1. 離散確率変数の復習とその積率母関数

2025 後期 3H606 数理科学入門（統計系）

Oct 7, 2025

数理環境論 Room 742
Email: ZHOU YI <zhouy@people.kobe-u.ac.jp>

講義情報

成績評価方法

- レポート (50%) と最終試験 (50%) の合計点を基準とし，必要に応じて調整する
- 最終試験: **対面**

授業の概要 (内容は進捗に応じて調整される場合がある)

- 上半期: 確率論 (単変量・２変量の分布など)
- 下半期: 数理統計 (標本データと推定量の性質など)

教科書

- 数理統計の基礎 (阪本雄二, 2024.4)

参考資料

- 現在数理統計学の基礎 (久保川達也)
- Statistical Inference(George Casella, Roger L. Berger)

スライド

Def 1.0　定義　definition, Def

「〜とは」論理的に議論するために，対象や概念を厳密に規定すること

Thm 1.0　定理　theorem, Th

「〜ならば」仮定 (前提条件) が満たされれば必ず結論が成り立つことを示した命題. (証明が与えられること)

Ex 1.0　例　example, Ex

Supp 1.0　補足資料 supplementary notes, Supp

課題 1.0　最後のスライド

提出期限: 次の火曜日 1:00AM

期限を過ぎた場合: メールでレポートを提出. ただし -10 点減点

- 教科書の記述と完全に一致しない部分がある．対応するページ番号を pp.xx-xx と付記している．教科書をお持ちでない方は無視してください．
- **(証)**: 重要な証明・導出

キーワード

確率空間

確率変数

離散確率変数
確率 (質量) 関数　分布関数[1]

期待値　分散　標準偏差
積率 積率母関数 代表的な分布の積率母関数

[1]Web application:
https://alain003.phs.osaka-u.ac.jp/mephas_web/1_2MFSdisdist_jp/

確率論 vs 統計学

確率論-理論数学

- 理論 (母集団 pp.1-2,73・モデル) から結果を予測[2]

統計学-応用数学

- 結果 (標本データ pp.1-2,73) から理論 (母集団・モデル) を推測[3]

Ex 1.1　確率・統計問題

1. 人体の体温は平均 37.0 度，標準偏差 0.8 の正規分布に従う．100 人の標本平均が 36.5 度未満となる確率を求めよ
2. A 大学の学生の 10% が左利きであるとする．35 人のクラスで少なくとも 5 人が左利きとなる確率を求めよ
3. 健康な人 100 人の無作為標本における標本平均体温は 36.5 度であった．健康な人の母平均体温が 37.0 度未満であるといえる根拠はあるか
4. A 大学の学生 35 人のクラスで 5 人が左利きであった．学生全体における左利きの割合を推定せよ

[2] 予測: これからどうなるかを当てる
[3] 推測: データの背後にあるものを理解する

試行・標本空間・事象 pp.22

試行 trial

- 偶然に結果が決まる観察や操作

Ex 1.2

試行 A: Trial^A = コインを 1 回投げること
試行 B: Trial^B = サイコロを 1 回振ること
試行 C: Trial^C = コインを 3 回投げること (Trial^A $\times\ 3$)

標本空間 sample space

- 試行で起こり得るすべての結果の集合 $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$(全事象)
- ω_i: 標本点

Ex 1.3

(続き)
Trial^A の標本空間: $\Omega^A = \{\omega_1 = \text{Head}, \omega_2 = \text{Tail}\} = \{H, T\}$
Trial^A の標本空間 *: $\Omega^{A*} = \{\omega_1 = H, \omega_2 = T, \omega_3 = \text{Stand-up}\} = \{H, T, S\}$
同じ試行でも、標本空間の取り方が異なれば確率の定義は変わる
Trial^B の標本空間: $\Omega = \{\omega_1 = 1, \omega_2 = 2, \ldots, \omega_6 = 6\} = \{1, 2, 3, 4, 5, 6\}$

試行・標本空間・事象 pp.22

事象 event

- 標本空間 Ω の部分集合
  - 空事象: $\varnothing$，余事象 (補集合):[4] A^c
  - 二つ (以上) の事象: 積事象 $A \cap B$，和事象 $A \cup B$，排反 $A \cap B = \varnothing$

Ex 1.4

(続き)
Trial^A の標本空間 Ω^A における事象の例

- $A = \{T\}$: T が出る; $A^c = \{H\}$: H が出る
- $\varnothing$: H/T が出ないこと (コインが消えた・立つ); Ω: H/T が出る

Trial^A の標本空間 Ω^{A*} における事象の例

- $A_1 = \{T\}$: T が出る; $A_2 = \{H\}$: H が出る; $A_3 = \{H, T\}$: コインが立たない

[4] 事象: 実際の偶然の出来事 (ただし数学的には集合で表される)

