連続型確率変数の期待値と分散, 大数の法則

赤間 陽二

Abstract. 確率変数の期待値とは, 確率変数の値の散らばりの位置を表す定数である. 確率変数の分散とは, 確率変数の値の散らばりの度合いを表す定数である. 連続型確率変数の確率密度関数を導入し, それを用いて期待値と分散を定義する. 連続型確率変数の期待値は加算と定数倍を保存することを証明する. 確率変数が独立であることを定義し, 独立な連続型確率変数の和の分散はそれぞれの確率変数の分散の和になることを証明する. また分散に関するさまざまな公式を証明する. これにより確率変数に対する大数の法則を説明する.

Contents

1. 5/14 の講義の流れ 1
2. 確率密度関数 2
3. 確率密度関数と期待値 4
4. 確率変数の分散 6
5. 多次元確率変数と同時密度関数 7
6. 確率変数の未知の期待値を反復独立試行により「推定」する 10
References 11
Appendix A. 区分求積法と期待値 12
Appendix B. 広義積分 12
B.1. 広義重積分 13

1. 5/14 の講義の流れ

原則, 証明なし.

(1) 定義 3（期待値）および $E(h(X))$ の定義
(2) 例 4（指数分布のモーメント）の最初 2 文, 例 3（コーシー分布の期待値なし）の最初の文,
(3) 例 6（標準正規分布のモーメント）の最初の文
(4) 定義 7（分散）
(5) 定理 3（分散公式）
(6) 例 7（指数分布, 標準正規分布, コーシー分布の分散）
(7) 定理 4（チェビシェフの不等式）
(8) 定義 5 （多次元確率変数）($p = 2$ のみ)
(9) 同時密度関数（$p = 2$ のみ）
(10) 定理 5（周辺確率密度関数）（$p = 2$ のみ）
(11) 系 1（期待値の線形性）※証明は下線部のみ説明
(12) 確率変数の独立性の復習

Date: 2013 年 5 月 7,14 日.

(13) 定理 6（連続型確率変数の独立性の必要十分条件）
(14) 定理 7（独立確率変数の積の期待値の公式）
(15) 定義 6（共分散）と系 2
(16) 定理 8（独立確率変数に対する分散の加法性）
(17) 定理 9（大数の弱法則）
(18) 定理 10（大数の強法則）

2. 確率密度関数

確率変数とは, 根元事象全体 Ω (全体事象) から実数全体 $\mathbb{R}$ への勝手な関数 X で, 勝手な実数 x に対して, X の観測値が x 以下になる確率 $F_X(x) = P(\{\omega \in \Omega \mid X(\omega) \le x\})$ が定義されるものであった. この単調増加関数 $F_X : \mathbb{R} \to \mathbb{R}$ を確率変数 X の分布関数といった. 勝手な実数 a, b $(a \le b)$ に対して, 確率変数 X が a より大きく b 未満の値を取る確率は $P(a < X \le b) = F_X(b) - F_X(a)$ となる. また連続型確率変数とは, $F_X(x)$ が連続関数となる確率変数のことであった. 例えば, 機器の故障から次の故障までの時間は連続型確率変数である.

定義 1. 非負値関数 $f(x)$ は連続型確率変数 X の 確率密度関数 *(probability density function)* であるというのは, 確率変数 X の分布関数 $F_X(x)$ が広義積分

$$F_X(x) = \int_{-\infty}^{x} f(t)dt. \qquad (x \in \mathbb{R}) \tag{1}$$

である時をいう *(*[1, 50 頁], [6, 70 頁]*)*.

重要なことは, 勝手な確率密度関数 $f(x)$ が

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{2}$$

を満たすことである. というのは, 分布関数 $F_X(x) = \int_{-\infty}^{x} f(t)dt$ が $x \to \infty$ の極限で 1 に収束するからである.

確率変数 X の分布関数 $F_X(x)$ が適当な連続非負値関数 $f_X(x)$ に対して (1) と書ける場合, [5, 定理 35] より, 分布関数 $F_X(x)$ は微分可能であり, $\frac{dF_X(x)}{dx} = f_X(x)$. 従って, 次の式が表すように, 確率密度関数は確率の密度を表す:

$$\lim_{h \to 0+} \frac{P(x < X \le x + h)}{h} = f_X(x).$$

定理 1 (指数分布の無記憶性)**.** *X が正値連続型確率変数であり, 全ての正の実数 s, t に対して*

$$P(X > s + t \mid X > s) = P(X > t) \quad (\underline{\text{無記憶性}})$$

が成立すると仮定する. 適当な正の実数 λ があって, 次の関数

$$f(t) = \begin{cases} 0, & (t \le 0); \\ \lambda e^{-\lambda t}, & (t > 0) \end{cases}$$

は, X の確率密度関数である. ただし $e(= 2.718 \cdots)$ は自然対数の底である. 一般に連続型確率変数 X がこのような確率密度関数を持つ場合, この X は 指数分布 (exponential distribution) に従うといい, このことを記号 $X \sim \mathrm{Exp}(\lambda)$ で表す.

Proof. 連続型確率変数の定義から X の分布関数 $F_X(t) = P(X \le t)$ は非負値連続関数になる. 従って関数 $G(t) = P(X > t)$ もそうである. 条件付き確率の定義から, 無記憶性は $G(t + s) = G(t)G(s)$ を意味する. 故に全ての正整数 n に対して

$G(n)=G(1)^n$ となり, $G\geq 0$ より全ての正整数 m に対して $G(n/m)=G(1)^{n/m}$. G が連続だから $G(t)=G(1)^t$ $(t>0)$. 確率変数 X の値は正だから, t が 0 以下ならば $G(t)=P(X>t)=1$. 分布関数 $F_X(t)$ が単調増加だから, $G(t)=1-F_X(t)$ は単調減少. 故に, 適当な正実数 λ が存在し, $G(t)=1$ $(t\leq 0)$; $e^{-\lambda t}$ $(t>0)$. 故に, 分布関数は $F_X(t)=0$ $(t\leq 0)$; $1-e^{-\lambda t}$ $(t>0)$. これは, ちょうど $\int_{-\infty}^t f(x)dx$ である. □

指数分布に従う連続型確率変数の例としては, 機器の故障から次の故障までの時間がある. 単位時間当たりの故障の回数という離散型確率変数は ポアソン分布 ([1, 54 頁-57 頁], [6, 80 頁]) という分布に従うことが知られていて [6, 184 頁] が, ポアソン分布は指数分布と密接な関係がある.

