

確率論 I

種村 秀紀

千葉大学理学部数学・情報数理学科

〒 263-8522 千葉県千葉市稲毛区弥生町 1-33

e-mail: tanemura@math.s.chiba-u.ac.jp

<http://www.math.s.chiba-u.ac.jp/~tanemura/index.html>

平成 27 年 4 月 15 日

目次

1	復習と準備	3
1.1	離散型確率空間	3
1.2	確率変数	4
1.3	期待値 (Expectation)	5
1.4	離散確率分布の例	6
1.4.1	ベルヌーイ分布 (Bernoulli distribution)	6
1.4.2	離散一様分布 (Uniform distribution)	7
1.4.3	二項分布 (Binomial distribution) : $B(n, p)$	8
1.4.4	幾何分布 (Geometric distribution) : $G(p)$	9
1.4.5	超幾何分布 (Hyper Geometric Distribution)	11
1.4.6	負の二項分布	13
1.4.7	ポアソン分布 (Poisson 分布) : $Po(\lambda)$	14
1.5	条件付確率と独立性	17
2	一般の確率空間	20
2.1	確率空間の定義と基本的性質	20
2.2	期待値	22
2.3	期待値に関する不等式	25
2.4	一様可積分性と収束定理	26
2.5	分布と分布関数	29
2.6	拡張定理	30
2.7	直積確率空間	34
2.8	連続確率分布の例	35
2.8.1	一様分布 (Uniform distribution)	36
2.8.2	指数分布 (Exponential distribution)	37
2.8.3	ガンマ分布 (Γ -distribution)	38
2.8.4	ベータ分布 (Beta distribution)	40
2.8.5	正規分布 (Normal distribution)	41
2.8.6	多次元正規分布 (Multi-dimensional Normal distribution)	43
2.8.7	コーシー分布 (Cauchy distribution)	43
2.8.8	対数正規分布 (log-normal distribution)	43
2.8.9	ワイブル分布 (Weibul distribution)	43

参考文献

- [1] 「確率論入門・1」, 池田信行・小倉幸雄・高橋陽一郎・真鍋昭治郎 著, 培風館.
- [2] 「確率論入門」, 笠原勇二 著, 数学書房
- [3] 「確率論」-新しい解析学の流れ, 熊谷 隆 著, 共立出版.
- [4] 「確率モデル」, 成田清正 著, 共立出版
- [5] 「ルベーグ積分から確率論」(21世紀の数学10), 志賀徳造 著, 共立出版.
- [6] 九州大学確率論講義ノート, 谷口説男 著.
- [7] 「確率論」, 西尾真喜子 著, 実教出版.
- [8] 「確率論」, 福島正俊 著裳華房.

1 復習と準備

1.1 離散型確率空間

- **標本空間** (Sample Space)

試行で得られる結果すべての集まり. 記号は Ω を用いる. そして標本空間に含まれる要素 $\omega \in \Omega$ を標本点 (Sample point) という. この章では標本空間は有限または可算無限集合を扱う. (非可算無限集合の場合は2章以降扱う.)

例)

- 1) 銅貨を1回投げるという試行に対して $\Omega = \{\text{表}, \text{裏}\}$.
- 2) 銅貨を裏が出るまで投げ続けるという試行に対して $\Omega = \{(\text{裏}), (\text{表}, \text{裏}), (\text{表}, \text{表}, \text{裏}), (\text{表}, \text{表}, \text{表}, \text{裏}), \dots\}$.
- 3) サイコロを1回投げるという試行に対して $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- 4) サイコロを3回投げるという試行に対して $\Omega = \{(i, j, k) : i, j, k = 1, 2, 3, 4, 5, 6\}$.

- **事象** (Event)

Ω の部分集合を事象という. A, B, C などの記号を用いる.

空事象: 標本点を1つも含まない場合も事象であり空事象と呼び \emptyset と書く.

全事象: 全ての標本点を含む場合も事象であり全事象と呼び, 標本空間と同じであるので Ω と書く.

根元事象: 1つの標本点からなる事象を根元事象と呼び, 例えば i を1つの標本点とすると $\{i\}$ となる.

そして事象に対してつぎの記号 (演算) を用いる.

和事象: A と B を事象としたとき A または B に属するすべての標本点からなる事象を A と B の和事象と呼び $A \cup B$ と書く. A_1, A_2, \dots, A_n の和事象は $\bigcup_{i=1}^n A_i$ と記す.

積事象: A と B を事象としたとき A かつ B に属するすべての標本点からなる事象を A と B の積事象とよび $A \cap B$ と書く. A_1, A_2, \dots, A_n の積事象は $\bigcap_{i=1}^n A_i$ と記す.

余事象: A に属するが B には属さない標本点全体からなる事象を $A - B$ と表すことにする. 特に $\Omega - A$ を A の余事象 (または補事象) とよび A^c で表す.

排反: $A \cap B = \emptyset$ のとき, 事象 A と B は排反という. この時に限り和事象 $A \cup B$ を $A + B$ と記し直和とよぶ. A_1, A_2, \dots, A_n の直和は $\sum_{i=1}^n A_i$ と記す.

対称差: $(A - B) \cup (B - A)$ で定義される事象を A と B の**対称差**とよび $A \Delta B$ で表す.

- **確率** (Probability)

標本空間上の関数で $p(\cdot)$ で

$$p(\omega) \in [0, 1], \quad \sum_{\omega \in \Omega} p(\omega) = 1,$$

を満足するものを**(離散) 確率密度関数**という. $p(\cdot)$ は標本点 ω の実現しやすさを測る量とみなせる. そして事象 A の (起こる) 確率は

$$(1.1) \quad P(A) \equiv \sum_{\omega \in A} p(\omega),$$

で与えられる. 任意の事象 $A_i, i = 1, 2, \dots, n$ に対して

$$P\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n P(A_i), \quad \text{劣加法性,}$$

A, B が排反であるとき

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i), \quad \text{有限加法性,}$$

が成り立つ. 対 (Ω, P) を離散型確率空間という.

例)

- 1) 公正な銅貨を 1 回投げるとい試行に対して $p(\text{表}) = 1/2, p(\text{裏}) = 1/2$.
- 2) 公正な銅貨を裏が出るまで投げ続けるとい試行に対して $p(\text{裏}) = 1/2,$
 $p(\text{表, 裏}) = (1/2)^2, p(\text{表, 表, 裏}) = (1/2)^3, p(\text{表, 表, 表, 裏}) = (1/2)^4, \dots$
- 3) 公正なサイコロを 1 回投げるとい試行に対して $p(\omega) = 1/6, \omega \in \Omega$.
- 4) 公正なサイコロを 3 回投げるとい試行に対して $p(\omega) = (1/6)^3, \omega \in \Omega$.

1.2 確率変数

確率変数 (random variable) X とは, 標本空間 Ω を定義域とする実数値関数である.
確率変数の例をつぎに挙げる.

- 1) 銅貨を 1 回投げるとい試行に対して表が出た回数 (表が出れば 1, 裏が出れば 0) X は確率変数である.
具体的に書いてみると $X(\text{表}) = 1, X(\text{裏}) = 0$, である.
- 2) 銅貨を裏が出るまで投げ続けるとい試行に対して裏がでるまでに表が出た回数 X は確率変数である.
具体的に書いてみると $X(\text{裏}) = 0, X(\text{表, 裏}) = 1,$
 $X(\text{表, 表, 裏}) = 2, X(\text{表, 表, 表, 裏}) = 3, \dots$, である.
- 3) サイコロを 1 回投げるとい試行に対して出た目の数 X は確率変数である.
- 4) サイコロを 3 回投げるとい試行に対して 1 回目に出た目の数 $X_1, 2$ 回目に出た目の数 $X_2, 3$ 回目に出た目の数 $X_3, 出た目の総数 X はすべて確率変数である.$

確率分布 離散確率空間上の確率変数 X の像 $X(\Omega) \equiv \{X(\omega) : \omega \in \Omega\}$ は明らかに可算集合である. $X(\Omega)$ を標本空間とし, その標本点 x にたいして

$$\mu_X(\{x\}) \equiv P(X^{-1}(x)) = P(\omega : X(\omega) = x)$$

と μ_X を定義すると,

$$\mu_X(\{x\}) \in [0, 1], \quad \sum_{x \in X(\Omega)} \mu_X(\{x\}) = 1,$$

を満たすので, μ_X は $X(\Omega)$ 上の確率であり, $(X(\Omega), \mu_X)$ は離散確率空間である. 確率 μ_X を確率変数 X の**確率法則**, または**確率分布**という.

1.3 期待値 (Expectation)

確率変数 X , 実数上で定義された関数 f に対して

$$\sum_{\omega \in \Omega} |f(X(\omega))| P(\{\omega\}) < \infty, \quad \text{絶対収束}$$

を満たすとき $f(X)$ は可積分であるといい

$$E[f(X)] \equiv \sum_{\omega \in \Omega} f(X(\omega)) P(\{\omega\})$$

を $f(X)$ の**期待値** とよぶ. また, $f(X)$ が可積分, A が事象であるとき

$$E[f(X), A] \equiv \sum_{\omega \in A} f(X(\omega)) P(\{\omega\})$$

と記す.

例)

- 1) $f(x) = x$ のとき, 期待値 $E[X] \equiv \mu$ は**平均** (Mean) である.
- 2) $f(x) = x^n$ のとき, 期待値 $E[X^n]$ は n -**次積率** (n-th moment) である.
- 3) $f(x) = (x - \mu)^2$ のとき, 期待値 $E[(X - \mu)^2] \equiv Var[X]$ は**分散** (Variance) である. 定義から明らかに $Var[X] \geq 0$ である.
- 4) $f(x) = e^{tx}$ のとき, 期待値 $E[e^{tX}] \equiv m(t)$ は**積率母関数** (Moment generating function) である.

X, Y を確率変数, f を 2次元ユークリッド空間上で定義された実数値関数としたとき, 期待値 $E[f(X, Y)]$ が定義できる. $f(x, y) = (x - E[X])(y - E[Y])$ のとき, 期待値 $E[(X - E[X])(Y - E[Y])] \equiv Cov[X, Y]$ は X, Y の**共分散** (Covariance) である. 定義から明らかに $Cov[X, Y] = Cov[Y, X]$, $Cov[X, X] = Var[X]$ である.

次に期待値に関する基本的な性質をまとめておく. 簡単であるので証明は省略する.

定理 1.1 X, Y を可積分確率変数とする.

(i) (線形性) 任意の実数 a, b に対して

$$E[aX + bY] = aE[X] + bE[Y].$$

(ii) 互いに排反である事象 A_1, A_2, \dots, A_n に対して

$$E[X, \bigcup_{i=1}^n A_i] = \sum_{i=1}^n E[X, A_i].$$

(iii) A を事象, X を $X(\omega) = a, \forall \omega$ である定数確率変数とすれば

$$E[X, A] = aP(A).$$

(iv) μ_X を確率変数 X の確率分布とする. $X(\Omega)$ 上で定義された任意の実数値関数 f に対して

$$E[f(X)] = \sum_{x \in X(\Omega)} f(x) \mu_X(\{x\}).$$

次に分散に関する基本的な性質をまとめておく.

定理 1.2 X, Y を二乗可積分確率変数とする.

(i)

$$\text{Var}[X] = E[X^2] - (E[X])^2, \quad \text{Cov}[X, Y] = E[XY] - E[X]E[Y].$$

(ii) 任意の実数 a, b に対して

$$\text{Var}[aX + bY] = a^2\text{Var}[X] + 2ab\text{Cov}[X, Y] + b^2\text{Var}[Y].$$

(iii)

$$|\text{Cov}[X, Y]|^2 \leq \text{Var}[X]\text{Var}[Y], \quad \text{シュワルツ (Schwarz) の不等式.}$$

証明 (i), (ii) は期待値の線形性を用いると簡単に示すことができる. (iii) は, (ii) で分散が非負であることに注意すれば, $b = 1$ として a の 2 次方程式の判別式から直ちに導かれる. \square

演習 1.1. 上述の定理の (iii) は, 数列 $\{a_k\}, \{b_k\}$ に関する不等式

$$\left(\sum_{k=1}^n a_k b_k \right)^2 \leq \sum_{k=1}^n a_k^2 \sum_{k=1}^n b_k^2$$

と同等である. この不等式を直接計算して示せ.

この節の最後にマルコフ (Markov) の不等式とよばれる評価式を紹介する.

定理 1.3 X を確率変数とし, n を正の整数とする. $E[|X|^n] < \infty$ (n 乗可積分) を仮定する. このとき

$$P(|X| \geq c) \leq \frac{E[|X|^n]}{c^n}, \quad \forall c > 0,$$

が成り立つ. 特に $n = 2$ の時 チェビセフ (Chebyshev) の不等式という.

証明

$$\begin{aligned} E[|X|^n] &= \sum_{\omega \in \Omega} |X(\omega)|^n p(\omega) \\ &= \sum_{\omega: |X(\omega)| \geq c} |X(\omega)|^n p(\omega) + \sum_{\omega: |X(\omega)| < c} |X(\omega)|^n p(\omega) \\ &\geq c^n \sum_{\omega: |X(\omega)| \geq c} p(\omega) \geq c^n P(|X| \geq c). \end{aligned}$$

両辺を c^n でわると定理が導かれる.

1.4 離散確率分布の例

1.4.1 ベルヌーイ分布 (Bernoulli distribution)

$$S = \{0, 1\}, \quad p(0) = 1 - p, \quad p(1) = p,$$

(ただし p は 0 以上 1 以下の実数) の時, 確率密度 $p(x)$ で定まる (離散) 確率分布をベルヌーイ分布という.

- 対応する試行

表が出る確率が p , 裏が出る確率が $1 - p$ であるコインを 1 回投げる. 表が出た回数の分布はベルヌーイ分布である.

- 平均 $\mu = 0 \cdot p(0) + 1 \cdot p(1) = p$
- 2次モーメント $m_2 = 0^2 \cdot p(0) + 1^2 \cdot p(1) = p$
- 分散 $V = m_2 - \mu^2 = p - p^2 = p(1 - p)$
- 標準偏差 $\sigma = \sqrt{V} = \sqrt{p(1 - p)}$
- 積率母関数 $m(t) = e^{t \cdot 0} p(0) + e^{t \cdot 1} p(1) = 1 \cdot (1 - p) + e^t p = 1 + (e^t - 1)p$

1.4.2 離散一様分布 (Uniform distribution)

$$S = \{1, 2, \dots, n\}, \quad p(x) = \frac{1}{n}, \quad x \in S$$

の時, 確率密度 $p(x)$ で定まる (離散) 確率分布を離散一様分布という.

• 対応する試行

各々の目が出る確率がすべて等しいサイコロを投げる. 出た目の数の分布は $n = 6$ の離散一様分布である.

- 平均 $\mu = \sum_{x=1}^n xp(x) = \frac{1}{n} \sum_{x=1}^n x = \frac{1}{n} \frac{n(n+1)}{2} = \frac{n+1}{2}$
- 2次モーメント $m_2 = \sum_{x=1}^n x^2 p(x) = \frac{1}{n} \sum_{x=1}^n x^2 = \frac{1}{n} \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$
- 分散 $V = m_2 - \mu^2 = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2$
 $= \frac{n+1}{12} \{2(2n+1) - 3(n+1)\} = \frac{n+1}{12} (n-1) = \frac{n^2-1}{12}$
- 標準偏差 $\sigma = \sqrt{V} = \sqrt{\frac{n^2-1}{12}}$
- 積率母関数 $m(t) = \sum_{x=1}^n e^{tx} p(x) = \frac{1}{n} \sum_{x=1}^n e^{tx} = \frac{e^t}{n} \cdot \frac{1 - e^{tn}}{1 - e^t}$

注意 ここで離散一様分布の積率の計算で鍵となる $\sum_{x=1}^n x^k$ に関する公式について議論する. $f(r) = \sum_{x=0}^n (1+r)^x$ とおくと, 二項定理 (後で述べる二項分布の項でも用いる) から

$$f(r) = \frac{(1+r)^{n+1} - 1}{r} = \sum_{x=1}^n \binom{n+1}{x} r^{x-1}$$

となることから, 任意の $k \in \mathbb{N}$ に対して

$$G(k) \equiv \sum_{x=0}^n x(x-1)\cdots(x-k+1) = \frac{\partial^k f}{\partial r^k}(r) \Big|_{r=0} = \binom{n+1}{k+1} k! = \frac{(n+1)!}{(n-k)!(k+1)}.$$

この結果を用いると $\sum_{x=1}^n x^k$, $k \in \mathbb{N}$ が計算できる. 例えば

$$\begin{aligned} \sum_{x=1}^n x &= G(1) = \frac{n(n+1)}{2}, \quad \sum_{x=1}^n x^2 = G(2) + G(1) = \frac{n(n+1)(2n+1)}{6}, \\ \sum_{x=1}^n x^3 &= G(3) + 3G(2) + G(1) = \left(\frac{n(n+1)}{2}\right)^2, \\ \sum_{x=1}^n x^4 &= G(4) + 6G(3) + 7G(2) + G(1) = \frac{n(n+1)(2n+1)(3n^2+3n-1)}{30}. \end{aligned}$$

が導かれる.