確率空間

試行: 偶然性 (不確からしさ) を伴う実験
確率: 試行の結果に関する「確からしさ」を数学的に表したもの

Def 1.1 可測集合族 $\mathcal{B}$ (σ-代数, sigma-algebra)

標本空間 Ω の部分集合族, s.t.[5]

1. $\Omega \in \mathcal{B}, \varnothing \in \mathcal{B}$
2. $A \in \mathcal{B} \ \Rightarrow \ A^c \in \mathcal{B}$
3. $A_1, A_2, \cdots \in \mathcal{B} \ \Rightarrow \ \bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$

- 集合 $=$ 要素の集まり
- 集合族 $=$ 集合の集まり
- 可測集合族 $\mathcal{B} =$ 事象の集合
- $\mathcal{B}$ の元 (可測集合)$=$ 事象

[5] s.t.:such that, 次の性質・条件などを満たす

Ex 1.5　2 枚のコイン投げる試行の可測集合族

標本空間: $\Omega = \{TT, HH, TH, HT\} = \{\omega_1, \omega_2, \omega_3, \omega_4\}$

- $\mathcal{B}_0 = \{\emptyset, \Omega\}$;
- $\mathcal{B}_1 = \{\emptyset, \Omega, \{TT, HH\}, \{TH, HT\}\}$
- $\mathcal{B}_m = \{\emptyset, \Omega, \omega_1, \omega_2, \omega_3, \omega_4, \{\omega_1, \omega_2\}, \dots,$(すべての組み合わせ)$, \dots\}$

確率空間

Def 1.2　確率 probabilitypp.21 **(コルモゴロフの公理 Kolmogorov Axioms)**

可測集合 A $(A \in \mathcal{B})$ に対して実数を対応させる関数 $P = P(.)$ で，次の 3 つの性質を満たすもの

1. 非負性: $\forall$[6] $A \in \mathcal{B}, P(A) \geq 0$
2. 正規化: $P(\Omega) = 1$
3. 可算加法性: $A_1, A_2, \ldots,$ が互いに排反であるとき，$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

- 確率は $\Omega, \mathcal{B}, P$ の 3 つの要素により決まる
- 2. を満たさない関数 $=$ 測度 measure，確率は測度の一つ

[6] $\forall$ ○○: for all, 任意の・すべての○○に対して

確率空間

Def 1.3　確率空間 probability space

$(\Omega, \mathcal{B}, P)$

- Ω: 標本空間 (すべての結果の集合)
- $\mathcal{B}$：Ω の部分集合族 (特例: すべての事象の集合)[7]
- P：確率測度

Ex 1.6　1 回コイン投げの確率 P

$P(H) = P(T) = \frac{1}{2}$ とする

1. 標本空間:　$\Omega = \{H, T\} \Rightarrow P(\{H\} \cup \{T\}) = 1$
2. 可測集合 $A = \{H\}, A = \{T\}, A = \{\}, A = \{H, T\}$ のとき，$\forall A \in \mathcal{B}, P(A) \geq 0$
3. $A_1 = \{H\}, A_2 = \{T\}$ が互いに排反であり，$P(A_1 \cup A_2) = P(A_1) + P(A_2)$

[7] は必ずしも「すべての部分集合」を含むわけではなく、その一部でもよい

確率空間

確率空間 vs 宇宙

- 宇宙 (Universe) = 標本空間 Ω
- 惑星 (Planet) = 標本点 ω_i
- 銀河の集まり (Galaxy set) = σ-代数 $\mathcal{B}$
- 銀河系の質量の割合関数 = 確率測度 P

確率空間 $(\Omega, \mathcal{B}, P)$ = 宇宙 + 銀河の集まり + 質の関数

確率空間:「結果・事象・確率」を明確に区別する

- 標本空間: 試行のすべての「結果」を整理する
- σ-代数: 確率測度の公理（非負・正規化・可算加法性）を守る
- 確率測度:「事象」の確からしさを測る

確率変数

Def 1.4　確率変数 random variable[pp.23]

Ω を全事象，$\mathcal{B}$ を Ω の可測集合族，P を $(\Omega, \mathcal{B})$ 上の確率とするとき標本点 $\omega \in \Omega$ に対して実数値 $X(\omega) \in \mathbb{R}$ を対応される関数 $X = X(.)$

確率変数: 確率 P が定義されている集合 Ω から実数への写像 X

- 確率変数 X の標本区間 (実現値の全体): $\mathcal{X} = \{X(\omega) \mid \omega \in \Omega\}$
- 1 つの $\omega_i \in \Omega$ に対して $X(\omega_i) = x_i$ なる X の値が定まる
- x: 実現値 (小文字，$_i$ を省略)，X: 確率変数 (大文字)
- $\forall x \in \mathbb{R}, X \leq x$ である確率: $P(X \leq x) = P(\{\omega \in \Omega \mid X(\omega) \leq x\})$