無記憶性の意味は, 時刻 s までに事象が生起しなかったという情報が与えられた時, その事象がさらに t 時間の間生起しない条件付き確率は, (時刻 s まで事象が生起しなかったという情報が完全に忘れ去られ, 改めてその時点から観測を始めて) t 時間の間事象が生起しない確率に一致する. 事象を死滅と考え, 寿命問題を指数分布を用いて研究するアプローチがある ([1, 62 頁],[6, 82 頁]).

例 1. $a<b$ とする. 連続型確率変数 X が閉区間 $[a,\ b]$ を一様に分布するならば, その分布関数は

$$F_X(x)=\begin{cases}0, & (x\leq a)\\ \frac{x-a}{b-a}, & (a<x\leq b)\\ 1, & (b<x)\end{cases} \tag{3}$$

となる. このような確率変数 X を $[a,\ b]$ 上の一様分布 *(uniform distribution)* に従うといい, このことを記号 $X\sim U(a,b)$ で表す *(*[1, 61 頁]*,* [6, 80 頁]*).*

このとき,

$$f_X(x)=\begin{cases}0, & (x\leq a)\\ \frac{1}{b-a}, & (a<x\leq b)\\ 0, & (b<x)\end{cases}$$

とおくと, この $f_X(t)$ は, $[a,\ b]$ 上の一様分布に従う 確率変数 X の確率密度関数である.

xy 座標平面に振幅 A の周期構造 $y=A\sin(x+\varphi)$ を作った時に, 位相 φ は $U(0,\theta)$ に従うことが多い. ただし θ は 2π 未満の適当な正定数.

例 2. 連続型確率変数 X が 標準正規分布 *(standard normal distribution)* に従う (記号では $X\sim \mathrm{N}(0,1)$) というのは, 非負値連続関数

$$f_X(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

が X の確率密度関数であるときを言う *(*[1, 63 頁]*,* [6, 84 頁]*).*

実際, $\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx=1$ が成立する. これを見るのに, $\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx=2\int_0^{\infty}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx$ を $t=\frac{x^2}{2}$ で置換積分をする. すると $dt=xdx=\sqrt{2t}dx$, つまり, $dx=\frac{1}{\sqrt{2}}t^{-1/2}dt$ が成立する. 従って $\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx=2\int_0^{\infty}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx=\int_{-\infty}^{\infty}\frac{1}{\sqrt{\pi}}t^{\frac{1}{2}-1}e^{-t}dt=\frac{1}{\sqrt{\pi}}\Gamma(1/2)=\frac{1}{\sqrt{\pi}}\sqrt{\pi}=1$. ここで Γ は ガンマ関数 である *(Appendix B.1* を参照のこと*.)*

標準正規分布は確率論・統計学では極めて重要である.

- 統計で重要な確率分布, 例えば, 自由度 n のカイ二乗分布 *(*[1, 86 頁]*,* [6, 88 頁]*,* このプリントの演習問題*),* 自由度 n の t 分布 *(*例 *3),* が標準正規分布

から作られる ([1, 第 4 章「母集団と標本」4 節「正規分布から導かれる分布」],[6, 88 頁]).

- *2* 項分布 $B(n,p)$ に従う確率変数 X を標準化して得られる確率変数 $(X-np)/\sqrt{npq}$ は, n を大きくすると, 標準正規分布に従う確率変数に近づく *(*ドモアブル・ラプラスの定理, 中心極限定理 *(*[1, 85 頁], [6, 102 頁]*).)*.
- 推定･検定において, 標準正規分布に従う確率変数の絶対値がある値より大きくなる確率が $\alpha \in (0,1)$ ならば, その値を z_α 点と言う. 例えば, $z_{0.05} = 1.96\cdots$ など *(*[1, 65 頁], [6, 80 頁]*)*. を推定・検定で非常によく使う.

例 3. 連続型確率変数 X が *Cauchy-Lorentz* 分布 *(*[1, 88 頁], [6, 116 頁]*)* に従うとは, X の確率密度関数が適当な正定数 γ と適当な実定数 x_0 に対して $f(x) = \frac{1}{\pi}\frac{\gamma}{(x-x_0)^2+\gamma^2}$ と書ける時をいう. 特に $x_0 = 0$, $\gamma = 1$ のときの *Cauchy-Lorentz* 分布を, 自由度 *1* の t 分布 と言い, 連続型確率変数 X がこの分布に従っている時, 記号 $X \sim t_1$ で表す.

Cauchy-Lorentz 分布に従う確率変数の確率密度関数 $f(x)$ を $-\infty$ から ∞ まで積分すると *1* である. 実際 $\int_{-\infty}^{\infty} f(x)dx = \left[\frac{1}{\pi}\arctan\left(\frac{x-x_0}{\gamma}\right)\right]_{-\infty}^{\infty} = (\frac{1}{\pi}\cdot\frac{\pi}{2} - (-\frac{1}{\pi}\cdot\frac{\pi}{2})) = 1$ が成立する [3, 60 頁].

正規分布 ([1, 62 頁], [6, 84 頁]) を考える.

定義 2. 確率変数 X が, 分散 σ^2 で期待値が μ の正規分布 *(normal distribution)* に従う *(*記号. $X \sim \mathrm{N}(\mu,\sigma^2)$*)* とは, 次の関数が X の確率密度関数であるときを言う.

$$f(x) = \frac{1}{(\sqrt{2\pi})\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{4}$$

正規分布は誤差の一般的な分布である. 身長, 体重は正規分布に従うといわれている.

定理 2. 関数 $f(x)$ が確率変数 X の確率密度関数ならば, *0* でない勝手な正定数 a と勝手な実定数 b に対して $g(x) = \frac{1}{a}f(\frac{x-b}{a})$ は $aX+b$ の確率密度関数である.