1.4.3 二項分布 (Binomial distribution) : $B(n, p)$

$$S = \{0, 1, 2, \dots, n\}$$

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x},$$

の時, 確率密度 $p(x)$ で定まる確率分布を二項分布という. ここで, $n! = n \cdot (n-1) \cdots 3 \cdot 2 \cdot 1$ であり n の階乗という. また $\binom{n}{x}$ は n 個の異なるものから x 個を選ぶ組み合わせの数という.

● 対応する試行

表が出る確率が p , 裏が出る確率 $1-p$ であるコインを n 回投げる. 表が出た回数の分布はパラメータ n, p の二項分布である.

二項分布の平均, 分散および積率母関数を計算するには二項定理を用いる.

$$\text{二項定理 : } (x+y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$$

$q = 1-p$ とおく. この二項定理を用いると,

$$P(S) = \sum_{x=0}^n p(x) = \sum_{x=0}^n \binom{n}{x} p^x q^{n-x} = (p+q)^n = 1$$

が直ちに導かれる.

● 平均

$$\mu = \sum_{x=0}^n x \cdot p(x) = \sum_{x=0}^n x \times \frac{n!}{x!(n-x)!} p^x q^{n-x} = \sum_{x=1}^n \frac{n!}{(x-1)!(n-x)!} p^x q^{n-x}$$

$\ell = x-1, m = n-1$ とおき, $n! = n \times (n-1)! = n \times m!$, $n-x = m-\ell$ となることに注意すると

$$= \sum_{\ell=0}^m \frac{n \times m!}{\ell!(m-\ell)!} p^{\ell+1} q^{m-\ell} = np \sum_{\ell=0}^m \frac{m!}{\ell!(m-\ell)!} p^{\ell} q^{m-\ell}$$

ここで二項定理を用いると

$$= np$$

を得る. 従って

二項分布の平均: $\mu = np$

● 分散

$V = m_2 - \mu^2$ よりまず 2 次元モーメント m_2 を計算する.

$$\begin{aligned} m_2 &= \sum_{x=0}^n x^2 p(x) = \sum_{x=0}^n x^2 \frac{n!}{x!(n-x)!} p^x q^{n-x} \\ &= \sum_{x=2}^n x(x-1) \frac{n!}{x!(n-x)!} p^x q^{n-x} + \sum_{x=1}^n x \frac{n!}{x!(n-x)!} p^x q^{n-x} \\ &= \sum_{x=2}^n \frac{n!}{(x-2)!(n-x)!} p^x q^{n-x} + np \end{aligned}$$

$\ell = x - 2, m = n - 2$ とおき, $n! = n(n-1)m!, n - x = m - \ell$ に注意すると

$$= \sum_{\ell=0}^m n(n-1)p^2 \cdot \frac{m!}{\ell!(m-\ell)!} p^\ell q^{m-\ell} + np$$

となり, ここで二項定理を用いると

$$= n(n-1)p^2 + np$$

を得る. よって,

$$V = \sigma^2 = n(n-1)p^2 + np - (np)^2 = n(p - p^2) = npq$$

となる. 従って

二項分布の分散 : $V = npq$

二項分布の標準偏差 : $\sigma = \sqrt{npq}$

● 積率母関数

二項分布の積率母関数は二項定理を用いることにより次のように簡単に計算できる.

$$\begin{aligned} m(t) &= \sum_{x=0}^n e^{tx} p(x) = \sum_{x=0}^n e^{tx} \binom{n}{x} p^x q^{n-x} \\ &= \sum_{x=0}^n \binom{n}{x} (pe^t)^x q^{n-x} = (pe^t + q)^n \end{aligned}$$

二項分布の積率母関数: $m(t) = (pe^t + q)^n$

積率母関数を β 階微分してから 0 を代入すると, β 次モーメント (β 次積率) となる. 二項分布の場合に平均について調べてみると

$$\begin{aligned} \frac{d}{dt} m(t) &= \frac{d}{dt} (pe^t + q)^n = n(pe^t + q)^{n-1} pe^t \\ &\quad t=0 \text{ とすれば} \\ &= n(p+q)^{n-1} p = np \end{aligned}$$

となることが確認できる. 2次モーメントについても同様に計算できる. (各自確認するように.) 上の計算では合成関数の微分に関する公式

$$\text{合成関数の微分} \quad \frac{d}{dt} F(g(t)) = F'(g(t))g'(t)$$

を $F(x) = x^n, g(t) = pe^t + q$ に対して適用している.

考察 二項分布は $n = 1$ という特別な場合はベルヌーイ分布である. そして一般の n の場合での積率母関数はベルヌーイ分布の積率母関数の n 乗となっている.

1.4.4 幾何分布 (Geometric distribution) : $G(p)$

S の標本点が可算無限個つまり

$$S = \{0, 1, 2, 3, \dots\}$$

の場合にも離散確率分布が定義できる. p を $0 < p < 1$ である実数とし

$$p(x) = p(1-p)^x, \quad x \in S$$

で定義される離散確率密度で定まる (離散) 確率分布を幾何分布という.

- 対応する試行

表が出る確率が p , 裏が出る確率が $1-p$ であるコインを投げる試行を行う. 1 回目の試行で

$$(*) \begin{cases} \text{表が出た場合そこで終了する.} \\ \text{裏が出た場合もう一度コインを投げる.} \end{cases}$$

そして同様に $(*)$ を表が出るまで繰り返し行う. このとき表がでるまでに裏がでた回数の分布は幾何分布である.

全事象の確率が 1 であることは等比級数の和の公式

$$\sum_{x=0}^{\infty} r^x = \frac{1}{1-r} \equiv f(r), \quad (\text{ただし } |r| < 1)$$

を用いると

$$P(S) = \sum_{x=0}^{\infty} p(x) = p \sum_{x=0}^{\infty} (1-p)^x = \frac{p}{1-(1-p)} = 1$$

直ちに導かれる.

- 平均

級数の公式

$$\sum_{x=1}^{\infty} x r^x = f'(r) r = \frac{r}{(1-r)^2}, \quad (\text{ただし } |r| < 1)$$

を用いると,

$$\mu = \sum_{x=1}^{\infty} x p(x) = \sum_{x=1}^{\infty} x p (1-p)^x = p \times \frac{1-p}{p^2} = \frac{1-p}{p}$$

従って,

$$\text{幾何分布の平均} : \mu = \frac{1-p}{p}$$

- 分散

級数の公式

$$\sum_{x=1}^{\infty} x^2 r^x = f''(r) r^2 + f'(r) r = \frac{r(1+r)}{(1-r)^3}, \quad (\text{ただし } |r| < 1)$$

を用いると,

$$\begin{aligned} V &= \sum_{x=1}^{\infty} x^2 p(x) - \mu^2 = \sum_{x=1}^{\infty} x^2 p (1-p)^x - \left(\frac{1-p}{p} \right)^2 \\ &= p \cdot \frac{1}{p^3} (1-p)(2-p) - \frac{(1-p)^2}{p^2} = \frac{1-p}{p^2} \{2-p-(1-p)\} = \frac{1-p}{p^2}. \end{aligned}$$

$$\text{幾何分布の分散} : V = \frac{1-p}{p^2}$$

$$\text{幾何分布の標準偏差} : \sigma = \frac{\sqrt{1-p}}{p}$$

- 積率母関数

最初の級数の公式を用いると

$$m(t) = \sum_{x=0}^{\infty} e^{tx} p(x) = p \sum_{x=0}^{\infty} e^{tx} (1-p)^x = \frac{p}{1-(1-p)e^t}, \quad t < -\log(1-p).$$

$$\text{幾何分布の積率母関数: } m(t) = \frac{p}{1-(1-p)e^t}, \quad t < -\log(1-p)$$

1.4.5 超幾何分布 (Hyper Geometric Distribution)

S を標本点が有限個つまり

$$S = \{0, 1, 2, 3, \dots, n\}$$

である標本空間であるとする. M を n 以上の整数, k を 0 以上 M 以下の整数とし

$$p(x) = \begin{cases} \frac{\binom{k}{x} \cdot \binom{M-k}{n-x}}{\binom{M}{n}} & x = 0, 1, \dots, n \\ 0 & \text{その他} \end{cases}$$

で定義される離散確率密度で定まる (離散) 確率分布を超幾何分布という.

● 対応する試行

k 個の赤玉, $M - k$ 個の白玉, 計 M 個の玉が箱の中に入っている. この箱の中から n 個の玉を非復元抽出法で取り出したとき赤玉の個数の分布は超幾何分布である.

注意 復元抽出法の場合は $p = \frac{k}{M}$ である二項分布である.

まず超幾何分布の計算のための公式を準備しておく.

$$\binom{a+b}{j} = \sum_{k=0}^j \binom{a}{k} \binom{b}{j-k}$$

証明 二項分布の定理より a, b が自然数のとき

$$(1+x)^a = \sum_{k=0}^a \binom{a}{k} x^k, \quad (1+x)^b = \sum_{k=0}^b \binom{b}{k} x^k, \quad (1+x)^{a+b} = \sum_{k=0}^{a+b} \binom{a+b}{k} x^k$$

が成り立つ. $(1+x)^a(1+x)^b = (1+x)^{a+b}$ に注意すると

$$(1.2) \quad \sum_{k=0}^a \binom{a}{k} x^k \sum_{\ell=0}^b \binom{b}{\ell} x^\ell = \sum_{j=0}^{a+b} \binom{a+b}{j} x^j,$$

が得られる. x^j の係数を比較すると

$$\begin{aligned} \text{右辺の } x^j \text{ の係数} &= \binom{a+b}{j} \\ \text{左辺の } x^j \text{ の係数} &= \sum_{k=0}^j \binom{a}{k} \binom{b}{j-k} \end{aligned}$$

よって上の関係式が得られる.

● 平均

$$\mu = \sum_{x=0}^n xp(x) = \sum_{x=0}^n x \times \frac{\binom{k}{x} \binom{M-k}{n-x}}{\binom{M}{n}} = \frac{nk}{M} \sum_{x=1}^n \frac{\binom{k-1}{x-1} \binom{M-k}{n-x}}{\binom{M-1}{n-1}} = \frac{nk}{M} \sum_{y=0}^{n-1} \frac{\binom{k-1}{y} \binom{(M-1)-(k-1)}{n-1-y}}{\binom{M-1}{n-1}}$$

ここで 3 番目の等式は次の計算から得られた.

$$x \frac{k!}{(k-x)!x!} \frac{n!(M-n)!}{M!} = \frac{k(k-1)!}{(k-x)!(x-1)!} \frac{n(n-1)!(M-n)!}{M(M-1)!} = \frac{kn}{M} \frac{\binom{k-1}{x-1}}{\binom{M-1}{n-1}}.$$

上述で示された公式を用いると

$$\sum_{y=0}^{n-1} \frac{\binom{k-1}{y} \binom{(M-1)-(k-1)}{n-1-y}}{\binom{M-1}{n-1}} = 1$$

となるので

$$\text{超幾何分布の平均} : = \frac{nk}{M}$$

注意 超幾何分布の平均は、 $p = \frac{k}{M}$ としたときの二項分布と同じ。

● 分散

$V = m_2 - \mu^2$ という関係式を用いる。 μ はすでに計算されているので $m_2 - \mu$ を計算してみると

$$\begin{aligned} m_2 - \mu &= \sum_{x=0}^n (x^2 - x)p(x) = \sum_{x=0}^n x(x-1) \frac{\binom{k}{x} \binom{M-k}{n-x}}{\binom{M}{n}} \\ &= \sum_{x=0}^n x(x-1) \frac{k!}{(k-x)!x!} \frac{n!(M-n)!}{M!} \binom{M-k}{n-x} \\ &= \sum_{x=0}^n \frac{k(k-1)(k-2)!}{(k-x)!(x-2)!} \frac{n}{M} \frac{n-1}{M-1} \frac{(n-2)!(M-n)!}{(M-2)!} \binom{M-k}{n-x} \\ &= n(n-1) \frac{k(k-1)}{M(M-1)} \sum_{x=2}^n \frac{\binom{k-2}{x-2} \binom{M-k}{n-x}}{\binom{M-2}{n-2}} = n(n-1) \frac{k(k-1)}{M(M-1)} \end{aligned}$$

最後の等式で上述の公式を用いた。したがって

$$\begin{aligned} V &= n(n-1) \frac{k(k-1)}{M(M-1)} + n \frac{k}{M} - n^2 \frac{k^2}{M^2} \\ &= \frac{M(n^2 - n)k(k-1) + nkM(M-1) - n^2k^2(M-1)}{M^2(M-1)} \\ &= \frac{Mk^2n^2 - Mkn^2 - Mk^2n + Mkn + M^2kn - Mkn - Mk^2n^2 + k^2n^2}{M^2(M-1)} \\ &= \frac{-Mkn^2 - Mk^2n + M^2kn + k^2n^2}{M^2(M-1)} \\ &= \frac{nk(-Mn - Mk + M^2 + nk)}{M^2(M-1)} \\ &= \frac{nk(M-k)(M-n)}{M^2(M-1)} = n \frac{k}{M} \frac{M-k}{M} \frac{M-n}{M-1} \end{aligned}$$

$$\text{超幾何分布の分散} \quad : \quad V = n \frac{k}{M} \frac{M-k}{M} \frac{M-n}{M-1}$$

$$\text{超幾何分布の標準偏差} \quad : \quad \sigma = \sqrt{n \frac{k}{M} \frac{M-k}{M} \frac{M-n}{M-1}}$$

[考察] 1) $p = k/M$, $q = 1 - p$ とおくと分散は $npq \times \frac{M-n}{M-1}$ となる。これは二項分布の分散の $\frac{M-n}{M-1}$ 倍になる。

2) 超幾何分布の積率母関数は超幾何関数 $F(a, b, c, z) = \sum_{x=0}^{\infty} \frac{(a, x)(b, x)}{(c, x)(1, x)} z^x$ (ここで $(a, x) = a(a+1)(a+2)\cdots(a+x-1)$, x は正の整数、 $(a, 0) = 1$) を用いると

$$m(t) = \frac{\binom{M-k}{n}}{\binom{M}{n}} F(-k, -n, M-k-n+1, e^t)$$

と表すことができる。

1.4.6 負の二項分布

負の二項分布は幾何分布の一般化である。 p を 0 以上 1 以下の実数, r を自然数とする。 (離散) 確率分布を負の二項分布の離散確率密度は

$$p(x) = \begin{cases} \binom{r+x-1}{x} p^r q^x & x \in S = \{0, 1, 2, 3, \dots\} \\ 0 & \text{その他} \end{cases}$$

で定義される。

- **対応する試行表**が出る確率が p , 裏が出る確率が $1-p$ であるコインを表が r 回出るまで投げ続ける。このとき裏が出た回数の分布は負の二項分布である。

注意 通常, 組み合わせ $\binom{n}{k}$ は $n \geq k \geq 0$ を満たす整数の組 (n, k) に対して定義されているが, すべての整数の組 (n, k) に対しても

$$\binom{n}{k} = \frac{1}{k!} \left(\frac{d}{dy} \right)^k y^n \Big|_{y=1}, \quad k \geq 0, \quad \binom{n}{k} = 0, \quad k < 0,$$

と定義することにより一般化できる。定義から α が負の整数であるとき

$$\binom{\alpha}{k} = (-1)^k \binom{k-\alpha-1}{k}$$

が成り立つことが分かる。

$$\text{テイラー展開: } F(y) = F(y_0) + \sum_{k=1}^{\infty} \left(\frac{d}{dy} \right)^k F(y_0) \frac{(y-y_0)^k}{k!}$$

を $F(y) = (y+1)^\alpha$ の場合に組み合わせの記号を用いて書き直すと

$$(y+1)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} y^k$$

を得る。この等式を $\alpha = -r, y = -q$ として用いると公式

$$p^{-r} = (1-q)^{-r} = \sum_{x=0}^{\infty} \binom{-r}{x} (-q)^x = \sum_{x=0}^{\infty} \binom{x+r-1}{x} q^x$$

が導かれ, この公式を用いると

$$P(S) = \sum_{x=0}^{\infty} p(x) = p^r \sum_{x=0}^{\infty} \binom{x+r-1}{x} q^x = 1$$

を得る。

● 平均

$$\mu = \sum_{x=0}^{\infty} x p(x) = \sum_{x=0}^{\infty} x \binom{r+x-1}{x} p^r q^x = r \sum_{x=1}^{\infty} \binom{r+x-1}{x-1} p^r q^x = \frac{qr}{p} \sum_{y=0}^{\infty} \binom{r+y}{y} p^{r+1} q^y = \frac{qr}{p}$$

最後の等式は上述の公式から導いた。従って

$$\text{負の二項分布の平均: } \mu = \frac{qr}{p}$$

● 分散

負の二項分布の場合も $V = m_2 - \mu^2$, $m_2 = \mu + \sum_{x=0}^{\infty} x(x-1)p(x)$ という関係を用いる

$$\sum_{x=0}^{\infty} x(x-1)p(x) = \sum_{x=0}^{\infty} x(x-1) \binom{r+x-1}{x} p^r q^x = r(r+1) \sum_{x=2}^{\infty} \binom{r+x-1}{x-2} p^r q^x = \frac{q^2 r(r+1)}{p^2}$$

となるので

$$V = \frac{q^2 r(r+1)}{p^2} + \frac{qr}{p} - \left(\frac{qr}{p}\right)^2 = \frac{qr}{p^2}$$

負の二項分布の分散 : $V = \frac{qr}{p^2}$
 負の二項分布の標準偏差 : $\sigma = \sqrt{\frac{qr}{p^2}} = \frac{\sqrt{qr}}{p}$

● 積率母関数

公式を用いると

$$m(t) = \sum_{x=0}^{\infty} e^{tx} p(x) = p^r \sum_{x=0}^{\infty} \frac{(r+x-1)!}{x!(r-1)!} (qe^t)^x = \left(\frac{p}{1-qe^t}\right)^r$$

が導かれる.