Supp 1.1　実数 $\mathbb{R}$ の例

- 整数: $-3, -2, -1, 0, 1, 2, 3, \ldots$
- 分数: $\frac{1}{2}$, $-\frac{3}{4}$, $\frac{22}{7}$
- 循環小数 (有理数): $0.333\cdots = \frac{1}{3}$
- 無理数 (循環しない小数): $\sqrt{2}$, π, e

確率変数

Ex 1.7　3 回コイン投げる試行

- 標本空間
  $\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\} = \{\omega_1, \ldots, \omega_8\}$
- 確率変数: $X =$ Hが出る個数
  $X(\Omega) = \{X(\{HHH\}), X(\{HHT\}), \ldots, X(\{TTT\})\} = \{x_1, \ldots, x_8\} = \{3, 2, \ldots, 0\}$
- X の標本区間
  $\mathcal{X} = \{0, 1, 2, 3\}$
- $X \leq 1$ の確率
  $P(X \leq 1) = P(\{TTT, HTT, THT, TTH\})$

i	1	2	3	4	5	6	7	8
ω_i	HHH	HHT	HTH	THH	TTH	THT	HTT	TTT
$X(\omega_i) = x_i$	3	2	2	2	1	1	1	0

確率変数: X^* =3 つのHが出ること，X^* の標本区間: $\mathcal{X}^* = \{0, 1\}$

離散確率変数

Def 1.5 離散確率変数 (離散変数)pp.24

数列 $\{x_k\}$ 上にだけ値をとる確率変数 X，s.t.

$$\sum_{k=0}^{\infty} P(X = x_k) = P(X = x_0) + P(X = x_1) + \cdots = 1$$

Ex 1.8 (Ex 1.7)

Supp 1.2 数列

- 数列: 自然数 $n = 1, 2, 3, \ldots$ に対し実数 a_n を対応させる写像
- 表記: $\{a_n\}_{n=1}^{\infty}$ $(a_1, a_2, a_3, \ldots)$
- 例: 等差数列 $a_n = n$，等比数列 $a_n = 2^n$

離散変数の分布

確率変数は変動するため、その分布 (結果が出るの可能性) は重要な情報

Def 1.6　確率質量関数 (確率関数) probability mass function, PMF[pp.24]

離散確率変数 X に対して

$$f_X(x) = \begin{cases} p_k = P(X = x_k) & x_k \in \mathcal{X} \\ 0 & x_k \notin \mathcal{X} \end{cases} \quad \text{s.t.} \sum_{x_k \in \mathcal{X}} p_k = 1, p_k \geq 0$$

Ex 1.9　(Ex 1.7)

$k = 0, 1, 2, 3 \Rightarrow x_0 = 0, x_1 = 1, x_2 = 2, x_3 = 3$ とする

$$p_0 = \frac{1}{8}, \quad p_1 = \frac{3}{8}, \quad p_2 = \frac{3}{8}, \quad p_3 = \frac{1}{8} \Rightarrow \sum_{x_k \in \mathcal{X}} p_k = \sum_{k=0}^{3} p_k = 1$$

x_k	具体的な事象	確率 $p_k = P(X = x_k)$
0	TTT	$1/8$
1	{HTT, THT, TTH}	$3/8$
2	{HHT, HTH, THH}	$3/8$
3	HHH	$1/8$

離散変数の分布

Def 1.7　累積分布関数 (分布関数) cumulative distribution function, CDFpp.24

(離散・連続) 確率変数 X の分布関数を $F_X(x)$ で表し，

$$F_X(x) = F(x) = P(X \leq x), \quad \forall x \in \mathbb{R}$$

- 離散変数 X に対して、分布関数は単調増加な階段関数

Thm 1.1　pp.24

関数 $F_X(x)$ がある確率変数の分布関数になるための**必要十分条件**

- 境界条件：$\lim_{x \to -\infty} F(x) = 0, \ \lim_{x \to +\infty} F(x) = 1$
- 単調増加：$x_1 \leq x_2 \ \Rightarrow \ F(x_1) \leq F(x_2)$
- 右連続：$\lim_{t \downarrow x} F(t) = F(x)$

Supp 1.3　右連続の定義

関数 f が点 $x = a$ で右連続であるとは $\lim_{t \downarrow a} f(t) = f(a)$ 右側からの極限値と関数値が一致

離散変数の分布

Ex 1.10 (Ex 1.7)

$$p_0 = \frac{1}{8}, \quad p_1 = \frac{3}{8}, \quad p_2 = \frac{3}{8}, \quad p_3 = \frac{1}{8}$$