Proof. $F_{aX+b}(y) = P(aX+b \le y) = P(X \le (y-b)/a) = \int_{-\infty}^{(y-b)/a} f(t)dt = \int_{-\infty}^{y}\frac{1}{a}f\left(\frac{t-b}{a}\right)dt = \int_{-\infty}^{y} g(t)dt.$ □

従って, $X \sim \mathrm{N}(\mu,\sigma^2)$ ならば, X の 標準化 (正規化) ([1, 55 頁], [6, 160 頁]) として得られる確率変数

$$\frac{X-\mu}{\sigma}$$

は $\mathrm{N}(0,1)$ に従うことが分かる. 故に,

$$P(a < X \le b) = P\left(\frac{a-\mu}{\sigma} < Z \le \frac{b-\mu}{\sigma}\right) \qquad (-\infty \le a < b \le \infty)$$

となる.

3. 確率密度関数と期待値

離散型確率変数の値の散らばりの位置の目安が期待値であったが, 連続型確率変数にも期待値を定義する ([1, 53 頁], [6, 90 頁]):

定義 3. 確率密度関数が $f_X(x)$ である連続型確率変数に対して，次の 広義積分 *(Appendix B.1* 参照のこと*)* が実数として値が定まれば，その値を X の 期待値 という．

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)dx. \tag{5}$$

さて a_i $(i = 1, 2, 3, \ldots)$ を値として取る離散型確率変数 X の期待値は，

$$E(X) = \sum_i a_i P(X = a_i) \tag{6}$$

と定義したが，式 (6) における総和記号が式 (5) における積分記号に対応し，式 (6) における a_i が式 (5) の x が対応し，式 (6) における $P(X = a_i)$ が式 (5) における確率密度関数 $f_X(x)$ に対応する．この離散型確率変数の期待値と連続型確率変数の期待値とのこの対応は，離散型確率変数 X が値 a_i を取る確率が

$$P(X = a_i) = F_X(a_i) - \lim_{x \to a_i - 0} F_X(x) \tag{7}$$

であることを注意すると，区分求積法 (Appendix A) である．

また $h : \mathbb{R} \to \mathbb{R}$ を任意の連続関数とするとき，

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

とする．

例 4 (定理 1 の続き)**.** 指数分布 $\mathrm{Exp}(\lambda)$ に従う確率変数 X の期待値は $E(X) = 1/\lambda$ である．より一般的に，$E(X^s) = \Gamma(s+1)/\lambda^s$ $(s > -1)$ である．

実際，$\int_0^\infty t\lambda e^{-\lambda t}dt = 1/\lambda$ である．というのは，$\lambda t = x$ で置換積分すると，$E(X^s) = \int_0^\infty x^s e^{-x}dx/\lambda^s = \Gamma(s+1)/\lambda^s$ である．

例 5 (例 3 の続き)**.** 連続型確率変数 X が *Cauchy-Lorentz* 分布に従うとき，X の期待値は定義されていない．X の確率密度関数は適当な正定数 γ と適当な実定数 x_0 に対して $f(x) = \frac{1}{\pi}\frac{\gamma}{(x-x_0)^2+\gamma^2}$ と書けるが，

$$\int_{-\infty}^{\infty} xf(x)dx = \left[\frac{\gamma}{2\pi}\log\left((x-x_0)^2+\gamma^2\right) + x_0\frac{1}{\pi}\arctan\left(\frac{x-x_0}{\gamma}\right)\right]_{-\infty}^{\infty} = \infty - \infty$$

となり定義されていないからである [1]．*Cauchy-Lorentz* 分布に対しては期待値ではなく 中央値 を考える．確率変数の 中央値 とは $P(X \le m) \ge 1/2$ かつ $P(X \ge m) \ge 1/2$ となる実数 m のことである．*Cauchy-Lorentz* 分布の中央値は x_0 である．

例 6 (例 2 の続き)**.** 確率変数 Z が標準正規分布 $\mathrm{N}(0,1)$ に従う時，Z の期待値 $E(Z)$ は *0* であり，$E(Z^2) = 1$, $E(Z^4) = 3$.

というのは，まず，非負整数 k に対して $E(Z^{2k+1}) = \int_{-\infty}^\infty z^{2k+1}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz$ は奇関数の広義積分だから，広義積分 $I = \int_0^\infty z^{2k+1}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz$ から I 自身を引いたものになる．広義積分 I が値として定義されていることを確認できれば，$E(Z^{2k+1}) = 0$ が従う．I を $x = z^2/2$ で置換積分すると，I はガンマ関数 $\Gamma(k+1) = \int_0^\infty x^{k+1-1}e^{-x}dx$ の定数倍であり，$k+1 > 0$ であるため値が定義されている *(Appendix B.1* を見よ*)*．従って $E(Z^{2k+1}) = 0$ になる．

---

[1] 期待値が定義されていない離散型確率変数として $P(X = n) = \frac{1}{n^2}/\zeta(2)$ $(n = 1, 2, 3, \ldots)$ がある．ここで $\zeta(s) = \sum_{n=1}^\infty \frac{1}{n^s}$ (s は 1 以外の勝手な複素数) はゼータ関数と呼ばれる現代数学で非常に重要な関数であり，$\zeta(2) = pi^2/6$ である．このため $\sum_{i=1}^\infty P(X = i) = 1$ が保証されている．また $\zeta(4) = \pi^4/90$ は統計物理でも使う．これらゼータ関数の値はフーリエ解析からも算出できる．確率・統計に基づく証明は [6] を参照されたい．

同様に, 非負整数 n に対し $E(Z^{2n}) = 2\int_0^\infty z^{2n}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz = \frac{1}{\sqrt{2\pi}}2^{n+\frac{1}{2}}\Gamma\left(n+\frac{1}{2}\right)$. $n+1/2>0$ であるため値が定義されている. ガンマ関数の性質 (16) より $E(Z^2)=1$, $E(Z^4)=3,\ldots,E(Z^{2n})=(2n-1)!!=(2n-1)(2n-3)(2n-5)\cdots$ *(n は正整数)* が成立する.

4. 確率変数の分散

確率変数 X の値の散らばりの度合いの目安として分散という定数 $V(X)$ を導入した (4/30 の授業) ([1, 54 頁], [6, 34 頁, 70 頁]):

定義 4. 確率変数 X に対して

$$V(X) = E\left((X-E(X))^2\right)$$

が定義されているならば, この定数 $V(X)$ を X の 分散 *(variance)* という.