負の二項分布の積率母関数: $m(t) = \left(\frac{p}{1-qe^t}\right)^r$

1.4.7 ポアソン分布 (Poisson 分布) : $Po(\lambda)$

ポアソン分布は標本空間

$$S = \{0, 1, 2, 3, \dots\}$$

上の分布であり, その分布密度は

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x \in S$$

で定義される. ここで $\lambda > 0$ である.

指数関数のテイラー展開

$$e^\lambda = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = 1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \dots$$

を用いると

$$P(S) = \sum_{x=0}^{\infty} p(x) = e^{-\lambda} e^\lambda = 1$$

が導かれる.

● ポアソン分布と関係のある現象

- (i) ある州における 1 週間あたりの交通事故死亡者数
- (ii) 単位時間当たりの放射性物質の放射の数
- (iii) ある素材の単位面積あたりの傷の数

● 平均

$$\mu = \sum_{x=0}^{\infty} xp(x) = e^{-\lambda} \sum_{x=0}^{\infty} x \frac{\lambda^x}{x!} = e^{-\lambda} \lambda \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda$$

ポアソン分布の平均: $\mu = \lambda$

● 分散

$$\sum_{x=0}^{\infty} x(x-1)p(x) = \sum_{x=2}^{\infty} e^{-\lambda} \frac{\lambda^x}{(x-2)!} = \lambda^2 \sum_{\ell=0}^{\infty} e^{-\lambda} \frac{\lambda^{\ell}}{\ell!} = \lambda^2$$

したがって

$$V = \sum_{x=0}^{\infty} x(x-1)p(x) + \mu - \mu^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

ポアソン分布の分散 : $V = \lambda$

ポアソン分布の標準偏差 : $\sigma = \sqrt{\lambda}$

● 積率母関数

$$m(t) = \sum_{x=0}^{\infty} e^{tx} p(x) = \sum_{x=0}^{\infty} e^{tx} e^{-\lambda} \frac{\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(e^t \lambda)^x}{x!} = \exp[e^t \lambda - \lambda]$$

ポアソン分布の積率母関数: $m(t) = \exp[e^t \lambda - \lambda]$

ポアソン分布に関する重要な定理を紹介する.

(i) 長さ $h (> 0)$ の短い時間内である出来事が正確に一つ起こる確率は λh に近似的に等しい. つまり

$$P[\text{長さ } h \text{ の時間の中で一つの出来事が起こる}] = \lambda h + o(h)$$

が成り立つ. ここで

$$\frac{f(h)}{h} \rightarrow 0, h \rightarrow 0 \iff f(h) = o(h)$$

を意味し, $o(h)$ はスモールオーダー h とよぶ. たとえば h^2, h^3 は $o(h)$ である.

(ii) 長さ $h (> 0)$ の短い時間内で 2 つ以上の出来事が起こる確率は $o(h)$, つまり無視できる.

(iii) 重なり合っていない時間の中で起きる出来事の回数は独立.

定理 1.4 上の 3 つの仮定が満たされるならば, 長さ t の時間の中で起こる出来事の回数はパラメータ λt のポアソン分布に従う.

証明 $s > 0$, にたいして $P_n(s) \equiv P[\text{長さ } s \text{ の時間の中で出来事が } n \text{ 回起こる}]$ と定義する. まず $n = 0$ の場合を調べてみる. (iii) より

$$\begin{aligned} P_0(t+h) &= P[\text{区間 } (0, t+h) \text{ で出来事が起こらない}] \\ &= P[\text{区間 } (0, t) \text{ で出来事が起こらない,} \\ &\quad \text{区間 } (t, t+h) \text{ で出来事が起こらない}] \\ &= P_0(t)P_0(h) \end{aligned}$$

を得る. (i),(ii) より

$$P[(t, t+h] \text{ で出来事が起こらない}] \\ = 1 - P[(t, t+h] \text{ で出来事が } 1 \text{ つ以上起こる}] = 1 - \lambda h + o(h)$$

従って $P_0(t+h) = P_0(t)\{1 - \lambda h + o(h)\}$ であるので

$$\lim_{h \rightarrow 0} \frac{P_0(t+h) - P_0(t)}{h} = \lim_{h \rightarrow 0} \frac{P_0(t) - \lambda h P_0(t) - P_0(t)}{h} = -\lambda P_0(t)$$

つまり $\frac{d}{dt} P_0(t) = -\lambda P_0(t)$ となる. 従って

$$\left. \begin{array}{l} \frac{d}{dt} P_0(t) = -\lambda P_0(t) \\ P_0(0) = 1 \end{array} \right\} \Rightarrow \boxed{P_0(t) = e^{-\lambda t}}$$

次に $n = 1$ の場合を調べてみる.

$$\begin{aligned} P_1(t+h) &= P[\text{区間 } (0, t+h] \text{ で出来事が一つ起きる}] \\ &= P[\text{区間 } (0, t] \text{ で一つ起き, } (t, t+h] \text{ で一つも起きない}] \\ &\quad + P[\text{区間 } (0, t] \text{ で一つも起きない, } (t, t+h] \text{ で一つ起きる}] \\ &= P_1(t)P_0(h) + P_0(t)P_1(h) \\ &= P_1(t)e^{-\lambda h} + e^{-\lambda t}P_1(h) \end{aligned}$$

従って $P_1(t+h) - P_1(t) = P_1(t)(e^{-\lambda h} - 1) + e^{-\lambda t}(\lambda h + o(h))$ であるので $e^{-\lambda h} = 1 - \lambda h + o(h)$ を考慮すると

$$\frac{d}{dt} P_1(t) = \lim_{h \rightarrow 0} \frac{P_1(t+h) - P_1(t)}{h} = -\lambda P_1(t) + \lambda e^{-\lambda t}$$

となる. 従って

$$\left. \begin{array}{l} \frac{d}{dt} P_1(t) = -\lambda P_1(t) + \lambda e^{-\lambda t} \\ P_1(0) = 0 \end{array} \right\} \Rightarrow \boxed{P_1(t) = \lambda t e^{-\lambda t}}$$

ここで微分方程式は定数変化法で解いている. つまり $P_1(t) = C_1(t)e^{-\lambda t}$ とおくと

$$\frac{d}{dt} P_1(t) = \frac{dC_1(t)}{dt} e^{-\lambda t} - \lambda C_1(t) e^{-\lambda t} = -\lambda P_1(t) + \frac{dC_1(t)}{dt} e^{-\lambda t}$$

従って

$$\lambda e^{-\lambda t} = \frac{dC_1(t)}{dt} e^{-\lambda t} \Rightarrow \lambda = \frac{dC_1(t)}{dt} \Rightarrow C_1(t) = \lambda t + C'$$

ゆえに

$$P_1(t) = (\lambda t + C')e^{-\lambda t}$$

$P_1(0) = 0$ より $C' = 0$ したがって $P_1(t) = \lambda t e^{-\lambda t}$ が示された.

一般の $n \geq 2$ の場合は

$$\frac{d}{dt} P_n(t) = -\lambda P_n(t) + \lambda P_{n-1}(t)$$

が得られ $P_n(t) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$ が示される.

この定理の系として, 二項分布 $Bi(n, \lambda/n)$ がポアソン分布 $Po(\lambda)$ に収束することが分かる.

1.5 条件付確率と独立性

定義 1.1 (事象の独立性) (i) 事象 A と事象 B が**独立**であるとは

$$P(A \cap B) = P(A)P(B)$$

が成り立つ事である.

(ii) 事象列 A_1, A_2, \dots, A_n が**独立**であるとは B_j を A_j または A_j^c とするとき

$$P\left(\bigcap_{j=1}^n B_j\right) = \prod_{j=1}^n P(B_j)$$

が成り立つ事である.

注意 3つの事象 A_1, A_2, A_3 が独立であるためには $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$ だけでは不十分であり, $P(A_1^c \cap A_2^c \cap A_3^c) = P(A_1^c)P(A_2^c)P(A_3^c)$ および

$$\begin{aligned} P(A_1^c \cap A_2 \cap A_3) &= P(A_1^c)P(A_2)P(A_3), & P(A_1 \cap A_2^c \cap A_3) &= P(A_1)P(A_2^c)P(A_3), \\ P(A_1 \cap A_2 \cap A_3^c) &= P(A_1)P(A_2)P(A_3^c), & P(A_1^c \cap A_2^c \cap A_3) &= P(A_1^c)P(A_2^c)P(A_3), \\ P(A_1^c \cap A_2 \cap A_3^c) &= P(A_1^c)P(A_2)P(A_3^c), & P(A_1 \cap A_2^c \cap A_3^c) &= P(A_1)P(A_2^c)P(A_3^c) \end{aligned}$$

の条件がすべて必要である. これは $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$ のもとでは

$$P(A_1 \cap A_2) = P(A_1)P(A_2), \quad P(A_1 \cap A_3) = P(A_1)P(A_3), \quad P(A_2 \cap A_3) = P(A_2)P(A_3),$$

という条件と必要十分である. $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$ をみたすが $P(A_1 \cap A_2) \neq P(A_1)P(A_2)$ となる例は, 簡単に作ることができるので, 各自で調べよ. 次の例は, $P(A_1 \cap A_2) = P(A_1)P(A_2)$, $P(A_1 \cap A_3) = P(A_1)P(A_3)$, $P(A_2 \cap A_3) = P(A_2)P(A_3)$ であるが, $P(A_1 \cap A_2 \cap A_3) \neq P(A_1)P(A_2)P(A_3)$ となる例である.

例 $\Omega = \{1, 2, 3, 4\}$, $P(\{i\}) = 1/4$ (4点集合上の離散一様分布), $A_1 = \{1, 2\}$, $A_2 = \{2, 3\}$, $A_3 = \{1, 3\}$ とする. A_1, A_2, A_3 は各々の2つは独立であるが3つ全体では独立ではない.

定義 1.2 (条件付確率) 事象 B が起こるという条件の下での事象 A が起こる条件付確率を

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

で定義する. ただし $P(B) = 0$ の時は定義しない.

$P(B) \neq 0$ の時 A と B が独立 $\iff P(A|B) = P(A)$ であることが分かる. また,

$$P(A \cap B) = P(A|B)P(B)$$

が成り立つ. この式を**乗法の公式**という.

つぎに確率変数列 $X_i, i = 1, 2, \dots, n$ の独立性を定義する.

定義 1.3 (確率変数の独立性) 確率変数列 $X_i, i = 1, 2, \dots, n$ が**独立**であるとは, 任意の $a_i \in X_i(\Omega)$, $i = 1, 2, \dots, n$ に対して

$$P(X_i = a_i, i = 1, 2, \dots, n) = \prod_{i=1}^n P(X_i = a_i)$$

が成り立つ事である.

注意 事象 A にたいして

$$\mathbf{1}_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

で定義される関数を指示関数 (indicator function) という. 事象列 $A_i, i = 1, 2, \dots, n$ が**独立**であることと確率変数列 $\mathbf{1}_{A_i}, i = 1, 2, \dots, n$ が**独立**であることは必要十分である.

定理 1.5 (期待値の乗法定理) 確率変数 X, Y が独立, f, g を $E[|f(X)|] < \infty, E[|g(Y)|] < \infty$, をみたす実数値関数とする. このとき $E[|f(X)g(Y)|] < \infty$ かつ

$$E[f(X)g(Y)] = E[f(X)]E[g(Y)]$$

である.

証明. 独立性と和の順序交換より

$$\begin{aligned} E[|f(X)g(Y)|] &= \sum_{a \in X(\Omega)} \sum_{b \in Y(\Omega)} |f(a)g(b)|P(X = a, Y = b) \\ &= \sum_{a \in X(\Omega)} \sum_{b \in Y(\Omega)} |f(a)g(b)|P(X = a)P(Y = b) \\ &= \left(\sum_{a \in X(\Omega)} |f(a)|P(X = a) \right) \left(\sum_{b \in Y(\Omega)} |g(b)|P(Y = b) \right) \\ &= E[|f(X)|]E[|g(Y)|] < \infty. \end{aligned}$$

をえる. 上と全くおなじ計算により $E[f(X)g(Y)] = E[f(X)]E[g(Y)]$ も示される.

系 1.1 (積率母関数の乗法定理) 確率変数 X, Y が独立であるとする. X の積率母関数を $m(t, X)$, Y の積率母関数を $m(t, Y)$ とおいたとき, 2つの積率母関数が定義されている t に対して $m(t, X+Y) = m(t, X)m(t, Y)$ が成り立つ.

演習 1.2. 確率変数 X, Y の積率母関数 $m(t, X), m(t, Y)$ が等しいときそれらの分布は等しいという性質をもちいて次のことが示せ.

- (i) $X_i, i = 1, 2, \dots, k$ が独立確率変数列で X_i の分布が2項分布 $B(n_i, p)$ のとき, $S = \sum_{i=1}^k X_i$ の分布は2項分布 $B(\sum_{i=1}^k n_i, p)$ である.
- (ii) $X_i, i = 1, 2, \dots, k$ が独立確率変数列で X_i の分布がポアソン分布 $Po(\lambda_i)$ のとき $S = \sum_{i=1}^k X_i$ の分布はポアソン分布 $Po(\sum_{i=1}^k \lambda_i)$ である.

系 1.2 (分散の加法性) 確率変数 X, Y が独立であるとする. $Var(X) < \infty, Var(Y) < \infty$ とき

$$Var(X + Y) = Var(X) + Var(Y)$$

が成り立つ.

証明. $f(x) = x - E[X], g(y) = y - E[Y]$ として乗法定理を用いると

$$E[(X - E[X])(Y - E[Y])] = E[X - E[X]]E[Y - E[Y]] = 0$$

となり

$$\begin{aligned} \text{Var}(X+Y) &= E[(X - E[X])^2] + E[(Y - E[Y])^2] + 2E[(X - E[X])(Y - E[Y])] \\ &= \text{Var}(X) + \text{Var}(Y) \end{aligned}$$

をえる.

たたみ込み (convolution) X, Y を確率変数とし, それぞれの確率密度関数を p_X, p_Y とおく.

$$\mathcal{R} = \{na + mb : n, m \in \mathbb{Z}, a \in X(\Omega), b \in Y(\Omega)\}$$

とおく.

$$p_X * p_Y(x) \equiv \sum_{y \in \mathcal{R}} p_X(x-y)p_Y(y), \quad x \in \mathcal{R}$$

で定義された確率分布は p_X と p_Y のたたみ込み (convolution) といい, $p_X * p_Y$ とかく.

X, Y を独立であるとし, $X+Y$ の密度関数を p_{X+Y} とおくと, $x \in \mathcal{R}$ に対して

$$\begin{aligned} p_{X+Y}(x) = P(X+Y=x) &= \sum_{y \in \mathcal{R}} P(X=x-y, Y=y) \\ &= \sum_{y \in \mathcal{R}} P(X=x-y)P(Y=y) \\ &= \sum_{y \in \mathcal{R}} p_X(x-y)p_Y(y) = p_X * p_Y(x) \end{aligned}$$

が成り立つ. したがって独立確率変数の和の確率密度関数は, それぞれの確率密度関数のたたみ込みになることがわかる.

2 一般の確率空間

2.1 確率空間の定義と基本的性質

Ω を集合とし, Ω のべき集合 (power set) を 2^Ω と書く: $2^\Omega = \{A | A \subset \Omega\}$.

