p_n \times (1/5) + (1-p_n) \times (2/5) = 2/5 - p_n/5$. (b) 明らかに $p_1 = 1/5$. また $p_{n+1} - 1/3 = (-1/5)[p_n - 1/3]$ なので $p_n - 1/3 = (1/5 - 1/3)(-1/5)^{n-1} = (2/3)(-1/5)^n$ より, $p_n = 1/3 + (2/3)(-1/5)^n$

参考文献

- [1] Feller, W. (1971) *An Introduction to Probability Theory and Its Applications, Vol.II*. John Wiley & Sons, New York.
- [2] Gallager, R. G. (1996) *Discrete Stochastic Processes*. Kluwer, Boston.
- [3] 高橋幸雄 (2008) 確率論 . 朝倉書店 .
- [4] 瀧澤精二, 微分積分学 (上), 廣川書店, 1976.
- [5] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*, Prentice Hall, Englewood Cliffs, NJ, 1989.