分散 $V(X)$ が小さければ, X がその期待値 $E(X)$ から a 外れている確率がより小さくなる. しかし, 次の定理に見るように, 確率変数 X の分散 $V(X)$ は X の散らばりの位置によらない.

定理 3. *勝手な確率変数 X と勝手な定数 α,β に対して, 次の各々の等式で, 両辺の一方が定義されているならば, 他方も同じ値に定義されている.*

$$V(X) = E(X^2) - E(X)^2. \quad (\text{分散公式}) \tag{8}$$

$$V(\alpha X+\beta) = \alpha^2 V(X). \tag{9}$$

また, $V(X)$ が定義されているならば $V(X)\ge 0$.

Proof. 分散の定義に現れる, 積分 (総和) の線形性よる. □

分散公式から次が従う.

例 7. (1) 指数分布 $\mathrm{Exp}(\lambda)$ に従う確率変数の分散は $1/\lambda^2$ である *(例 4 による).*
(2) 標準正規分布 $\mathrm{N}(0,1)$ に従う確率変数 X の分散は 1 である *(例 6 による).*
(3) *Cauchy-Lorentz* 分布に従う確率変数の分散は定義されない *(例 5).*

以下, 根元事象全体を Ω で表す. 以下の形をチェビシェフの不等式と呼ぶことがある [2] [1, 99 頁].

定理 4. *X を確率変数で, 分散 $V(X)$ が定義されているものとする. 次が成立する:*

$$P\left(\{\omega\in\Omega \mid |X(\omega)-E(X)|\ge a\}\right) \le \frac{V(X)}{a^2} \qquad (a>0).$$

Proof. $A=\{\omega\in\Omega \mid |X(\omega)-\mu|\ge\varepsilon\}$ とし, 関数 $1_A:\Omega\to\{0,1\}$ を次で定義する:

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

このとき, 集合 $\{\omega\in\Omega \mid 1_A(\omega)\le a\}$ は, $a<0$ の場合は空事象, $0\le a<1$ の場合は事象 A の余事象であり, それ以外の場合は全体事象 Ω となる $(a\ge 1)$. 従って, 関数 1_A は確率変数である. また

$$P(|X-E(X)|\ge\varepsilon) = P(A) = E(1_A).$$

[2] より一般的な次の不等式をチェビシェフの不等式と呼ぶ方が多い [6, 104 頁]: “h が非負関数で $a>0$ の時 $P(h(X)\ge a)\le E(h(X))/a$.”

ここで

$$1_A \leq \frac{(X - E(X))^2}{\varepsilon^2}$$

が成立している. 従ってこの式の両辺の期待値を取ると, 分散の定義より

$$P(|X - E(X)| \geq \varepsilon) \leq E\left(\frac{(X - E(X))^2}{\varepsilon^2}\right) = \frac{V(X)}{\varepsilon^2}.$$

□

5. 多次元確率変数と同時密度関数

確率変数の和の期待値や分散をより形式的に議論するのに, 多次元確率変数を導入する.

定義 5. p 次元確率変数とは p 個の確率変数の組 $(X_1, \dots, X_p)$ で

$$F_{X_1,\dots,X_p}(x_1, \dots, x_p) = P(X_1 \leq x_1 \ \& \cdots \& \ X_p \leq x_p) \quad (x_1, \dots, x_p \in \mathbb{R})$$

となる関数が定義されているものである. この関数を 同時分布関数 という.

同時分布関数 $F_{X_1,\dots,X_p}(x_1, \dots, x_p)$ が適当な非負値関数 $f_{X_1,\dots,X_p}(x_1, \dots, x_p)$ の次の形の広義重積分 *(Appendix B.1* 参照のこと*)*

$$F_{X_1,\dots,X_p}(x_1, \dots, x_p) = \int \cdots \int_D f_{X_1,\dots,X_p}(t_1, \dots, t_p) dt_1 \cdots dt_p,$$
$$D = \{(t_1, \dots, t_p) \mid t_i \leq x_i \ (i = 1, \dots, p) \ \}$$

と書ける場合, 関数 $f_{X_1,\dots,X_p}(x_1, \dots, x_p)$ を p 次元確率変数 $(X_1, \dots, X_p)$ の 同時密度関数 *(joint probability density function)* という *(*[1, 70 頁],[6, 110 頁]*)* という.

同時密度関数 $f_{X_1,\dots,X_p} : \mathbb{R}^p \to \mathbb{R}$ が $\mathbb{R}^p$ で連続ならば, [3, 定理 5.1.2] と広義重積分の性質 (命題 1) [3, 定理 5.5.6] より, 同時分布関数は累次積分として書ける:

$$F_{X_1,\dots,X_p}(x_1, \dots, x_p) = \int_{-\infty}^{x_1} dt_1 \cdots \int_{-\infty}^{x_p} dt_p \ f_{X_1,\dots,X_p}(t_1, \dots, t_p) \tag{10}$$

が成立する.

同時密度関数をある引数たちを除いて他の全ての引数たちを全空間で積分すると, 除外した変数についての同時密度関数になる ([1, 71 頁], [2, 77 頁]):

定理 5. 多次元確率変数 $(X_1, \dots, X_p)$ の同時密度関数 $f_{X_1,\dots,X_p}(t_1, \dots, t_p)$ に対して, $1 \leq i_1 < i_2 < \dots < i_q \leq p$, $1 \leq j_1 < j_2 < \cdots < j_r \leq p$, $\{i_1, \dots, i_q\} \cap \{j_1, \dots, j_r\} = \emptyset$, かつ $\{i_1, \dots, i_q\} \cup \{j_1, \dots, j_r\} = \{1, \dots, p\}$ とする. このとき広義積分

$$\int_{-\infty}^{\infty} dt_{i_1} \int_{-\infty}^{\infty} dt_{i_2} \cdots \int_{-\infty}^{\infty} dt_{i_q} \ f_{X_1,\dots,X_p}(t_1, \dots, t_p) \tag{11}$$

は r 次元確率変数 $(X_{j_1}, X_{j_2}, \dots, X_{j_r})$ の同時密度関数になる.