定義 2.1 $\mathcal{F} \subset 2^\Omega$ が σ -加法族であるとは次の 3 条件が満たされることをいう.

- (a) $\Omega, \emptyset \in \mathcal{F}$,
- (b) $A \in \mathcal{F}$ ならば $A^c \in \mathcal{F}$,
- (c) $A_n \in \mathcal{F}, n = 1, 2, \dots$, ならば $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Ω を集合, \mathcal{F} を σ 加法族とするとき, 組 (Ω, \mathcal{F}) を **可測空間** (measurable space) とよぶ.

注意 $\mathcal{A} \subset 2^\Omega$ が上で述べた (a), (b) に加えて次の (c') がみたされるとき, \mathcal{A} は**加法族** であるという.

- (c') $A_1, A_2 \in \mathcal{A}$ ならば $A_1 \cup A_2 \in \mathcal{A}$.

- 例** (i) $\mathcal{F} = \{\emptyset, \Omega\}$ は σ -加法族である.
(ii) $\mathcal{F} = 2^\Omega$ は σ -加法族である.
(iii) $A \subset \Omega$ であるとき, $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$ は σ -加法族である.

- 演習 2.1.** (i) \mathcal{F}, \mathcal{G} が σ -加法族であるとき, $\mathcal{F} \cap \mathcal{G}$ は σ -加法族であることを示せ.
(ii) \mathcal{F}, \mathcal{G} が σ -加法族であるが, $\mathcal{F} \cup \mathcal{G}$ は σ -加法族ではない例を与えよ.
(iii) $\mathcal{A} \subset 2^\Omega$ とするとき,

$$\sigma(\mathcal{A}) \equiv \{A \subset \Omega; A \subset \mathcal{F} \text{ なる任意の } \sigma\text{-加法族 } \mathcal{F} \text{ にたいし } A \in \mathcal{F}\}$$

は σ -加法族であることを示せ. ($\sigma(\mathcal{A})$ は \mathcal{A} を含む最小の σ -加法族であり, \mathcal{A} により生成される σ -加法族という.)

定義 2.2 Ω が位相空間であり \mathcal{O} をその開部分集合全体とする. $\sigma(\mathcal{O})$ を $\mathcal{B}(\Omega)$ と表し, Ω の**ボレル σ -加法族** という.

- 演習 2.2.** (i) Ω を位相空間とし, \mathcal{C} をその閉部分集合全体とする. $\sigma(\mathcal{C}) = \mathcal{B}(\Omega)$ となることを示せ.
(ii) $\Omega = \mathbb{R}^2$ とし,

$$\mathcal{A} = \{(a, b) \times (c, d) \subset \mathbb{R}^2; -\infty < a < b < \infty, -\infty < c < d < \infty\}$$

とおく. $\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R}^2)$ となることを示せ.

次に集合族が σ -加法族になるための条件について纏めておく.

定義 2.3 (単調族) Ω の部分集合族 \mathcal{M} が, 増大列, 減少列にたいして閉じているとき, 単調族という. すなわち, $A_n \in \mathcal{M}, n \in \mathbb{N}$ に対し

(M1) $A_n \subset A_{n+1}, n = 1, 2, \dots$ ならば $\bigcup_{n=1}^{\infty} A_n \in \mathcal{M}$.

(M2) $A_n \supset A_{n+1}, n = 1, 2, \dots$ ならば $\bigcap_{n=1}^{\infty} A_n \in \mathcal{M}$.

\mathcal{C} を Ω の部分集合族とする. 演習 2.1 (iii) で用いた議論と同様な議論を用いると \mathcal{C} を含む最少の単調族が唯一存在することが分かる. これを $\mathcal{M}(\mathcal{C})$ と書くことにする. さらに次の定理が導かれる.

定理 2.1 \mathcal{A} を加法族とする. \mathcal{A} が σ -加法族となるための必要十分条件は, \mathcal{A} が単調族になることである. そして $\sigma(\mathcal{A}) = \mathcal{M}(\mathcal{A})$ が成り立つ.

演習 2.3. 定理 th:monotone を証明せよ. (西尾真喜子 [7]31 ページ、二章、定理 4, 5 の証明を参照せよ.)

定義 2.4 (Dynkin 族) Ω の部分集合族 \mathcal{D} がつぎの 3 条件をみたすとき Dynkin 族という:

- (D1) $\Omega \in \mathcal{D}$.
- (D2) $A_1, A_2 \in \mathcal{D}$, $A_2 \subset A_1$ ならば $A_1 \setminus A_2 \in \mathcal{D}$.
- (D3) $A_n \in \mathcal{D}$ が増大するとき $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.

定義からただちに, σ -加法族は Dynkin 族であり, Dynkin 族は単調族であることが分かる. \mathcal{C} を Ω の部分集合族とする. 演習 2.1 (iii) で用いた議論と同様な議論を用いると \mathcal{C} を含む最少の Dynkin 族が唯一存在することが分かる. これを $\mathcal{D}(\mathcal{C})$ と書くことにする.

定理 2.2 \mathcal{C} が条件「 $A_1, A_2 \in \mathcal{C}$ ならば $A_1 \cap A_2 \in \mathcal{C}$ 」をみたすとき $\sigma(\mathcal{C}) = \mathcal{D}(\mathcal{C})$ が成り立つ.

証明 σ -加法族は, Dynkin 族であるので $\mathcal{D}(\mathcal{C}) \subset \sigma(\mathcal{C})$ であることは, 明らか. $A \in \mathcal{C}$ に対して $\mathcal{C}_A = \{B \in \mathcal{D}(\mathcal{C}) : A \cap B \in \mathcal{D}(\mathcal{C})\}$ とおけば \mathcal{C}_A は \mathcal{C} を含む Dynkin 族である. よって $\mathcal{C}_A = \mathcal{D}(\mathcal{C})$ である. $B \in \mathcal{C}_A$ に対して $\mathcal{C}_B = \{A \in \mathcal{D}(\mathcal{C}) : A \cap B \in \mathcal{D}(\mathcal{C})\}$ とおけば \mathcal{C}_B は \mathcal{C} を含む Dynkin 族である. よって $\mathcal{C}_B = \mathcal{D}(\mathcal{C})$ である. したがって $A, B \in \mathcal{D}(\mathcal{C})$ であれば $A \cap B \in \mathcal{D}(\mathcal{C})$, ゆえに $\mathcal{D}(\mathcal{C})$ は加法族である. Dynkin 族は単調族であるので, 定理 2.1 より σ -加法族であることが分かる. ゆえに $\mathcal{D}(\mathcal{C}) \supset \sigma(\mathcal{C})$ となる. したがって, 定理が示された.

定義 2.5 σ -加法族 \mathcal{F} 上の写像 P が可測空間 (Ω, \mathcal{F}) 上の確率測度 (Probability measure) であるとは次の条件がみたされることをいう.

- (a) $P(\Omega) = 1$,
- (b) $0 \leq P(A) \leq 1$, $A \in \mathcal{F}$,
- (c) (σ -加法性) $A_n \in \mathcal{F}, n = 1, 2, \dots$, かつ $A_n \cap A_m = \emptyset, n \neq m$ ならば

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

Ω を集合, \mathcal{F} を σ 加法族, P を確率とすると, 三つの組 (Ω, \mathcal{F}, P) を**確率空間** (Probability space) とよぶ. 以降, (Ω, \mathcal{F}, P) は確率空間を表すものとする.

演習 2.4. (i) $P(\emptyset) = 0$, が常に成り立つことを示せ.
(ii) $A_n \in \mathcal{F}, A_n \subset A_{n+1}, n = 1, 2, \dots$, ならば $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n)$ であることを示せ.
(iii) $A_n \in \mathcal{F}, A_n \supset A_{n+1}, n = 1, 2, \dots$, ならば $P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} P(A_n)$ であることを示せ.

定義 2.6 命題 $A(\omega)$ が与えられていて $\omega \in \Omega$ を決めるごとに命題が正しいか誤りが判定できるとする. $T \equiv \{A(\omega) \text{ が正しい}\}$ が可測であり, つまり $T \in \mathcal{F}$ であり, さらに $P(T) = 1$ となるとき, 命題 A は**ほとんど確実になりたつ** といひ, $A, a.s.$ と表す.

例 $\Omega = [0, 1], \mathcal{F} = \mathcal{B}([0, 1]), P$ をルベーグ測度 (Lebesgue measure) とする. $A(\omega)$ を「 ω は無理数である」という命題であるとする. $P(T) = 1$ である. したがって, ほとんど確実に「 ω は無理数である」.

2.2 期待値

定義 2.7 (確率変数, *random variable*) 可測空間 (Ω, \mathcal{F}) 上で定義された実数値関数 X が **確率変数** であるとは, 任意の $B \in \mathcal{B}(\mathbb{R})$ にたいして,

$$X^{-1}(B) \equiv \{\omega \in \Omega; X(\omega) \in B\} \in \mathcal{F}$$

が成り立つことをいう.

演習 2.5. (i) 任意の $a \in \mathbb{R}$ にたいして $X^{-1}((-\infty, a]) \in \mathcal{F}$ を満たせば X は確率変数であることを示せ.

(ii) X, Y を確率変数, $a, b \in \mathbb{R}$ とする. $aX + bY, XY$ もまた確率変数であることを示せ.

(iii) $X_n, n = 1, 2, \dots$, を確率変数とすれば, $\sup_{n \in \mathbb{N}} X_n, \inf_{n \in \mathbb{N}} X_n, \limsup_{n \in \mathbb{N}} X_n, \liminf_{n \in \mathbb{N}} X_n$ はすべて確率変数であることを示せ.

確率変数の期待値を定義するために階段関数 (step function) (または単純関数 (simple function)) と呼ばれる関数を定義する. 確率変数 X が階段関数であるとは, 適当な $n \in \mathbb{N}, a_1, a_2, \dots, a_n \in \mathbb{R}, A_1, A_2, \dots, A_n \in \mathcal{F}$ を選んで

$$X(\omega) = \sum_{k=1}^n a_k \mathbf{1}_{A_k}(\omega), \quad \omega \in \Omega$$

と表すことができることである. 階段関数全体の集合を \mathbb{SF} と書くことにする.

確率変数の期待値 $E[X]$ は次のように定義される.

(i) $X \in \mathbb{SF}$ に対して

$$E[X] \equiv \sum_{i=1}^n a_i P(A_i)$$

で定義する.

(ii) 非負確率変数 X に対して

$$E[X] \equiv \sup\{E[Y]; Y \in \mathbb{SF}, 0 \leq Y \leq X\}$$

で定義する.

(iii) 一般の確率変数 X に対しては, 2つの非負確率変数

$$X^+(\omega) \equiv \max\{0, X(\omega)\}, \quad X^-(\omega) \equiv \max\{0, -X(\omega)\}$$

を用いる. そして $E[X^+] < \infty$ または $E[X^-] < \infty$ のとき

$$E[X] = E[X^+] - E[X^-]$$

で定義する. したがって, $E[X^+] < \infty$ かつ $E[X^-] = \infty$ のときは $E[X] = -\infty$ となり, $E[X^+] = \infty$ かつ $E[X^-] < \infty$ のときは $E[X] = \infty$ となる. $E[X^+] = \infty$ かつ $E[X^-] = \infty$ のときは期待値は存在しない.

$E[X^+] < \infty$ かつ $E[X^-] < \infty$ のとき, X は**可積分である** という. $A \in \mathcal{F}$ であるとき $E[\mathbf{1}_A X]$ を $E[X, A]$ と記し, A 上の期待値と呼ぶ.

定理 2.3 X, Y を可積分確率変数とする.

(i) (線形性) 任意の $a, b \in \mathbb{R}$ に対して $aX + bY$ も可積分確率変数であり

$$E[aX + bY] = aE[X] + bE[Y]$$

がなりたつ.

- (ii) (正值性) $X \geq Y$ a.s. を満たせば $E[X] \geq E[Y]$ がなりたつ.
- (iii) $X \geq Y$ a.s. かつ $E[X] = E[Y]$ であれば $X = Y$ a.s. がなりたつ.

証明

- (Step 1) $X, Y \in \mathbb{SF}$ のとき, (i),(ii),(iii) は明らか. (各自確認)
- (Step 2) $\{X_n\}_{n \in \mathbb{N}}, \{Y_n\}_{n \in \mathbb{N}} \subset \mathbb{SF}$ が単調増加列であり, $\lim_{n \rightarrow \infty} X_n = \lim_{n \rightarrow \infty} Y_n$ であるとき $\lim_{n \rightarrow \infty} E[X_n] = \lim_{n \rightarrow \infty} E[Y_n]$ がなりたつ.

\therefore) $m \in \mathbb{N}$ を固定しする. $Y_m \in \mathbb{SF}$ であるので $Y_m = \sum_{k=1}^M a_k \mathbf{1}_{A_k}$ と互いに素である事象 $A_k, k = 1, 2, \dots, M$ で表すことができる. 任意の $\varepsilon > 0$ を固定して

$$A_{k,n} = \{\omega \in A_k; X_n \geq (1 - \varepsilon)a_k\}$$

とおく. $A_{k,n}$ は n について単調増加であり $\bigcup_{n=1}^{\infty} A_{k,n} = A_k$ であることが $\lim_{n \rightarrow \infty} X_n \geq Y_m$ よりわかる.

さらに定義より $X_n \geq (1 - \varepsilon) \sum_{k=1}^M a_k \mathbf{1}_{A_{k,n}}$ であることに注意すると

$$\begin{aligned} \lim_{n \rightarrow \infty} E[X_n] &\geq (1 - \varepsilon) \liminf_{n \rightarrow \infty} \sum_{k=1}^M a_k P(A_{k,n}) \\ &= (1 - \varepsilon) \sum_{k=1}^M a_k P(A_k) = (1 - \varepsilon) E[Y_m] \end{aligned}$$

が得られる. ここで最初の等号は演習 2.4 (ii) を使って得られた. 任意の $\varepsilon > 0$ について上式が成立するので $\lim_{n \rightarrow \infty} E[X_n] \geq E[Y_m]$ となる. 従って $\lim_{n \rightarrow \infty} E[X_n] \geq \lim_{m \rightarrow \infty} E[Y_m]$ が導かれる. 逆の不等式も全く同様にして導かれる.

- (Step 3) $\{X_n\}_{n \in \mathbb{N}} \subset \mathbb{SF}$ が非負単調増加列であり, $X = \lim_{n \rightarrow \infty} X_n$ であるとき $E[X] = \lim_{n \rightarrow \infty} E[X_n]$ がなりたつ.

\therefore) 期待値の定義より

$$\lim_{n \rightarrow \infty} E[X_n] \leq E[X]$$

であること, また $Y_n \leq X, E[X] = \lim_{n \rightarrow \infty} E[Y_n]$ を満たす非負階段関数列 $\{Y_n\}_{n \in \mathbb{N}} \subset \mathbb{SF}$ が存在することが分かる. 非負単調増加階段関数列 $\{Z_n\}_{n \in \mathbb{N}}$ を $Z_n = \max\{X_n, Y_1, \dots, Y_n\}$ で定義すると $\lim_{n \rightarrow \infty} Z_n = X$ が成り立つ. Step 1) と Step 2) を用いると

$$\lim_{n \rightarrow \infty} E[X_n] = \lim_{n \rightarrow \infty} E[Z_n] \geq \lim_{n \rightarrow \infty} E[Y_n] = E[X]$$

が導かれる.

- (Step 4) $X, Y \geq 0, a, b \geq 0$ のとき, (i),(ii) が成り立つ.

\therefore) 任意の確率変数 Z に対して可測事象列 $A_{n,k}(Z)$ を

$$A_{n,k}(Z) = \{\omega; k2^{-n} \leq Z(\omega) < (k+1)2^{-n}\}$$

で定義し

$$X_n = \sum_{k=0}^{n2^n} k2^{-n} \mathbf{1}_{A_{n,k}(X)}, \quad Y_n = \sum_{k=0}^{n2^n} k2^{-n} \mathbf{1}_{A_{n,k}(Y)}$$

とおく. $\{X_n\}_{n \in \mathbb{N}}, \{Y_n\}_{n \in \mathbb{N}}$ は非負単調増加階段関数列であり $\lim X_n = X, \lim Y_n = Y$ が成り立つ. $a, b \geq 0$ であるとき $\{aX_n + bY_n\}_{n \in \mathbb{N}}$ も非負単調増加階段関数列であり $\lim_{n \rightarrow \infty} (aX_n + bY_n) = aX + bY$ が成り立つことは明らかであるので Step 1) と Step 3) を用いれば (i) は導かれる.

$X \geq Y$ ならば $X_n \geq Y_n$ であるから Step 1) と Step 3) を用いれば (ii) は導かれる.