同時密度関数から得られることを強調するために, (11) を $(X_{j_1}, X_{j_2}, \dots, X_{j_r})$ の 周辺確率密度関数 ともいう.

Proof. 同時密度関数 $f_{X_1,\dots,X_p}(t_1, \dots, t_p)$ が連続である場合のみを証明する. その他の場合は煩雑であるため省略する. この場合, 積分 (10) は累次積分の順番によらないので, $P(X_1 \leq x_1 \ \& \cdots \& \ X_p \leq x_p)$ は

$$= \int_{-\infty}^{x_{j_1}} dt_{j_1} \cdots \int_{-\infty}^{x_{j_r}} dt_{j_r} \ \int_{-\infty}^{x_{i_1}} dt_{i_1} \cdots \int_{-\infty}^{x_{i_q}} dt_{i_q} \ f_{X_1,\dots,X_p}(t_1, \dots, t_p)$$

従って, 確率の連続性などから $F_{X_{j_1},\dots,X_{j_r}}(x_{j_1},\dots,x_{j_r})$ は

$$= \lim_{x_{i_1}\to\infty,\dots,x_{i_q}\to\infty} P(X_1 \le x_1 \ \&\cdots\&\ X_p \le x_p) = \int_{-\infty}^{x_{j_1}} dt_{j_1} \cdots \int_{-\infty}^{x_{j_r}} dt_{j_r} \ (\text{式 } (11))$$

□

勝手な連続関数 $h:\mathbb{R}^p \to \mathbb{R}$ に対して

$$E(h(X_1,\dots,X_p)) = \int\cdots\int_{\mathbb{R}^p} h(x_1,\dots,x_p) f_{X_1,\dots,X_p}(x_1,\dots,x_p) dx_1\cdots dx_p$$

と定義する.

次は重要である ([1, 定理 3.1, 定理 3.6], [2, 定理 4.2, 定理 6.5]):

系 1 (期待値の線形性)**.** 勝手な実数 $\alpha_1,\dots,\alpha_n,\gamma$ と勝手な確率変数 $X_1,\dots,X_n$ に対して, 次の等式の一方が定義されていれば他方も同じ値に定義されている:

$$E(\alpha_1X_1+\alpha_2X_2+\cdots+\alpha_nX_n+\gamma) = \alpha_1E(X_1)+\alpha_2E(X_2)+\cdots+\alpha_nE(X_n)+\gamma.$$

Proof. $n=2$ の場合をやる. その他の場合は同様に証明できる. X,Y が連続型確率変数とし, 同時密度関数を $f(x,y)$, それぞれの確率密度関数を $f_X(x)$, $f_Y(y)$ とする. f, f_X, f_Y が連続である場合のみを示す. 定理 5 を使う.

$$\begin{aligned}
&E(\alpha X+\beta Y+\gamma)\\
&= \int\int_{\mathbb{R}^2} (\alpha x+\beta y+\gamma) f_{X,Y}(x,y)dxdy\\
&= \alpha\int\int_{\mathbb{R}^2} x f_{X,Y}(x,y)dxdy + \beta\int\int_{\mathbb{R}^2} y f_{X,Y}(x,y)dxdy + \gamma\int\int_{\mathbb{R}^2} f_{X,Y}(x,y)dxdy
\end{aligned}$$

$xf_{X,Y}(x,y)$ と $yf_{X,Y}(x,y)$ は $\mathbb{R}^2$ で連続だから, それらの 重積分を累次積分 にできるため

$$= \alpha\int_{-\infty}^{\infty} x\left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)dy\right)dx + \beta\int_{-\infty}^{\infty} y\left(\int_{-\infty}^{\infty} f_{X,Y}(x,y)dx\right)dy + \gamma$$

二個の内側の積分は周辺確率密度関数だから(定理 5),

$$= \alpha\int_{-\infty}^{\infty} x f_X(x)dx + \beta\int_{-\infty}^{\infty} y f_Y(y)dy + \gamma = \alpha E(X)+\beta E(Y)+\gamma.$$

X,Y が離散型確率変数の場合は, プリント「離散型確率変数の期待値と分散, 二項分布, 推定」(4/23)) で行った. □

確率変数 $X_1,\dots,X_p$ が独立というのは,

$$\begin{aligned}
&P(X_{i_1}\le a_{i_1}\ \&\cdots\&\ X_{i_q}\le a_{i_q}) = P(X_{i_1}\le a_{i_1})\cdots P(X_{i_q}\le a_{i_q})\\
&(1\le i_1<\cdots<i_q\le p)\quad (a_{i_1},\dots,a_{i_q}\in\mathbb{R})
\end{aligned}$$

が成立することであった (プリント「離散型確率変数の期待値と分散, 二項分布, 推定」(4/23)).

定理 6 (連続型確率変数の独立性の必要十分条件)**.** 関数 $f_{X_1,\dots,X_p}(x_1,\dots,x_p)$ が p 次元連続型確率変数 $(X_1,\dots,X_p)$ の同時密度関数で, 関数 f_{X_i} が連続型確率変数 X_i の確率密度関数とする. このとき, $X_1,\dots,X_p$ が独立であるための必要十分条件は

$$f_{X_1}(x_1)\cdots f_{X_p}(x_p) = f_{X_1,\dots,X_p}(x_1,\dots,x_p) \qquad (x_1,\dots,x_p\in\mathbb{R}) \tag{12}$$

が成立することである.

Proof. $p=2$ の場合のみを示す．その他の場合は同様．十分性を証明する．(12) が成立すると仮定する．このとき，

$$P(X \le a \,\&\, Y \le c) = \int\int_D f_{X,Y}(x,y)dxdy,$$
$$D = \{(x,y) \mid x \le a\, y \le c\}$$

が成立する．仮定と広義重積分に関する命題 1 ([3, 定理 5.5.7]) より，

$$\int\int_D f_X(x) f_Y(y) dxdy = \int_{-\infty}^{a} f_X(x)dx \int_{-\infty}^{c} f_Y(y)dy.$$

右辺は, $P(X \le a)P(Y \le c)$.

必要性の証明: $f_{X,Y}(x,y), f_X(x), f_Y(y)$ はそれぞれ $(x,y), x, y$ の近傍で連続である場合のみをここでは示す．そうでない場合は証明が煩雑であるので省略する．

仮定より, h の関数

$$g(h) = P(x < X \le x+h) = \int_x^{x+h} f_X(s)ds$$

は, 連続な導関数 $g'(h) = f_X(x+h)$ を持つ．$g(h)$ を漸近展開 [3, 定理 2.4.5] すると, $g(h) = g(0) + g'(0)h + o(h) = 0 + f_X(x)h + o(h)$ $(h \to 0)$．ただし, $o(h)$ はランダウの式 [3, 49 頁] で, $\lim_{h\to 0} u(h)/h = 0$ であるとき $u(h) = o(h)$ と書くための評価式である．故に

$$P(x < X \le x+h) = f_X(x)h + o(h), \quad (h \to 0) \tag{13}$$

が成立する．同様に

$$P(y < Y \le y+k) = f_Y(y)k + o(k), \quad (k \to 0)$$

と書ける．X, Y が独立とすると, $P(x < X \le x+h, y < Y \le y+k) = P(x < X \le x+h)P(y < Y \le y+k)$ が成立する．従って,

$$\int_x^{x+h} ds \int_y^{y+k} f_{X,Y}(s,t)dt = (f_X(x)h + o(h))(f_Y(y)k + o(k)).$$

$h, k \ne 0$ より両辺を hk で割り算すれば

$$\frac{1}{h}\int_x^{x+h} ds \frac{1}{k}\int_y^{y+k} f_{X,Y}(s,t)dt = \left(f_X(x) + \frac{o(h)}{h}\right)\left(f_Y(y) + \frac{o(k)}{k}\right).$$

$h \to 0$ とすると, 左辺の極限は, 累次積分 $\int_x^u ds \frac{1}{k}\int_y^{y+k} f_{X,Y}(s,t)dt$ の u に関する偏微分の $u = x$ での値になる．従って

$$\frac{1}{k}\int_y^{y+k} f_{X,Y}(x,t)dt = f_X(x)\left(f_Y(y) + \frac{o(k)}{k}\right).$$

$k \to 0$ とすると, 同様の議論により結論 (12) が従う． □

次に [1, 定理 3.5] の一般化である [2, 定理 6.4] を見る．この定理は, 中心極限定理を証明するときにも使う．

定理 7 (独立確率変数の積の期待値の公式)**.** $X_1, \ldots, X_p$ を独立な確率変数で各 $h_i(x)$ は実数値連続関数とする．次の等式の両辺の一方が定義されているならば他方も同じ値で定義されている．

$$E(h_1(X_1)\cdots h_p(X_p)) = E(h_1(X_1))\cdots E(h_p(X_p)) \tag{14}$$

Proof. $p=2$ の場合のみを証明する．その他の場合は同様にできる．X_1, X_2 が離散型の場合は, プリント「離散型確率変数の期待値と分散, 二項分布, 推定」(4/23) で行った．

X_1, X_2 が連続型の場合を証明する．$E(h_1(X)h_2(Y)) = \int\int_{\mathbb{R}^2} h_1(x)h_2(y)f_{X,Y}(x,y)dxdy$. X と Y が独立だから, (12) が成立する．従って,

$$E(h_1(X)h_2(Y)) = \int\int_{\mathbb{R}^2} h_1(x)h_2(y)f_X(x)f_Y(y)dxdy$$

被積分関数は連続である．広義重積分に関する命題 1 ([3, 定理 5.5.7]) より

$$= \int_{-\infty}^{\infty} h_1(x)f_X(x)dx \int_{-\infty}^{\infty} h_2(y)f_Y(y)dy = E(h_1(X))E(h_2(Y)).$$

□

二つの確率変数の関わりの強さを表すものとして 共分散 [1, 71 頁] を導入する．

定義 6. 確率変数の 共分散 とは,

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

である．

系 2. 確率変数 X_1, X_2 が独立で, $\mathrm{Cov}(X_1, X_2)$ が定義されているならば, $\mathrm{Cov}(X_1, X_2) = 0$ が成立する．

定理 8 (独立確率変数に対する分散の加法性)**.** 確率変数 $X_1, X_2, \ldots, X_n$ が与えられたときに, 次の等式の一方が定義されていれば他方も定義されている．

$$V(X_1+\cdots+X_n) = V(X_1)+\cdots+V(X_n) \qquad (\mathrm{Cov}(X_i, X_j) = 0 \ (i \neq j))$$

が成立する．

Proof. $n=2$ のときを証明する．その他の場合は同様に証明できる．分散公式と期待値の線形性 (系 1) より, $V(X+Y) = E((X+Y)^2) - (E(X+Y))^2 = E(X^2 + 2XY + Y^2) - (E(X)+E(Y))^2 = (E(X^2)-E(X)^2) + (E(Y^2)-E(Y)^2) + 2E(XY) - 2E(X)E(Y) = V(X)+V(Y)+2(E(XY)-E(X)E(Y))$. 仮定より, $V(X+Y) = V(X)+V(Y)$ が従う． □

二項分布 $B(n,p)$ に従う期待値と分散がそれぞれ np, $np(1-p)$ であることを, 前回, 系 1, 定理 8, および例 6 から導いた．この方法を連続型変数に対して練習する演習問題を最後に添付した．

6. 確率変数の未知の期待値を反復独立試行により「推定」する

確率変数たちが同一の分布に従うというのは, それらの分布関数が同じであるときを言う．

次の定理は 大数の弱法則 ([1, 84 頁], [6, 100 頁]) という．

定理 9. 同一の分布に従う連続型確率変数 X_i $(i = 1, 2, \ldots, n)$ が, $i \neq j$ ならば $\mathrm{Cov}(X_i, X_j) = 0$ と仮定する．また X_1 は期待値と分散を持つものとする．確率変数 $X_1, \ldots, X_n$ の算術平均を

$$\overline{X}_n := \frac{X_1+\cdots+X_n}{n}$$

とおく．このとき正の実数 ε に対して,

$$\lim_{n\to\infty} P\left(\left\{\omega \in \Omega \ : \ \left|\overline{X}_n(\omega) - E(X_1)\right| \geq \varepsilon\right\}\right) = 0.$$

Proof. 系 (1) より $E\left(\overline{X}_n\right) = E(X_1)$. 定理 3 および定理 8 より, $V\left(\overline{X}_n\right) = \frac{V(X_1)}{n}$. これとチェビシェフの不等式により,

$$0 \le P\left(\left\{\omega \in \Omega \ : \ \left|\overline{X}_n(\omega) - E(X_1)\right| \ge \varepsilon\right\}\right) \le \frac{V(X_1)}{n\varepsilon^2}$$

が成立し, ここから題意が従う. □

現象を仕切り直しながら繰り返し同じ観測してその算術平均を取れば, 繰り返し回数が大きくなるに従って, 算術平均はその現象の期待値に収束する.