- (Step 5) 一般の可積分確率変数 $X, Y, a, b \in \mathbb{R}$ に対して (i), (ii) が成り立つ.

\therefore $a \geq 0$ のときは $(aX)^\pm = aX^\pm$ であることより Step 4) を用いると

$$E[aX] = E[(aX)^+] - E[(aX)^-] = aE[X^+] - aE[X^-] = aE[X]$$

となり, $a < 0$ のときは $(aX)^\pm = -aX^\mp$ であることより Step 4) を用いると

$$E[aX] = E[(aX)^+] - E[(aX)^-] = -aE[X^-] + aE[X^+] = aE[X]$$

を得るので, 任意の $a \in (-\infty, \infty)$ と可積分確率変数 X に対して $E[aX] = aE[X]$ が成り立つことがわかる. したがって (i) を示すためには任意の可積分確率変数 X, Y に対して

$$(2.1) \quad E[X - Y] = E[X] - E[Y]$$

を示せばよい.

$$(2.2) \quad \begin{aligned} E[X - Y] &= E[(X - Y)^+] - E[(X - Y)^-] \\ &= E[(X - Y)\mathbf{1}_{\{X > Y\}}] + E[(X - Y)\mathbf{1}_{\{Y > X\}}] \end{aligned}$$

となるが, Step 4) を用いると

$$\begin{aligned} E[(X - Y)\mathbf{1}_{\{X > Y > 0\}}] + E[Y\mathbf{1}_{\{X > Y > 0\}}] &= E[X\mathbf{1}_{\{X > Y > 0\}}] \\ E[(X - Y)\mathbf{1}_{\{X > Y, Y < 0\}}] &= E[X\mathbf{1}_{\{X > Y, Y < 0\}}] - E[Y\mathbf{1}_{\{X > Y, Y < 0\}}] \end{aligned}$$

を得る. この2つの等式を合わせると

$$E[(X - Y)\mathbf{1}_{\{X > Y\}}] = E[X\mathbf{1}_{\{X > Y\}}] - E[Y\mathbf{1}_{\{X > Y\}}]$$

が導かれる. 同様にして

$$E[(X - Y)\mathbf{1}_{\{X \leq Y\}}] = E[X\mathbf{1}_{\{X \leq Y\}}] - E[Y\mathbf{1}_{\{X \leq Y\}}]$$

が導かれ, 上の2つの式を (2.2) に代入すると (2.1) を得る.

線形性より「 $X \geq 0$ a.s. であれば $E[X] \geq 0$ 」であることを示せば十分であるが, これは $X \geq 0$ a.s. という条件があるので Step 4) で示されている.

- (Step 6) 一般の可積分確率変数 X, Y に対して (iii) が成り立つ.

\therefore 線形性より「 $X \geq 0$ a.s. かつ $E[X] = 0$ であれば $X = 0$ a.s.」であることを示せば十分である. $E_n = \{\omega; X(\omega) \geq n^{-1}\}$ とおく. $X \geq n^{-1}\mathbf{1}_{E_n}$ であるから

$$0 = E[X] \geq n^{-1}P(E_n) \geq 0.$$

よって $P(E_n) = 0$. $n \rightarrow \infty$ とすれば, $P(X > 0) = 0$. \square

定義 2.8 X, Y を X^2, Y^2 が可積分な確率変数とする. X の分散 $Var(X)$, X, Y の共分散 $Cov(X, Y)$ を次で定義する.

$$Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2], \quad Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

演習 2.6.

- (i) $Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2, Cov(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$ となることを示せ.
- (ii) $Var(X) \geq 0$ であり, 等号成立は $\mu \in \mathbb{R}$ が存在し $X = \mu, a.s.$ となるときに限ることを証明せよ.
- (iii) X_1, \dots, X_n を 2 乗可積分確率変数とし, $C_{ij} = Cov(X_i, X_j)$ とおく. 行列 $C = (C_{ij})_{1 \leq i, j \leq n}$ は対称非負定値となることを証明せよ.

2.3 期待値に関する不等式

1. シュバルツ (Schwarz) の不等式: X^2, Y^2 が可積分となるような確率変数 X, Y に対し, XY もまた可積分であり, 次式が成り立つ.

$$|\mathbb{E}[XY]| \leq (\mathbb{E}[X^2])^{1/2}(\mathbb{E}[Y^2])^{1/2}.$$

証明 $|XY| \leq (|X|^2 + |Y|^2)/2$ より, XY は可積分である. 期待値の正值性より,

$$0 \leq \mathbb{E}[(X + tY)^2] = (\mathbb{E}[Y^2])t^2 + 2(\mathbb{E}[XY])t + \mathbb{E}[X^2], \forall t \in \mathbb{R}$$

となるから, 2 次方程式の判別式から望む不等式を得る. \square

2. ヘルダー (Hölder) の不等式: p, q を $\frac{1}{p} + \frac{1}{q} = 1$ をみたす 1 より大きい実数とする. X が p 乗可積分, Y が q 乗可積分であるとき

$$(2.3) \quad \mathbb{E}[|XY|] \leq \mathbb{E}[|X|^p]^{1/p} \mathbb{E}[|Y|^q]^{1/q}$$

が成り立つ. 特に $p = q = 2$ のときがシュバルツの不等式である.

証明 任意の正の実数 a, b に対して

$$(2.4) \quad \frac{a^p}{p} + \frac{b^q}{q} \geq ab$$

が成り立つ (ヤングの不等式) ので, これを $a = |X|/\mathbb{E}[|X|^p]^{1/p}, b = |Y|/\mathbb{E}[|Y|^q]^{1/q}$ として使えば

$$\frac{1}{p} \frac{|X|^p}{\mathbb{E}[|X|^p]} + \frac{1}{q} \frac{|Y|^q}{\mathbb{E}[|Y|^q]} \geq \frac{|XY|}{\mathbb{E}[|X|^p]^{1/p} \mathbb{E}[|Y|^q]^{1/q}}$$

となる. 両辺の期待値をとれば, 求める不等式が直ちに得られる. \square

3. ミンコフスキー (Minkowski) の不等式: p を 1 以上の実数とする.

$$(2.5) \quad \mathbb{E}[|X + Y|^p]^{1/p} \leq \mathbb{E}[|X|^p]^{1/p} + \mathbb{E}[|Y|^p]^{1/p}$$

が成り立つ.

演習 2.7.

- (i) ヤングの不等式 (2.4) を証明せよ.
- (ii) ヘルダーの不等式を用いて, $0 < p < p'$ のとき $\mathbb{E}[|X|^p]^{1/p} \leq \mathbb{E}[|X|^{p'}]^{1/p'}$ が成り立つことを示せ.
この不等式は, ある次数の積率が存在すれば, それより低次の成立が存在することをしめしている.
- (iii) ヘルダーの不等式を用いて, ミンコフスキーの不等式を証明せよ.

4. チェビシェフ (Tchebichev) の不等式: φ を \mathbb{R} 上の非負偶関数で $(0, \infty)$ 上で正かつ増大とする. $\varphi(X)$ が可積分のとき, 任意の $x > 0$ に対して

$$(2.6) \quad P(|X| \geq x) \leq \frac{E[\varphi(X)]}{\varphi(x)}$$

が成り立つ.

証明

$$E[\varphi(X)] = E[\varphi(X) : |X| \geq x] + E[\varphi(X) : |X| < x] \geq \varphi(x)P(|X| \geq x)$$

より, 直ちに導かれる. \square

5. エンセン (Jensen) の不等式: φ を \mathbb{R} 上の凸関数で X と $\varphi(X)$ が可積分のとき,

$$(2.7) \quad \varphi(E[X]) \leq E[\varphi(X)]$$

が成り立つ.

証明 $m = E[X]$ とおく. $\varphi(X)$ は凸関数であるので, ある $c \in \mathbb{R}$ で

$$\varphi(x) \geq \varphi(m) + c(x - m)$$

をみたすものが存在する. 従って

$$E[\varphi(X)] \geq \varphi(m) + cE[X - m] = \varphi(m)$$

が導かれる. \square

2.4 一様可積分性と収束定理

定理 2.4 (収束定理) $X, X_n, n \in \mathbb{N}$ は確率変数とする.

(i) (単調収束定理) $0 \leq X_n \leq X_{n+1}$, $\lim_{n \rightarrow \infty} X_n = X$ a.s. であれば

$$\lim_{n \rightarrow \infty} E[X_n] = E[X].$$

(ii) (Fatou の補題) $X_n \geq 0$ a.s. であれば

$$E[\liminf_{n \rightarrow \infty} X_n] \leq \liminf_{n \rightarrow \infty} E[X_n].$$

(iii) (Lebesgue の収束定理) Y が非負可積分確率変数で, $|X_n| \leq Y$ a.s. であれば

$$X_n \rightarrow X \text{ a.s.} \quad \text{ならば} \quad \lim_{n \rightarrow \infty} E[X_n] = E[X].$$

(iv) (Lebesgue の有界収束定理) $M > 0$ が存在して $|X_n| \leq M$ a.s. であれば

$$X_n \rightarrow X \text{ a.s.} \quad \text{ならば} \quad \lim_{n \rightarrow \infty} E[X_n] = E[X].$$

証明. (i) $X_0 = 0$ とする. $X_k - X_{k-1} \geq 0$ より $Y_{k,n} \in \mathbb{SF}$ で非負単調増加であり $\lim_{n \rightarrow \infty} Y_{k,n} = X_k - X_{k-1}$

となるものが存在する. $Z_n = \sum_{k=1}^n Y_{k,n}$ とおくと $\{Z_n\}_{n \in \mathbb{N}}$ は非負単調増加階段関数列であり $\lim_{n \rightarrow \infty} Z_n = X$

であることがわかる. したがって前定理の証明での Step 3), および (ii) を用いると

$$E[X] = \lim_{n \rightarrow \infty} E[Z_n] \leq \lim_{n \rightarrow \infty} E[X_n] \leq E[X]$$

となり (i) をえる.

(ii) $Y_n = \inf_{m \geq n} X_m$ とおくと Y_n は単調増加であるので (i) を用いれば

$$E[\liminf_{n \rightarrow \infty} X_n] = \lim_{n \rightarrow \infty} E[Y_n]$$

が成り立つ. 一方 $Y_n \leq X_n$ より $E[Y_n] \leq E[X_n]$. したがって $\lim_{n \rightarrow \infty} E[Y_n] \leq \liminf_{n \rightarrow \infty} E[X_n]$ となるので, これを上等の式に代入すると (ii) をえる.

(iii) $X_n + Y, -X_n + Y$ が非負確率変数であり, それぞれ $X + Y, -X + Y$ に概収束することより Fatou の補題と線形性より

$$\begin{aligned} E[X] &= E[X + Y] - E[Y] \leq \liminf_{n \rightarrow \infty} E[X_n + Y] - E[Y] = \liminf_{n \rightarrow \infty} E[X_n], \\ -E[X] &= E[-X + Y] - E[Y] \leq \liminf_{n \rightarrow \infty} E[-X_n + Y] - E[Y] = -\limsup_{n \rightarrow \infty} E[X_n] \end{aligned}$$

となり, 合わせると (iii) をえる.

(iv) 定数確率変数は可積分であるので (iv) は (iii) の特別な場合である. \square

定義 2.9 確率変数列 $X_n, n \in \mathbb{N}$, が一様可積分であるとは

$$(2.8) \quad \lim_{a \rightarrow \infty} \sup_{n \in \mathbb{N}} E[|X_n|, |X_n| \geq a] = 0,$$

が成り立つことである.

一様可積分確率変数列の例

1) ある可積分確率変数 Y に対して $\sup_{n \in \mathbb{N}} |X_n| \leq Y$, a.s. が成り立つ時, $X_n, n \in \mathbb{N}$ は一様可積分である.

$$(\because) \sup_{n \in \mathbb{N}} E[|X_n|, |X_n| \geq a] \leq \sup_{n \in \mathbb{N}} E[|X_n|, Y \geq a] \leq E[Y, Y \geq a] \rightarrow 0, \quad a \rightarrow \infty.$$

2) ある $p > 1$ に対して $\sup_{n \in \mathbb{N}} E[|X_n|^p] \leq K < \infty$ が成り立つ時, $X_n, n \in \mathbb{N}$ は一様可積分である.

(\because) チェビセフの不等式より

$$P(|X_n| \geq a) \leq a^{-p} E[|X_n|^p] \leq K a^{-p}$$

となる. ヘルダーの不等式を用いると ($1/p + 1/q = 1$)

$$\begin{aligned} \sup_n E[|X_n|, |X_n| \geq a] &\leq \sup_n \left\{ E[|X_n|^p]^{1/p} E[\mathbf{1}(|X_n| \geq a)]^{1/q} \right\} \\ &\leq K^{1/p} (K a^{-p})^{1/q} \leq K a^{-p/q} \rightarrow 0, \quad a \rightarrow \infty. \end{aligned}$$

確率変数列 $X_n, n \in \mathbb{N}$, が一様可積分であるとき, 次の性質がなりたつ.

$$(U1) \quad \sup_{n \in \mathbb{N}} E[|X_n|] \equiv K < \infty$$

(U2) 任意の $\varepsilon > 0$ にたいして $\delta > 0$ が存在して

$$P(\Lambda) < \delta, \quad \Lambda \in \mathcal{F} \implies E[|X_n|, \Lambda] < \varepsilon, \quad n \in \mathbb{N}$$

(U3) $X_n \rightarrow X$, a.e. $\implies X$ は可積分

(U4) $E[|Y|] < \infty \implies X_n - Y, n \in \mathbb{N}$ は一様可積分

(\because) (U.1): a を固定したとき (2.8) より $A \equiv \sup_n E[|X_n|, |X_n| \geq a] < \infty$. したがって $E[|X_n|] \leq A + a$.

(U.3): (U.1) と Fatou の補題を合わせると $E[X] \leq \liminf_{n \rightarrow \infty} E[|X_n|] \leq A + a$. (U.2), (U.4) の証明は各自で確認すること.

定理 2.5 $X_n, n \in \mathbb{N}$ が一様可積分 \iff (U.1), (U.2) をみたす

証明 (\implies) はすでに示されている. (\impliedby) を示す. $X_n, n \in \mathbb{N}$ が (U.1), (U.2) をみたすとする. チェビセフの不等式より

$$P(|X_n| \geq a) \leq a^{-1} \sup_{n \in \mathbb{N}} E[|X_n|] \leq Ka^{-1}, \quad n \in \mathbb{N}$$

となる. ゆえに (U.2) の δ に対して a を $K/a < \delta$ となるように十分大きくとれば, $P(|X_n| \geq a) < \delta$ となり, (U.2) より

$$E[|X_n|, |X_n| \geq a] < \varepsilon, \quad n \in \mathbb{N}.$$

□

定理 2.6 確率変数列 $X_n, n \in \mathbb{N}$ が X に概収束しているとする. このとき次の 3 条件は同値である.

- (1) $X_n, n \in \mathbb{N}$ は一様可積分である.
- (2) $\lim_{n \rightarrow \infty} E[|X_n - X|] = 0$.
- (3) $\lim_{n \rightarrow \infty} E[|X_n|] = E[|X|] < \infty$.

証明 (1) \implies (2): (U.3), (U.4) により $X - X_n$ は, 一様可積分.

$$\begin{aligned} E[|X_n - X|] &= E[|X_n - X|, |X - X_n| > \varepsilon/2] + E[|X_n - X|, |X - X_n| \leq \varepsilon/2] \\ &\leq E[|X_n - X|, |X - X_n| > \varepsilon/2] + \varepsilon/2 \end{aligned}$$

(U.2) より, 任意の $\varepsilon > 0$ に対して $\delta > 0$ が存在して,

$$P(\Lambda) < \delta, \Lambda \in \mathcal{F} \implies E[|X_n - X|, \Lambda] < \varepsilon, \quad n \in \mathbb{N}$$

一方, $X_n \rightarrow X$ a.s. より, ある自然数 N が存在して

$$P(|X_n - X| > \varepsilon/2) < \delta, \quad n \geq N$$

となるから, $E[|X_n - X|, |X - X_n| > \varepsilon/2] < \varepsilon/2, n \geq N$ となる. 以上より $E[|X_n - X|] < \varepsilon$.

(2) \implies (3): $|E[|X_n|] - E[|X|]| \leq E[||X| - |X_n||] \leq E[|X - X_n|] \rightarrow 0, n \rightarrow \infty$.