次の定理は 大数の強法則 ([6, 100 頁]) という. 大数の強法則の結論から大数の弱法則の結論が従うためである.

定理 10. 同一の分布に従う 独立な 確率変数の列 X_i $(i = 1, 2, \ldots, n)$ が与えられ X_1 は期待値を持つものとする. 確率変数 $X_1, \ldots, X_n$ の算術平均を

$$\overline{X}_n := \frac{X_1 + \cdots + X_n}{n}$$

とおく. すると

$$P\left(\left\{\omega \in \Omega \ : \ \lim_{n\to\infty} \overline{X}_n(\omega) = E(X_1)\right\}\right) = 1.$$

プリント「離散型確率変数の期待値と分散, 二項分布, 推定」(4/23) の議論により,

定理 11. 二項分布 $B(n, p)$ に従う確率変数 B_n に対して

$$P\left(\left\{\omega \in \Omega \ : \ \lim_{n\to\infty} \frac{1}{n} B_n(\omega) = p\right\}\right) = 1.$$

同様の議論を今回の演習で行う.

定義 7. 同一の分布に従う 独立な 確率変数の列 X_i $(i = 1, 2, \ldots, n)$ を 標本 という.

$$\overline{X}_n := \frac{X_1 + \cdots + X_n}{n}$$

を 標本平均 という.

大数の強法則は, 標本平均は母平均 (この後の講義で説明する) の一致推定量 (この後の講義で説明する) であることを主張している.

次回は中心極限定理を Taylor 展開を用いて証明する.

References

[1] 稲垣宣生, 山根芳和, 吉田光雄. 統計学入門. 裳華房, 第 22 版, 2012.
[2] 穴井克則. 講義: 確率・統計. 学術図書出版社, 第 2 版, 2012.
[3] 三宅敏恒. 入門微分積分. 培風館, 1992.
[4] 杉浦光夫. 解析入門 I. 東京大学出版会, 1980.
[5] 高木貞治. 解析概論. 岩波書店, 改訂第 3 版, 1961.
[6] 藤田岳彦. 大学生の確率・統計. 東京図書, 2010.

Appendix A. 区分求積法と期待値

定積分

$$\int_c^d x f_X(x)dx$$

は, 区間 $[c,d]$ の分割を用いて定義できる (区分求積法 [3, 3.4 節]). いいかえれば定積分はリーマン和の極限である: 閉区間 $[a,c]$ の勝手な分割

$$\Delta : c = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = d$$

と, 勝手な実数 ξ_i $(x_i \leq \xi_i \leq x_{i+1};\ 0 \leq i \leq n)$ に対して, 極限

$$\lim_{\min_i (x_{i+1}-x_i)\to 0} \sum_{i=0}^{n} \xi_i f_X(\xi_i)(x_{i+1}-x_i) \tag{15}$$

が定義されて, その値が分割 Δ と $(\xi_i)_i$ によらない時, その極限値を $\int_c^d x f_X(x)dx$ と書いた.

ここで, f_X は非負値 $\underline{\text{連続}}$ 関数とする. すると, 分布関数 $F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(t)dt$ は連続かつ微分可能である. 平均値の定理 [4] から, 適当な実数 ξ_i $(x_i < \xi_i < x_{i+1})$ があって,

$$F'_X(\xi_i) = \frac{F_X(x_{i+1}) - F_X(x_i)}{x_{i+1}-x_i}.$$

が成立する. f_X は非負値連続関数であるため, 分布関数 $F_X(x)$ の導関数は確率密度関数 $f_X(x)$ である. このことと, 分布関数の定義から

$$f_X(\xi_i) = \frac{P(x_i < X \leq x_{i+1})}{x_{i+1}-x_i}$$

が成立する. 平均値の定理で見つけた ξ_i をリーマン和 (15) の中の ξ_i として取ると,

$$\int_c^d x f_X(x)dx = \lim_{\min_i (x_{i+1}-x_i)\to 0} \sum_{i=0}^{n} \xi_i P(x_i < X \leq x_{i+1}).$$

一方, 離散型確率変数 X の期待値 (6) は

$$E(X) = \lim_{\min_i (x_{i+1}-x_i)\to 0} \sum_{i=0}^{n} \xi_i P(x_i < X \leq x_{i+1}). \qquad (x_i < \xi_i < x_{i+1})$$

と書ける. というのは, 式 (7) が成立するからである.

以上, 離散型確率変数の期待値と連続型確率変数の期待値の対応を分布関数を用いて説明した. この対応は分布関数についてのリーマン・スティルチェス積分 [4, 17 節] により厳密化できる. リーマン・スティルチェス積分はリーマン積分の一般化である. 確率変数の期待値を分布関数についてのリーマン・スティルチェス積分としても定義できる.