(3) \implies (1): ξ が $(-\infty, \infty)$ で定義された実数値関数であるとき, $\xi^a(\omega) = \xi(\omega)\mathbf{1}(|\xi(\omega)| < a)$ とおくことにする. $P(|X| = a) = 0$ の場合, $X_n^a \rightarrow X^a$ a.s. となる. したがって $|X_n^a| \rightarrow |X^a|$ a.s., さらに $|X_n^a| \leq a$ となるため, 有界収束定理を用いれば

$$\lim_{n \rightarrow \infty} E[|X_n|, |X_n| \leq a] = \lim_{n \rightarrow \infty} E[|X_n^a|] = E[|X^a|] = E[|X|, |X| \leq a]$$

となる. $|X|$ の可積分性より, 任意の $\varepsilon > 0$ に対して, $A > 0$ が存在して

$$E[|X|, |X| > a] < \varepsilon, \quad a \geq A.$$

となるので, $a \geq A, P(|X| = a) = 0$ のとき,

$$E[|X_n|, |X_n| > a] = E[|X_n|] - E[|X_n|, |X_n| \leq a] < 2\varepsilon, \quad n \geq N, \quad \max_{1 \leq k \leq N} E[|X_k|, |X_k| \geq b] < 2\varepsilon$$

をみたす 自然数 N , 実数 $b \geq a$ が存在する. したがって

$$E[|X_n|, |X_n| \geq b] < 2\varepsilon, \quad n \in \mathbb{N}$$

となるので, $X_n, n \in \mathbb{N}$ は一様可積分となる. □

上の定理から次の定理が導かれる.

定理 2.7 確率変数列 $X_n, n \in \mathbb{N}$ が一様可積分であり, X に概収束しているれば, $\lim_{n \rightarrow \infty} E[X_n] = E[X] < \infty$.

2.5 分布と分布関数

この節を通じて, (Ω, \mathcal{F}, P) を確率空間とする.

定義 2.10 (Ω', \mathcal{F}') を可測空間とする. 写像 $T: \Omega \rightarrow \Omega'$ が可測であるとは

$$T^{-1}(A) = \{\omega \in \Omega : T(\omega) \in A\} \in \mathcal{F}, \forall A \in \mathcal{F}'$$

が成り立つことをいう.

演習 2.8.

(i) (逆像の性質) $T^{-1}(\Omega' \setminus A) = \Omega \setminus T^{-1}(A)$, $T^{-1}(\bigcup_{n=1}^{\infty} A_n) = \bigcup_{n=1}^{\infty} T^{-1}(A_n)$ となることを示せ.

(ii) $\mathcal{G} \equiv \{T^{-1}(A) : A \in \mathcal{F}'\}$ が σ -加法族であることを示せ. また, T が可測であることは $\mathcal{G} \subset \mathcal{F}$ と同値であることを示せ.

補題 2.1 (Ω', \mathcal{F}') を可測空間, $T: \Omega \rightarrow \Omega'$ を可測写像とする.

$$P'(A) = P(T^{-1}(A)), A \in \mathcal{F}'$$

とおけば, P' は確率測度である.

演習 2.9. 上の補題を証明せよ.

定義 2.11 上の P' を $P \circ T^{-1}$ と書き, T の (Ω', \mathcal{F}') に誘導する確率測度 (induced measure by T) という. とくに, $(\Omega', \mathcal{F}') = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ で, $X: \Omega \rightarrow \mathbb{R}$ が確率変数となっているとき, $P \circ X^{-1}$ を P^X と表し, X の分布と呼ぶ.

定理 2.8 (Ω', \mathcal{F}') を可測空間, $T: \Omega \rightarrow \Omega'$, $f: \Omega' \rightarrow \mathbb{R}$ を可測写像とする ($(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ を考えている). このとき $g = f \circ T: \Omega \rightarrow \mathbb{R}$ は可測となる. さらに g が P について可積分となるためには f が $P \circ T^{-1}$ について可積分となる必要かつ十分であり, さらにこのとき

$$(2.9) \quad \mathbb{E}^P[g] = \mathbb{E}^{P \circ T^{-1}}[f]$$

となる. ただし \mathbb{E}^P は P に関する期待値を表す.

証明 $A \in \mathcal{B}(\mathbb{R})$ に対し,

$$g^{-1}(A) = (f \circ T)^{-1}(A) = T^{-1}(f^{-1}(A)) \in \mathcal{F}$$

となるから, g は可測である.

$f \geq 0$ と仮定する. $f_n: \Omega' \rightarrow \mathbb{R}, \in \mathbb{SF}$ を $0 \leq f_n \leq f_{n+1} \nearrow f$ となるようにとる.

$$\sum_{k=1}^m \mathbf{1}_{A_k}(T\omega) = \sum_{k=1}^m \mathbf{1}_{T^{-1}A_k}(\omega)$$

であることに注意すると, $g_n = f_n \circ T$ とおけば, $g_n \in \mathbb{SF}$ で, $0 \leq g_n \leq g_{n+1} \nearrow g$. 定義より明らかに

$$\mathbb{E}^P[g_n] = \mathbb{E}^{P \circ T^{-1}}[f_n]$$

単調収束定理より, $n \rightarrow \infty$ として (2.9) を得る. これを $|f|$ に適用して可積分性の同値性を得, さらに f^{\pm} に適用して一般の f に対する (2.9) を得る. \square

演習 2.10. X を確率変数とし, $\xi: \mathbb{R} \rightarrow \mathbb{R}$ を恒等写像, $\xi(x) = x (\forall x \in \mathbb{R})$ とする. このとき次の等式を示せ.

$$\mathbb{E}^P[X] = \mathbb{E}^{P^X}[\xi].$$

注 3.1. 可測空間 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度を考えるとき, 恒等写像 $\xi: \mathbb{R} \rightarrow \mathbb{R}$ は自然に確率変数と見なすことができる.

定義 2.12 X を確率変数, P^X をその分布とする.

$$F_X(x) = P^X((-\infty, x]) = P(\{\omega | X(\omega) \leq x\}), x \in \mathbb{R}$$

とおき, この F_X を X の**分布関数** という. μ が $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度のとき, 恒等写像 $\xi: \mathbb{R} \rightarrow \mathbb{R}$ の分布関数を F_μ と書き, μ の分布関数と呼ぶ.

次の定理は, 確率分布と分布関数が 1 対 1 であることを示している.

定理 2.9 X, Y を確率変数とする. $P^X = P^Y$ が成り立つためには $F_X = F_Y$ が成り立つことが必要十分である.

証明 必要性は明らかであるので十分性を示す. $\mathcal{A} = \{B \in \mathcal{B}(\mathbb{R}) | P^X(B) = P^Y(B)\}$ とおく. 確率測度の σ -加法性より, \mathcal{A} は Dynkin 族 (定義 2.4 参照) となることは明らかである. $\mathcal{C} = \{(-\infty, x], x \in \mathbb{R}\}$ とおくと, $F_X = F_Y$ より $\mathcal{C} \subset \mathcal{A}$ が成り立つ. 定理 2.2 を適用すると $\mathcal{D}(\mathcal{C}) = \sigma(\mathcal{C}) = \mathcal{B}(\mathbb{R})$ となるので前述の結果と合わせると $\mathcal{B}(\mathbb{R}) \subset \mathcal{A}$ をえる. 定義から $\mathcal{A} \subset \mathcal{B}(\mathbb{R})$ であるから, $\mathcal{B}(\mathbb{R}) = \mathcal{A}$ となり, 十分性が示された. \square

定理 2.10 分布関数 F_X は非減少関数で, 右連続かつ左極限を持つ. さらに $\lim_{x \rightarrow -\infty} F_X(x) = 0, \lim_{x \rightarrow \infty} F_X(x) = 1$ となる. とくに不連続点は高々加算個.

証明 確率測度の単調性と $\bigcap_{\varepsilon > 0} (-\infty, x + \varepsilon] = (-\infty, x], \bigcup_{\varepsilon > 0} (-\infty, x - \varepsilon] = (-\infty, x)$ という関係式より明らかである. $0 \leq F_X \leq 1$ と非減少性より, $F_X(x+) - F_X(x-) \geq 1/n$ なる x は高々 n 個しか存在しない. よって不連続点は高々加算個となる. \square

F が非減少関数で, 右連続かつ左極限を持ち $\lim_{x \rightarrow \infty} F(x) = 1, \lim_{x \rightarrow -\infty} F(x) = 0$ が成り立つとき, **分布関数** と呼ぶ. 次の節の定理 2.12 より, F が分布関数であれば, $F = F_\mu$ をみたす確率分布 μ が存在することが分かる. さらに μ の一意性は, 定理 2.9 から導かれる.

2.6 拡張定理

定理 2.11 (Caratheodory の拡張定理) Ω を集合とし, $\mathcal{A} \subset 2^\Omega$ は加法族

- (i) $\emptyset, \Omega \in \mathcal{A}$
- (ii) $A \in \mathcal{A}$ ならば $A^c \in \mathcal{A}$
- (iii) $A, B \in \mathcal{A}$ ならば $A \cap B, A \cup B \in \mathcal{A}$.

であり, $P: \mathcal{A} \rightarrow [0, 1]$ は

- (a) $P(\Omega) = 1$

(b) $A_1, A_2, \dots \in \mathcal{A}$, $A_i \cap A_j = \emptyset, i \neq j$, $\bigcup_{j=1}^{\infty} A_j \in \mathcal{A}$ ならば

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j)$$

を満たすとする. このとき $(\Omega, \sigma(\mathcal{A}))$ 上の確率測度 Q で,

$$Q(A) = P(A), \quad (\forall A \in \mathcal{A})$$

となるものがただ一つ存在する.

注意 (b) は次と同値である.

(b') $A_n \in \mathcal{A}$ が $A_n \supset A_{n+1} \searrow \emptyset$ ならば, $P(A_n) \rightarrow 0 (n \rightarrow \infty)$.

証明.

$$P^*(B) = \inf \left\{ \sum_{j=1}^{\infty} P(A_j) : \bigcup_{j=1}^{\infty} A_j \supset B, A_j \in \mathcal{A} \right\}, \quad B \in 2^\Omega$$

とおき, 外測度と呼ぶ.

Step 1) P^* は P の拡張となっている.

$$(2.10) \quad P^*(A) = P(A), \quad \forall A \in \mathcal{A}.$$

(\because) $A \in \mathcal{A}$ とする. $A \subset A$ より, $P^*(A) \leq P(A)$ は明らか. $\varepsilon > 0$ とし, $A_j \in \mathcal{A}$ を $\bigcup_{j=1}^{\infty} A_j \supset A$, $\sum_{j=1}^{\infty} P(A_j) \leq P^*(A) + \varepsilon$ となるようにとる. $C_j = A \cap \left\{ A_j \setminus \left(\bigcup_{k=1}^{j-1} A_k \right) \right\}$ とすれば, $C_j \in \mathcal{A}$, $\bigcup_{j=1}^{\infty} C_j = A$, $C_i \cap C_j = \emptyset$ となる. 条件 (b) より

$$P(A) = \sum_{j=1}^{\infty} P(C_j) \leq \sum_{j=1}^{\infty} P(A_j) \leq P^*(A) + \varepsilon.$$

$\varepsilon \rightarrow 0$ とすれば $P(A) \leq P^*(A)$.

Step 2)

$$(2.11) \quad P^*\left(\bigcup_{j=1}^{\infty} B_j\right) \leq \sum_{j=1}^{\infty} P^*(B_j), \quad \forall B_1, B_2, \dots \in 2^\Omega.$$

(\because) $\varepsilon > 0$ を任意に固定し, $A_{n,k} \in \mathcal{A}$ を $\bigcup_{k=1}^{\infty} A_{n,k} \supset B_n$, $\sum_{k=1}^{\infty} P(A_{n,k}) < P^*(B_n) + \varepsilon 2^{-n}$ となるように選ぶ.

このとき, $\bigcup_{n,k=1}^{\infty} A_{n,k} \supset \bigcup_{n=1}^{\infty} B_n$ であるから,

$$P^*\left(\bigcup_{n=1}^{\infty} B_n\right) \leq \sum_{n,k=1}^{\infty} P(A_{n,k}) \leq \sum_{n=1}^{\infty} \{P^*(B_n) + \varepsilon 2^{-n}\} \leq \sum_{n=1}^{\infty} P^*(B_n) + \varepsilon.$$

$\varepsilon \downarrow 0$ とすればよい.

Step 3) $\mathcal{M} = \{B \subset \Omega \mid P^*(G) = P^*(B \cap G) + P^*(B^c \cap G), \forall G \subset \Omega\}$ とおく. このとき

$$\begin{aligned}
& \text{(a)} \emptyset, \Omega \in \mathcal{M} \\
& \text{(b)} A \in \mathcal{M} \implies A^c \in \mathcal{M} \\
& \text{(c)} B_k \in \mathcal{M}, k = 1, 2, \dots, n, \implies \bigcup_{k=1}^n B_k \in \mathcal{M} \\
& \text{(d)} B_k \in \mathcal{M}, k = 1, 2, \dots, n, B_i \cap B_j = \emptyset, i \neq j \implies \\
& \quad P^* \left(\bigcup_{k=1}^n (B_k \cap G) \right) = \sum_{k=1}^n P^*(B_k \cap G), \quad \forall G \subset \Omega.
\end{aligned}$$

が成り立つ。

(\because) (b) は明らかである。 $P^*(\emptyset) = 0$ と (b) より, (a) が従う。 (c) については, de Morgan の法則, (b) と帰納法により $B_1, B_2 \in \mathcal{M}$ のときに $B_1 \cap B_2 \in \mathcal{M}$ を示せばよい。 $B_1, B_2 \in \mathcal{M}$ より

$$\begin{aligned}
& P^*((B_1 \cap B_2)^c \cap G) + P^*((B_1 \cap B_2) \cap G) \\
&= P^*((B_1^c \cup B_2^c) \cap G) + P^*((B_1 \cap B_2) \cap G) \\
&= P^*(B_2 \cap \{(B_1^c \cup B_2^c) \cap G\}) + P^*(B_2^c \cap \{(B_1^c \cup B_2^c) \cap G\}) + P^*((B_1 \cap B_2) \cap G) \\
&= P^*(B_2 \cap B_1^c \cap G) + P^*(B_2^c \cap G) + P^*((B_1 \cap B_2) \cap G) \\
&= P^*(B_2 \cap G) + P^*(B_2^c \cap G) = P^*(G)
\end{aligned}$$

したがって $B_1 \cap B_2 \in \mathcal{M}$.

(d) $n = 2$ のときに示せばよい。 $B_1, B_2 \in \mathcal{M}, B_1 \cap B_2 = \emptyset, G \subset \Omega$ とする。このとき $B_1 \in \mathcal{M}$ であるから,

$$\begin{aligned}
& P^*((B_1 \cup B_2) \cap G) \\
&= P^*(B_1 \cap \{(B_1 \cup B_2) \cap G\}) + P^*(B_1^c \cap \{(B_1 \cup B_2) \cap G\}) \\
&= P^*(B_1 \cap G) + P^*(B_2 \cap G).
\end{aligned}$$

Step 4) 次が成り立つ。

$$\begin{aligned}
& \text{(e)} B_n \in \mathcal{M}, n = 1, 2, \dots \implies \bigcup_{n=1}^{\infty} B_n \in \mathcal{M} \\
& \text{(f)} B_n \in \mathcal{M}, n = 1, 2, \dots, B_i \cap B_j = \emptyset \implies \\
& \quad P^* \left(\bigcup_{n=1}^{\infty} B_n \right) = \sum_{n=1}^{\infty} P^*(B_n).
\end{aligned}
\tag{2.12}$$

(\because) $B_n \in \mathcal{M}$ は互いに素であるとし, $G \subset \Omega$ とする。 $B = \bigcup_{n=1}^{\infty} B_n$ とおく。 $G = (B \cap G) \cup (B^c \cap G)$ であるから, (2.11) より

$$P^*(G) \leq P^*(B \cap G) + P^*(B^c \cap G).
\tag{2.13}$$

$M_n = \bigcup_{k=1}^n B_k$ とおく。 **step3(c)** より $M_n \in \mathcal{M}$ である。 よって

$$P^*(G) = P^*(M_n \cap G) + P^*(M_n^c \cap G).$$

$M_n \subset B$ より, $M_n^c \supset B^c$ である. よって $P^*(M_n^c \cap G) \geq P^*(B^c \cap G)$. また, **step3(d)** より $P^*(M_n \cap G) = \sum_{k=1}^n P^*(B_k \cap G)$. これらをあわせると

$$P^*(G) \geq \sum_{k=1}^n P^*(B_k \cap G) + P^*(B^c \cap G).$$

$n \rightarrow \infty$ とし, さらに (2.11) より $\sum_{k=1}^{\infty} P^*(B_k \cap G) \geq P^*(B \cap G)$ となることに注意すれば,

$$P^*(G) \geq \sum_{k=1}^{\infty} P^*(B_k \cap G) + P^*(B^c \cap G) \geq P^*(B \cap G) + P^*(B^c \cap G)$$

となる. (2.13) とあわせて

$$P^*(G) = \sum_{k=1}^{\infty} P^*(B_k \cap G) + P^*(B^c \cap G) = P^*(B \cap G) + P^*(B^c \cap G).$$

よって $B \in \mathcal{M}$ となる. この等式で $G = B$ とおけば (f) も従う.

Step 5) $\mathcal{A} \subset \mathcal{M}$. 特に, $\sigma(\mathcal{A}) \subset \mathcal{M}$ である.

$A \in \mathcal{A}$, $G \subset \Omega$ とする. $A_n \in \mathcal{A}$ を $\bigcup_{n=1}^{\infty} A_n \supset G$ ととる. このとき

$$A \cap G \subset \bigcup_{n=1}^{\infty} (A \cap A_n), \quad A^c \cap G \subset \bigcup_{n=1}^{\infty} (A^c \cap A_n)$$

であるから,

$$\begin{aligned} \sum_{n=1}^{\infty} P(A_n) &= \sum_{n=1}^{\infty} P(A \cap A_n) + \sum_{n=1}^{\infty} P(A^c \cap A_n) \\ &\geq P^*(A \cap G) + P^*(A^c \cap G). \end{aligned}$$

ここで $\{A_n\}$ について下限をとれば

$$P^*(G) \geq P^*(A \cap G) + P^*(A^c \cap G).$$

逆の不等式は P^* の劣加法性より従うので,

$$P^*(G) = P^*(A \cap G) + P^*(A^c \cap G).$$

すなわち, $A \in \mathcal{M}$.

以上の結果をまとめると, $Q(B) = P^*(B)$ ($B \in \sigma(\mathcal{A})$) とおけば望む Q を得る.

Step 6) (一意性) Q_1, Q_2 が $\sigma(\mathcal{A})$ の確率測度であり, $Q_1(A) = Q_2(A)$, $A \in \mathcal{A}$, を満たすとする. このとき $\{A \in \sigma(\mathcal{A}) : Q_1(A) = Q_2(A)\}$ は, $\sigma(\mathcal{A})$ と一致する. (証明は演習とする.) よって, 一意性が示された. \square

定理 2.12 $F: \mathbb{R} \rightarrow \mathbb{R}$ は非減少かつ右連続で $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow \infty} F(x) = 1$ を満たすとする. このとき $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の確率測度 μ が存在し, $F = F_\mu$ となる.

証明.

$$\mathcal{A} = \left\{ \bigcup_{j=1}^n (a_j, b_j] : n \in \mathbb{N}, -\infty \leq a_1 < b_1 < a_2 < b_2 < \cdots < a_n < b_n \leq \infty \right\}$$

とおく. $P: \mathcal{A} \rightarrow [0, 1]$ を

$$P\left(\bigcup_{j=1}^n (a_j, b_j]\right) = \sum_{j=1}^n (F(b_j) - F(a_j))$$

と定義する. \mathcal{A} は加法族であり, P は定理 2.11 の (a) を満たす. (b) を満たすことを見るには, 任意の $A_j \in \mathcal{A}$, $A_j \supset A_{j+1} \searrow \emptyset$ に対し, $P(A_j) \rightarrow 0$ を示せばよい.

$\inf_j P(A_j) \geq \delta > 0$ と仮定し, 矛盾を導く. $\ell > 0$ を $1 - F(\ell) + F(-\ell) < \delta/2$ となるようにとる. $B_j = A_j \cap [-\ell, \ell]$ とおく. $A_j \setminus B_j \subset (-\infty, -\ell) \cup (\ell, \infty)$ より $P(A_j) - P(B_j) \leq 1 - F(\ell) + F(-\ell) < \delta/2$. したがって $P(B_j) \geq \delta/2$ である. $C_j \subset B_j$ を $D_j = \overline{C_j} \subset B_j$,

$$P(B_j \setminus C_j) \leq \frac{\delta}{10 \cdot 2^j}, \quad \forall j \in \mathbb{N}$$

となるように選ぶ. $E_j = \bigcap_{k=1}^j C_k$, $F_j = \bigcap_{k=1}^j D_k$ とおく. $F_j \supset F_{j+1}$, F_j はコンパクト集合である. $F_j \supset E_j$, $P(E_j) \geq (\delta/2) - \sum_{j=1}^{\infty} P(B_j \setminus C_j) \geq 4\delta/10$. したがって $F_j \neq \emptyset$. 有限交叉性より $\bigcap_{j=1}^{\infty} F_j \neq \emptyset$. ところが $A_j \supset F_j$ より, $F_j \searrow \emptyset$. であつたので, 矛盾する. \square

2.7 直積確率空間

定義 2.13 $(\Omega_1, \mathcal{F}_1)$, $(\Omega_2, \mathcal{F}_2)$ を可測空間とする. $\Omega_1 \times \Omega_2$ 上の σ 加法族

$$\sigma\{A_1 \times A_2 \mid A_i \in \mathcal{F}_i, i = 1, 2\}$$

を $\mathcal{F}_1 \times \mathcal{F}_2$ と表し, \mathcal{F}_1 と \mathcal{F}_2 の直積 σ 加法族という.

定理 2.13 $(\Omega_1, \mathcal{F}_1, P_1)$, $(\Omega_2, \mathcal{F}_2, P_2)$ を確率空間とし, $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$ とおく. (Ω, \mathcal{F}) 上の確率測度 Q で

$$Q(A_1 \times A_2) = P_1(A_1)P_2(A_2), \quad \forall A_i \in \mathcal{F}_i, i = 1, 2$$

を満たすものがただ一つ存在する.

証明 $\mathcal{A}_0 = \{A_1 \times A_2 : A_i \in \mathcal{F}_i, i = 1, 2\}$,

$$\mathcal{A} = \left\{ \bigcup_{j=1}^n B_j : B_j \in \mathcal{A}_0, B_i \cap B_j = \emptyset, i \neq j, n \in \mathbb{N} \right\}$$

とおく. \mathcal{A} が

(i) $\emptyset, \Omega \in \mathcal{A}$, (ii) $B \in \mathcal{A}$ ならば $B^c \in \mathcal{A}$, (iii) $B, C \in \mathcal{A}$ ならば $B \cup C \in \mathcal{A}$ を満たすことは容易に分かる. さらに定義より明らかに $\mathcal{F} = \sigma[\mathcal{A}]$ となる.

$P: \mathcal{A} \rightarrow [0, 1]$ を

$$P\left(\bigcup_{j=1}^n A_{1,j} \times A_{2,j}\right) = \sum_{j=1}^n P_1(A_{1,j})P_2(A_{2,j})$$

と定義する. ただし $A_{i,j} \in \mathcal{F}_i$, $i = 1, 2$, で $(A_{1,k} \times A_{2,k}) \cap (A_{1,j} \times A_{2,j}) = \emptyset$, $k \neq j$. このとき $P(\Omega) = 1$ であることは明らかである..

$E_n \in \mathcal{A}$ は $E_n \supset E_{n+1} \searrow \emptyset$ とする. $\omega_2 \in \Omega_2$ に対し,

$$E_{n,\omega_2} = \{\omega_1 : (\omega_1, \omega_2) \in E_n\}$$

とおく. $E_{n,\omega_2} \in \mathcal{F}_1$ であり, $E_{n,\omega_2} \supset E_{n+1,\omega_2} \searrow \emptyset$ であるので $P_1(E_{n,\omega_2}) \rightarrow 0$ となる. \mathcal{A} と P の定義より, Ω から $[0, 1]$ への写像 $X_n^2 : \omega_2 \mapsto P_1(E_{n,\omega_2})$ は可測であり, $P(E_n) = \mathbb{E}^{P_2}[X_n^2]$ となる. したがって有界収束定理より, $P(E_n) \rightarrow 0$ が成り立つ.

よって定理 2.11 より, \mathcal{A} 上では P と一致する (Ω, \mathcal{F}) 上の確率測度 Q が唯一存在する. \square

定義 2.14 定理 2.13 の Q を $P_1 \times P_2$ と表し, P_1 と P_2 の **直積確率測度** という. また $(\Omega, \mathcal{F}, P_1 \times P_2)$ を $(\Omega_1, \mathcal{F}_1, P_1)$ と $(\Omega_2, \mathcal{F}_2, P_2)$ の **直積確率空間** という.

定理 2.14 (Fubini の定理) $(\Omega_i, \mathcal{F}_i, P_i)$ を確率空間, $(\Omega, \mathcal{F}, P_1 \times P_2)$ をその直積確率空間とする. 確率変数 $X : \Omega \rightarrow \mathbb{R}$ に対し,

$$Y_{\omega_1}(\omega_2) = Z_{\omega_2}(\omega_1) = X(\omega_1, \omega_2), \quad \omega_i \in \Omega_i, \quad i = 1, 2$$

とおく.

- (i) 各 ω_1, ω_2 に対し, $Y_{\omega_1} : \Omega_2 \rightarrow \mathbb{R}$, $Z_{\omega_2} : \Omega_1 \rightarrow \mathbb{R}$ は可測である.
- (ii) X が可積分ならば, P_1 に関してほとんどすべての ω_1 について Y_{ω_1} は P_2 -可積分であり, P_2 に関してほとんどすべての ω_2 について Z_{ω_2} は P_1 -可積分である.
- (iii) X が可積分ならば, 写像 $\Omega_1 \ni \omega_1 \mapsto \mathbb{E}^{P_2}[Y_{\omega_1}]$, $\Omega_2 \ni \omega_2 \mapsto \mathbb{E}^{P_1}[Z_{\omega_2}]$ はともに可測である.
- (iv) X が可積分ならば,

$$\mathbb{E}^{P_1 \times P_2}[X] = \mathbb{E}^{P_1}[\mathbb{E}^{P_2}[Y_{\omega_1}]] = \mathbb{E}^{P_2}[\mathbb{E}^{P_1}[Z_{\omega_2}]].$$

もし $X \geq 0$ ならば上の主張は可積分性の仮定なしに成立する.

証明.

$$\mathcal{G} = \{A \in \mathcal{F} : 1_A \text{ が (i)~(iv) を満たす}\}$$

とおく. まず (a) $\emptyset, \Omega \in \mathcal{G}$, (b) $A \in \mathcal{G}$ ならば $A^c \in \mathcal{G}$, は定義より明らかである.

(c) $A_j \in \mathcal{G}$, $j \in \mathbb{N}$, ならば, 定義関数の分解を用いて $\bigcup_{j=1}^n A_j \in \mathcal{G}$ となり, さらに単調収束定理より, $\bigcup_{j=1}^{\infty} A_j \in \mathcal{G}$ となる. (d) $A_i \in \mathcal{F}_i$, $i = 1, 2$, ならば, 直ちに $A_1 \times A_2 \in \mathcal{G}$ となる. したがって $\mathcal{G} = \mathcal{F}$ である.

線形性より $X \in \mathbb{S}\mathbb{F}$ は (i)~(iv) を満たすといえる. $X \geq 0$ ならば, $X_n \in \mathbb{S}\mathbb{F}$ を $0 \leq X_n \leq X_{n+1} \nearrow X$ となるようにとり, 単調収束定理を用いればよい. 一般の X については線形性より主張が従う. \square

2.8 連続確率分布の例

定義 (連続確率密度) p が $\mathbb{R} = (-\infty, \infty)$ 上の確率密度であるとは

- (i) $p(x) \geq 0$, $x \in \mathbb{R}$
- (ii) $\int_{-\infty}^{\infty} p(x) dx = 1$

が成り立つ事である. A を \mathbb{R} 上の事象としたとき

$$P(A) = \int_A p(x) dx$$

と定義すると P は確率となる.

連続確率分布の平均, 2次モーメント, 分散, 標準偏差は次で定義される.

- 平均

$$\mu = \int_{-\infty}^{\infty} xp(x)dx$$

- 二次モーメント

$$m_2 = \int_{-\infty}^{\infty} x^2p(x)dx$$

- 分散

$$V = \int_{-\infty}^{\infty} (x - \mu)^2p(x)dx = m_2 - \mu^2$$

- 標準偏差

$$\sigma = \sqrt{V}$$

- 積率母関数

$$\mu(t) = \int_{-\infty}^{\infty} e^{tx}p(x)dx$$

2.8.1 一様分布 (Uniform distribution)

$-\infty < a < b < \infty$ に対して $S = (a, b)$,

$$p(x) = \begin{cases} \frac{1}{b-a} & (a < x < b) \\ 0 & (\text{その他}) \end{cases}$$

と定義したとき, $p(x)$ を確率密度とする分布を区間 (a, b) 上の一様分布という.

一様分布の平均, 2次モーメント, 分散, 標準偏差の計算

- 平均

$$\mu = \int_{-\infty}^{\infty} xp(x)dx = \frac{1}{b-a} \int_a^b xdx = \frac{1}{2} \times \frac{b^2 - a^2}{b-a} = \frac{a+b}{2}$$

- 2次モーメント

$$m_2 = \int_{-\infty}^{\infty} x^2p(x)dx = \frac{1}{b-a} \int_a^b x^2dx = \frac{1}{b-a} \times \frac{b^3 - a^3}{3} = \frac{a^2 + ab + b^2}{3}$$

- 分散

$$V = m_2 - \mu^2 = \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{a^2 - 2ab + b^2}{12} = \frac{(a-b)^2}{12}$$

- 標準偏差

$$\sigma = \sqrt{V} = \frac{|a-b|}{2\sqrt{3}}$$

- 積率母関数

$$m_t(X) = \int_{-\infty}^{\infty} e^{tx}p(x)dx = \frac{1}{b-a} \int_a^b e^{tx}dx = \frac{1}{b-a} \left[\frac{1}{t} e^{tx} \right]_a^b = \frac{e^{tb} - e^{ta}}{(b-a)t}$$

2.8.2 指数分布 (Exponential distribution)

確率密度関数 $p(x)$ が

$$p(x) = \begin{cases} \lambda e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

である分布を (パラメータ $\lambda > 0$ の) 指数分布という.

● 対応するモデル

- (i) 電球の寿命などのさまざまな事物の存続期間
- (ii) 一定の時間間隔における出来事回数がポアソン分布であるとき, つぎつぎに起こる出来事の時間間隔 (詳しくは次のガンマ分布のところで説明する.)

指数分布の性質 X の分布が指数分布であるとき

$$(2.14) \quad P(X \geq t+h | X \geq t) = P(X \geq h)$$

が成り立つ. つまり X を電球の寿命とすると時刻 t まで電球が切れていない時, 時刻 $t+h$ まで電球が切れない確率は t に無関係である. また, (2.14) を満たし, $P(X \leq h) = h\lambda + o(h)$, $h \rightarrow 0$ が成り立てば X の分布はパラメータ λ の指数分布である.

証明 前半は演習とする. 後半を証明する.

$$\begin{aligned} P(X \geq t+h) &= P(X \geq t+h | X \geq t) \times P(X \geq t) \\ \text{寿命が } t+h \text{ 以上} & \quad \text{寿命が } t \text{ 以上である} \quad \text{寿命が } t \text{ 以上} \\ & \quad \text{条件の下で } t+h \text{ 以上} \\ &= P(X \geq h)P(X \geq t) \end{aligned}$$

をえる. 両辺を微分すると

$$\begin{aligned} \frac{d}{dt} P(X \geq t) &= \lim_{h \rightarrow 0} \frac{P(X \geq t+h) - P(X \geq t)}{h} \\ &= P(X \geq t) \lim_{h \rightarrow 0} \frac{P(X \geq h) - 1}{h} \\ & \quad \downarrow \\ & \quad -\lambda \\ &= -\lambda P(X \geq t) \end{aligned}$$

となる. したがって $P(X \geq t) = ce^{-\lambda t}$ が導かれるが, $t = 0$ のとき $P(X \geq 0) = 1$ であることに注意すると $c = 1$ をえるので

$$P(X \geq t) = e^{-\lambda t} = \int_t^{\infty} p(x) dx$$

両辺を t で微分すれば $-\lambda e^{-\lambda t} = -p(t)$ つまり $p(t) = \lambda e^{-\lambda t}$ をえる.

● 平均

$$\mu = \int_0^{\infty} p(x)x dx = [-xe^{-\lambda x}]_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx = \left[-\frac{1}{\lambda} e^{-\lambda x}\right]_0^{\infty} = \frac{1}{\lambda}$$

- 2次モーメント

$$m_2 = \int_0^{\infty} p(x)x^2 dx = \lambda \int_0^{\infty} x^2 e^{-\lambda x} dx = [-x^2 e^{-\lambda x}]_0^{\infty} + 2 \int_0^{\infty} x e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

- 分散

$$V = m_2 - \mu^2 = \frac{1}{\lambda^2}$$

- 標準偏差

$$\sigma = \sqrt{V} = \frac{1}{\lambda}$$

2.8.3 ガンマ分布 (Γ -distribution)

確率密度関数 $p(x)$

$$p(x) = \begin{cases} \frac{\lambda}{\Gamma(r)} (\lambda x)^{r-1} e^{-\lambda x} & (x \geq 0) \\ 0 & (x < 0) \end{cases}$$

を持つ分布を (パラメータ $\lambda > 0, r > 0$ の) ガンマ分布という.