Appendix B. 広義積分

以下は, [3] に基づく. 関数 $f(x)$ は区間 $[a,b)$ (b は実数または ∞) で連続関数とする. 極限 $\lim_{\beta\to b-0} \int_a^\beta f(x)dx$ が収束する時, $f(x)$ は区間 $[a,b)$ で積分可能といい (積分が存在する, 積分が収束するともいう)

$$\int_a^b f(x)dx = \lim_{\beta\to b-0} \int_a^\beta f(x)dx$$

とおく．区間 $(a,b]$ (a は実数または $-\infty$) における積分についても，全く同様に極限を用いて定義される．

$f(x)$ が開区間 (a,b) で連続であるとき，$f(x)$ が (a,b) において積分可能であるとは，a と b の間に適当に点 c を取ったときに，$f(x)$ が区間 $(a,c]$ および $[c,b)$ のいずれにおいても積分可能であるときに言い

$$\int_a^c f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$$

とおく．この定義が c の取り方によらないことは明らかである．

関数 $f(x)$ が区間 (a,b) において，有限個の点 $c_1 < c_2 < \cdots < c_k$ を除いて連続とする．このとき $\int_a^b f(x)dx$ が存在するとは，(a,b) を $c_1, c_2, \ldots, c_k$ で分割して得られる，有限個の区間における広義積分が全て収束するときに言い

$$\int_a^b f(x)dx = \int_a^{c_1} f(x)dx + \int_{c_1}^{c_2} f(x)dx + \cdots + \int_{c_k}^b f(x)dx$$

とおく．

例 8. ガンマ関数 *([6, 76 頁], [3, 70 頁, 134 頁]).*

$$\Gamma(s) = \int_0^\infty x^{s-1}e^{-x}dx. \quad (s>0)$$

$0<s<1$ の範囲では，被積分関数 $x^{s-1}e^{-x}$ は $x \to 0+$ で発散するが，ガンマ関数は $s>0$ で定義されていて，次が成立する．

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}. \qquad \Gamma(s+1) = s\,\Gamma(s),\ (s>0). \qquad \Gamma(1)=1. \tag{16}$$

B.1. **広義重積分.** D は次の性質 (#) を持つ有界閉領域の列 $\{D_n\}_{n=1}^\infty$ が存在するようなものとする:

$$D_1 \subset D_2 \subset D_3 \subset \cdots \subset D = \bigcup_{n=1}^\infty D_n$$

D に含まれる有界閉領域は，いずれかの D_n に含まれる．

$\mathbb{R}^p$ の有界閉集合のどんな列 $D_1 \subset D_2 \subset \cdots \subset D$ に対しても，重積分 $\int\int_{D_i} f(x,y)dxdy$ が，領域の列によらない値に収束するときに，その値をこの広義積分 $\int\int_D f(x,y)dxdy$ で表す [5, 94 節]．実を言うと，$f_{X,Y}$ が非負値であるため，(#) を満たす列 $D_1 \subset D_2 \subset D_3 \subset \cdots$ が一つでもあればよい [3, 定理 5.5.6]．

次の命題 [3, 定理 5.5.7] が成立する:

命題 1. *$\int_a^\infty |f(x)|dx$, $\int_c^\infty |g(y)|dy$ が存在するならば，$f(x)g(y)$ は $D=\{(x,y) \mid a \le x < \infty, c \le y < \infty\}$ で積分可能で*

$$\int\int_D f(x)g(y)dxdy = \left(\int_a^\infty f(x)dx\right)\left(\int_c^\infty g(y)dy\right).$$

区間 $[a,\infty)$, $[c,\infty)$ 以外にも，全く同様に区間 $(-\infty,a]$, $(-\infty,c]$ についても成立する．

数理統計学 第5回演習 (2013/5/13) 所属　　　　　　　　氏名

成績に入らないので, 自由に話し合いながら解いて下さい. 質問も自由に書いて下さい.

(1) $n(\geq 1)$ 個の独立な確率変数 $Z_1, \ldots, Z_n$ が標準正規分布に従う時, 確率変数 $Z_1^2, \ldots, Z_n^2$ が独立であることを用いて, 確率変数 $Y = Z_1^2 + \cdots + Z_n^2$ の期待値と分散を求めよ (ヒント. 系 1, 定理 8, および例 6). このとき Y は自由度 n のカイ二乗分布に従う といい, このことを記号 $Y \sim \chi_n^2$ で表す. カイ二乗分布は, 分散の推定・検定や, データの適合度の検定 ([1, 141 頁], [2, 12 章], [6, 152 頁]) で極めて重要である.

(2) 自由度 n のカイ二乗分布に従う確率変数 Y に対して,

$$P\left(\left\{\omega \in \Omega \mid \lim_{n\to\infty} \frac{Y(\omega)}{n} = 1\right\}\right) = 1$$

が成立することを, 大数の強法則を用いて証明せよ. なおこの事実を使うと, 自由度 n が大きい時 t_n 分布が標準正規分布 $\mathrm{N}(0,1)$ で近似できる [1, 89 頁], ことが証明できる.

解. (1) おのおのの確率変数 Z_i は標準正規分布に従うから $E(Z_i) = 0$ かつ $V(Z_i) = 1$ である. 従って $E(Z_i^2) = V(Z_i) + (E(Z_i))^2 = 1$ である. 期待値の線形性から $E(Y) = E(Z_1^2) + \cdots + E(Z_n^2) = n$. 仮定から $Z_1^2, \ldots, Z_n^2$ は独立である.

例 6 より $E(Z_i^4) = 3!! = 3$. 従って $V(Z_i^2) = E(Z_i^4) - (E(Z_i^2))^2 = 3 - 1^2 = 2$. $Z_1^2, \ldots, Z_n^2$ が独立だから分散の性質から $V(Y) = V(Z_1^2) + \cdots + V(Z_n^2) = 2n$.

(2) $Z_1, \ldots, Z_n$ が標準正規分布に従う独立な確率変数たちであるため, $Z_1^2, \ldots, Z_n^2$ も独立な確率変数たちである. ここで Z_1^2 はより正確には, 全体事象 Ω の各要素 ω に対して, $(Z_1(\omega))^2$ を割り当てる確率変数である. 大数の強法則 (定理 10) より

$$P\left(\left\{\omega \in \Omega \mid \lim_{n\to\infty} \frac{Z_1^2(\omega) + \cdots + Z_n^2(\omega)}{n} = E(Z_1^2)\right\}\right) = 1$$

である. $E(Z_1^1) = V(Z_1) + E(Z_1)^2 = 1 + 0 = 1$ であるから,

$$P\left(\left\{\omega \in \Omega \mid \lim_{n\to\infty} \frac{Z_1^2(\omega) + \cdots + Z_n^2(\omega)}{n} = 1\right\}\right) = 1$$

$Y = Z_1^2 + \cdots + Z_n^2$ であるから

$$P\left(\left\{\omega \in \Omega \mid \lim_{n\to\infty} \frac{Y(\omega)}{n} = 1\right\}\right) = 1$$

である.