ここで $\Gamma(r)$ は, ガンマ関数

$$\Gamma(r) = \int_0^{\infty} x^{r-1} e^{-x} dx = \int_0^{\infty} \lambda (\lambda x)^{r-1} e^{-\lambda x} dx$$

であり,

$$\begin{aligned} \Gamma(n+1) &= n! && : n \text{ は } 0 \text{ 以上の整数} \\ \Gamma\left(\frac{1}{2}\right) &= \sqrt{\pi} && : \left(-\frac{1}{2}\right)! = \sqrt{\pi} \\ \Gamma(z+1) &= z\Gamma(z) && : z \text{ は実数} \end{aligned}$$

という性質が知られている.

注) 指数分布はガンマ分布の $r = 1$ の場合に対応する.

- 対応するモデル (r 個の電球の寿命の和) i 番目の電球の寿命を X_i とする. (X_i の分布は指数分布)

このとき $Y = \sum_{i=1}^r X_i$ の分布はガンマ分布となる.

ガンマ分布の平均, 分散, 積率母関数の計算.

まず積率母関数を計算してから, 平均, 分散を計算する.

- 積率母関数

$$\begin{aligned} m_t(Y) &= \int_0^{\infty} p(x)e^{tx} dx = \int_0^{\infty} \frac{\lambda^r}{\Gamma(r)} e^{tx} x^{r-1} e^{-\lambda x} dx \\ &= \left(\frac{\lambda}{\lambda-t}\right)^r \int_0^{\infty} \frac{(\lambda-t)^r}{\Gamma(r)} x^{r-1} e^{-(\lambda-t)x} dx = \left(\frac{\lambda}{\lambda-t}\right)^r \end{aligned}$$

- 平均

$$\begin{aligned} \mu &= \frac{\partial}{\partial t} m_t(Y) \Big|_{t=0} = \frac{\partial}{\partial t} \left(\frac{\lambda}{\lambda-t}\right)^r \Big|_{t=0} \\ &= \lambda^r \frac{\partial}{\partial t} (\lambda-t)^{-r} \Big|_{t=0} = r\lambda^r (\lambda-t)^{-r-1} \Big|_{t=0} = \frac{r}{\lambda} \end{aligned}$$

• 2次モーメント

$$m_2 = \left. \frac{\partial}{\partial t^2} m_t(Y) \right|_{t=0} = \left. \frac{\partial}{\partial t} \{r\lambda^r(\lambda-t)^{-r-1}\} \right|_{t=0}$$

$$= r\lambda^r(r+1)\lambda^{-r-2} = \frac{r(r+1)}{\lambda^2}$$

• 分散

$$V(Y) = E[Y^2] - EY^2 = \frac{r(r+1)}{\lambda^2} - \left(\frac{r}{\lambda}\right)^2 = \frac{r}{\lambda^2}$$

ガンマ分布と指数分布の関係を表にしてみる.

	ガンマ分布	指数分布
積率母関数: $m_t(Y)$	$\left(\frac{\lambda}{\lambda-t}\right)^r$	$\frac{\lambda}{\lambda-t}$
平均: $E[Y]$	$\frac{r}{\lambda}$	$\frac{1}{\lambda}$
分散: $V(Y)$	$\frac{r}{\lambda^2}$	$\frac{1}{\lambda^2}$

演習 2.11. Z_α をパラメータ $\lambda > 0, \alpha > 0$ のガンマ分布に従う確率変数, Z_β をパラメータ $\lambda > 0, \beta > 0$ のガンマ分布に従う確率変数とし, Z_α と Z_β は独立であるとする. この時, $Z_\alpha + Z_\beta$ の分布は, パラメータ $\lambda > 0, \alpha + \beta > 0$ のガンマ分布に従うことを示せ.

ポアソン分布との関係 Z がパラメータ λ のポアソン分布を持つとする. つまり

$$P(Z = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

が成り立つ. 一方, $X_i, i = 1, 2, \dots, k$ を (互いに独立で) パラメータ λ 指数分布を持つとする. すでに述べたように $X_1 + X_2 + \dots + X_k = Y$ はパラメータ $\lambda > 0, k$ のガンマ分布をもつ. Y が 1 以下である確率を計算してみると

$$P(X_1 + X_2 + \dots + X_k \leq 1) = P(Y \leq 1)$$

$$= \int_0^1 \frac{\lambda(\lambda x)^{k-1}}{\Gamma(k)} e^{-\lambda x} dx$$

$$= \frac{\lambda^k}{\Gamma(k)} \int_0^1 x^{k-1} e^{-\lambda x} dx$$

$$= \frac{\lambda^{k-1}}{\Gamma(k-1)} \int_0^1 x^{k-2} e^{-\lambda x} dx - \frac{\lambda^{k-1}}{\Gamma(k)} e^{-\lambda}$$

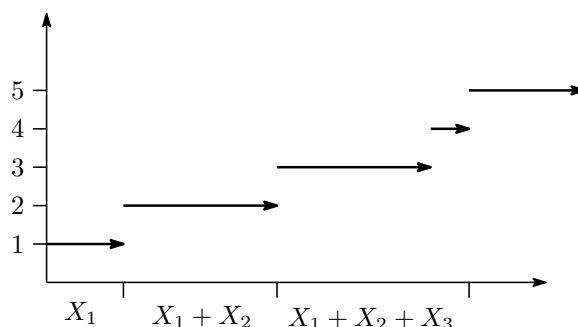
$$= 1 - \sum_{j=1}^k \frac{\lambda^{j-1}}{\Gamma(j)} e^{-\lambda}$$

$$= \sum_{j=k}^{\infty} \frac{\lambda^j}{\Gamma(j+1)} e^{-\lambda}$$

となる。したがって

$$\begin{aligned} & P(X_1 + X_2 + \cdots + X_n \leq 1 < X_1 + X_2 + \cdots + X_{n+1}) \\ &= P(X_1 + X_2 + \cdots + X_n \leq 1) - P(X_1 + X_2 + \cdots + X_{n+1} \leq 1) \\ &= P(Z = n) \end{aligned}$$

が成り立つ。電球の寿命のモデルでこの関係を説明してみると、「時刻 1 までに切れる電球の個数の分布はポアソン分布である」ということになる。



2.8.4 ベータ分布 (Beta distribution)

$\alpha > 0, \beta > 0$ に対して,

$$p(x) = \begin{cases} \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} & x \in S = (0, 1) \\ 0 & (\text{その他}) \end{cases}$$

と定義したとき, $p(x)$ を確率密度とする分布をパラメータ $\alpha > 0, \beta > 0$ のベータ分布という. $B(\alpha, \beta)$ は積分して 1 にするための規格化定数であって,

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$$

で定義される. これをベータ関数とよび, ガンマ関数 $\Gamma(x)$ と

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}$$

という関係をもつ. とくに $\alpha = 1, \beta = 1$ のときには, ベータ分布が一様分布と一致することがわかる. ベータ関数の特徴は α, β の値によっていろいろな形をとることである. 確率分布のおおよその形がわかっているその形に関数をあてはめたいときに, ベータ関数が用いられる. 確率密度関数の形より平均, 分散は簡単に計算できる.

- 平均 $\mu = \frac{\alpha}{\alpha + \beta}$.
- 分散 $V = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$.

演習 2.12. Z_α をパラメータ $\lambda > 0, \alpha > 0$ のガンマ分布に従う確率変数, Z_β をパラメータ $\lambda > 0, \beta > 0$ のガンマ分布に従う確率変数とし, Z_α と Z_β は独立であるとする. この時, $Z_{\alpha, \beta} \equiv \frac{Z_\alpha}{Z_\alpha + Z_\beta}$ はパラメータ $\alpha > 0, \beta > 0$ のベータ分布に従い, $Z_\alpha + Z_\beta$ と独立であることを示せ.

2.8.5 正規分布 (Normal distribution)

$\sigma > 0, -\infty < \mu < \infty$ とする. 確率密度関数 $p(x)$ が

$$p(x) = p_{\mu\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

で定まる分布をパラメータ μ, σ の正規分布という. ($N(\mu, \sigma^2)$ と書く) $\mu = 0, \sigma = 1$ のときの正規分布を標準正規分布 (規準正規分布) という. このとき確率密度関数は,

$$p(x) = p_{01}(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\}$$

となる. まず $p(x)$ が確率密度関数であること, つまり

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

を示しておく.

証明 変数変換を用いると

$$\begin{aligned} \int_{-\infty}^{\infty} p(x) dx &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{y^2}{2}\right\} \sigma dy \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy \end{aligned}$$

となるので,

$$(2.15) \quad \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \sqrt{2\pi}$$

を示せばよい.

演習 2.13. 等式 (2.15) を証明せよ.

平均, 分散, 標準偏差の計算は積分公式

$$(2.16) \quad \int_{-\infty}^{\infty} \exp\left\{-\frac{y^2}{2}\right\} y dy = 0,$$

$$(2.17) \quad \int_{-\infty}^{\infty} \exp\left\{-\frac{y^2}{2}\right\} y^2 dy = \int_{-\infty}^{\infty} \exp\left\{-\frac{y^2}{2}\right\} dy = \sqrt{2\pi}$$

を用いて示される.

演習 2.14. 等式 (2.16), (2.17) を証明せよ.

● 平均

$y = \frac{x-\mu}{\sigma}$ で変数変換を行うと ($dy = \frac{1}{\sigma} dx, x = \mu + \sigma y$)

$$\begin{aligned} \int_{-\infty}^{\infty} x p_{\mu\sigma}(x) dx &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mu + \sigma y) \exp\left\{-\frac{y^2}{2}\right\} dy \\ &= \mu \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{y^2}{2}\right\} dy + \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y \exp\left\{-\frac{y^2}{2}\right\} dy \\ &= \mu \end{aligned}$$

を得る. 最後の等号を導くために積分公式 (2.16) を用いた. 従って,

$$N(\mu, \sigma^2) \text{ の平均} = \mu$$

• 分散

$y = \frac{x - \mu}{\sigma}$ で変数変換を行うと,

$$\begin{aligned} \int_{-\infty}^{\infty} (x - \mu)^2 p_{\mu\sigma}(x) dx &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (x - \mu)^2 \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \sigma^2 y^2 \exp\left\{-\frac{y^2}{2}\right\} dy \\ &= \sigma^2 \end{aligned}$$

を得る. 最後の等式を導くために積分公式 (2.17) を用いた. 従って,

$$\begin{array}{ll} N(\mu, \sigma^2) \text{ の分散} & : \quad \sigma^2 \\ N(\mu, \sigma^2) \text{ の標準偏差} & : \quad \sigma \end{array}$$

• 積率母関数

$$\begin{aligned} \int_{-\infty}^{\infty} e^{tx} p(x) dx &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp\left\{tx - \frac{(x - \mu)^2}{2\sigma^2}\right\} dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp\left\{-\frac{(x - \mu - \sigma^2 t)^2}{2\sigma^2}\right\} \exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\} dx \\ &= \exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\} \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp\left\{-\frac{(x - \mu - \sigma^2 t)^2}{2\sigma^2}\right\} dx \\ &= \exp\left\{\frac{\sigma^2 t^2}{2} + \mu t\right\} \end{aligned}$$

2行目の等式では,

$$\begin{aligned} tx - \frac{(x - \mu)^2}{2\sigma^2} &= -\frac{1}{2\sigma^2} \{(x - \mu)^2 - 2\sigma^2 tx\} \\ &= -\frac{1}{2\sigma^2} (x^2 - 2\mu x + \mu^2 - 2\sigma^2 tx) \\ &= -\frac{1}{2\sigma^2} \{(x - \mu - \sigma^2 t)^2 - 2\mu\sigma^2 t - \sigma^4 t^2\} \\ &= -\frac{1}{2\sigma^2} (x - \mu - \sigma^2 t)^2 + \frac{\sigma^2 t^2}{2} + \mu t \end{aligned}$$

を用いた.

次に正規分布に関する重要な性質を示しておく.

定理 2.15 (標準化, 規準化) X の分布がパラメータ μ, σ の正規分布 $N(\mu, \sigma^2)$ であるとき, $Y = \frac{X - \mu}{\sigma}$ の分布は標準正規分布 $N(0, 1)$ である.

証明)

$y = \frac{x - \mu}{\sigma}$ として変数変換を行うと,

$$\begin{aligned} P(a \leq Y \leq b) &= P(a\sigma + \mu \leq X \leq b\sigma + \mu) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{a\sigma + \mu}^{b\sigma + \mu} \exp\left\{-\frac{(x - \mu)^2}{2\sigma^2}\right\} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_a^b \exp\left\{-\frac{y^2}{2}\right\} dy \end{aligned}$$

2.8.6 多次元正規分布 (Multi-dimensional Normal distribution)

$\mathbf{m} = (m_1, m_2, \dots, m_n) \in \mathbb{R}^n$, $V = (v_{ij})_{1 \leq i, j \leq n}$ を $n \times n$ 正定値対称行列とする. 確率密度関数 $p(\mathbf{x}) = p(x_1, x_2, \dots, x_n)$ が

$$p(\mathbf{x}) = \sqrt{\frac{1}{(2\pi)^n \det V}} \exp \left\{ -\frac{1}{2} \sum_{i,j=1}^n (x_i - m_i)(V^{-1})_{ij}(x_j - m_j) \right\}$$

で定まる分布をパラメータ \mathbf{m}, \mathbf{v} の n 次元正規分布という. ($N(\mathbf{m}, V)$ と書く)

演習 2.15. $\mathbf{X} = (X_1, X_2, \dots, X_n)$ を $N(\mathbf{m}, V)$ に従う確率変数とする. 対称行列は, 直交行列により対角化できることを用いて, $E[X_i] = m_i$, $Cov(X_i, X_j) = V_{ij}$, $i, j = 1, 2, \dots, n$, となることを計算せよ.

2.8.7 コーシー分布 (Cauchy distribution)

$\alpha > 0, \lambda \in (-\infty, \infty)$ に対して,

$$p(x) = \begin{cases} \frac{\alpha}{\pi\{\alpha^2 + (x - \lambda)^2\}} & x \in S = (-\infty, \infty) \\ 0 & (\text{その他}) \end{cases}$$

と定義したとき, $p(x)$ を確率密度とする分布をコーシー分布という. コーシー分布の密度関数の形は正規分布のそれと似ているが, 詳しく調べるとまったく異なっている. 最も大きく異なることは, 平均も分散も存在しないことである.

2.8.8 対数正規分布 (log-normal distribution)

$\mu \in (-\infty, \infty), \sigma > 0$ に対して,

$$p(x) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma x}} \exp\left\{-\frac{(\log x - \mu)^2}{2\sigma^2}\right\} & x \in S = (0, \infty) \\ 0 & (\text{その他}) \end{cases}$$

と定義したとき, $p(x)$ を確率密度とする分布を対数正規分布という. その平均と分散は

平均 $\mu = \exp\left(\mu + \frac{\sigma^2}{2}\right)$,

分散 $V = \exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2)$,

である.

ランダムに世帯を選びその年間所得 X を調べると, 低い方は一定限度があるが高い方には明確な限度がない. このような場合は, 対数をとると有効である. この例では $\log X$ の分布が正規分布に近いことが知られている. $\log X$ が正規分布に従うときもとの X は対数正規分布に従う.

2.8.9 ワイブル分布 (Weibul distribution)

$a, b > 0$ に対して,

$$p(x) = \begin{cases} \frac{bx^{b-1}}{a^b} \exp\left\{-\left(\frac{x}{a}\right)^b\right\} & x \in S = (0, \infty) \\ 0 & (\text{その他}) \end{cases}$$

と定義したとき, $p(x)$ を確率密度とする分布をワイブル分布という.

一般に耐用年数や寿命は確率変数だが、故障が偶発故障なら瞬間故障率は一定になり、確率変数は指数分布に従う。もし劣化が進行し故障率が増加するときは IFR (Increasing Failure Rate) といい、指数分布に従わない。またいわゆる「初期故障」の時期には故障率の減少が起こる。これを DFR (Decreasing Failure Rate) といい、このときも指数分布に従わない。これらの現象での分布はワイブル分布に近いことが知られている。

ワイブル分布の平均と分散は

平均 $\mu = a\Gamma\left(1 + \frac{1}{b}\right),$

分散 $V = a^2\left\{\Gamma\left(2 + \frac{1}{b}\right) - \Gamma\left(1 + \frac{1}{b}\right)^2\right\},$

である。

母数 a, b はそれぞれ尺度母数, 形状母数とよばれ b の値をかえると分布の形が変化する。 b が大きいときのワイブル分布は正規分布に近づく。正規分布に似ているが厳密には正規分布ではない場合の精密なあてはめにも用いられる。