したがって，分布関数 $F(x) = P(X \leq x)$ は

$$F(x) = \begin{cases} 0 & (x < 0) \\ \frac{1}{8} & (0 \leq x < 1) \\ \frac{4}{8} & (1 \leq x < 2) \\ \frac{7}{8} & (2 \leq x < 3) \\ 1 & (x \geq 3) \end{cases}$$

分布関数は右連続なので、「ジャンプ前の値」が得られる

離散変数の期待値・分散・標準偏差

確率変数は変動するため、その中心とばらつきは重要な記述指標

Def 1.8　期待値 expected value[pp.27]

確率 X の関数 $h(X)$ に対して，$E[h(X)] = \mu_h = \sum_{x_k \in \mathcal{X}} h(x_k) p_k$

X に対して，$E[X] = \mu_X = \sum_{x_k \in \mathcal{X}} x_k \cdot p_k$

Ex 1.11　(Ex 1.7)

$$P(X=0) = \frac{1}{8}, \quad P(X=1) = \frac{3}{8}, \quad P(X=2) = \frac{3}{8}, \quad P(X=3) = \frac{1}{8}$$

X の期待値

$$\mathbb{E}[X] = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = 1.5$$

$h(X) = X^2$ の期待値

$$\mathbb{E}[X^2] = 0^2 \cdot \frac{1}{8} + 1^2 \cdot \frac{3}{8} + 2^2 \cdot \frac{3}{8} + 3^2 \cdot \frac{1}{8} = \frac{24}{8} = 3$$

離散変数の期待値・分散・標準偏差

Thm 1.2 pp.27

a, b, c, d を定数とし，関数 $g(X), g_1(X), g_2(X)$ の期待値が存在すると仮定する

1. $E[c] = c, E[1] = 1$
2. 線形性: $E[a \cdot g_1(X) + b \cdot g_2(X) + c] = a \cdot E[g_1(X)] + b \cdot E[g_2(X)] + c$
3. $\forall x, g(x) \geq 0 \Rightarrow E[g(x)] \geq 0$
4. $\forall x, g_1(x) \geq g_2(x) \Rightarrow E[g_1(x)] \geq E[g_2(x)]$
5. $|E[g(x)]| \leq E[|g(x)|]$

Supp 1.4

$E[g(X)] = \sum_x g(x)P(X = x)$

絶対値をとる: $|E[g(x)]| = \left|\sum_x g(x)P(X = x)\right|$

三角不等式 ($|a + b| \leq |a| + |b|$) より

$|E[g(x)]| \leq \sum_x |g(x)|P(X = x) = E[|g(X)|]$

離散変数の期待値・分散・標準偏差

Def 1.9　分散 variance，標準偏差 standard deviation, SD pp.27-28

$g(X) = (X - E[X])^2 < \infty$ のとき，$E[g(X)]$ を X の分散という

X に対して，$Var[X] = E[(X - \mu_X)^2]$

$h(X)$ に対して，$Var[h(X)] = E[(h(X) - \mu_h)^2]$

- 標準偏差: $\sigma_X = \sqrt{Var[X]}, \sigma_h = \sqrt{Var[h(X)]}$

分散: 中心化された確率変数の 2 乗の期待値

- b を定数とし，$\min_b E[(X-b)^2] = E[(X - E[X])^2]$，二乗距離の最小化
- $Var[X] = E[X^2] - \{E[X]\}^2$

離散変数の期待値・分散・標準偏差

Thm 1.3 pp.27-28

定数 a, b と関数 $g(x)$ に対して，$Var[a \cdot h(X) + b] = a^2 \, Var[h(X)]$

特に，$Var[a \cdot X + b] = a^2 \, Var[X]$，$Var[b] = 0$

- 分散は平行移動には不変である
- 尺度を変えると尺度の二乗倍の影響を受ける
- 分散=0: 確率変数は定数となり、期待値に固定される

Ex 1.12 **標準化変量**

$$Z = \frac{X - E[X]}{\sqrt{Var[X]}}$$

に対して，$E[Z] = 0, \quad Var[Z] = 1$

モーメント

Def 1.10　モーメント・積率 moment[pp.67]

$k = 1, 2, \dots$ に対して，

原点まわりの k 次モーメント: $\mu'_k = E[X^k]$

平均まわりの k 次モーメント: $\mu_k = E[(X-\mu)^k], \quad \mu = \mu'_1 = E[X]$

Ex 1.13　モーメントの特例

- $\mu'_1 = E[X] = \mu$: 期待値
- $\mu_2 = E[(X-\mu)^2]$: 分散

歪度 skewness:[pp.29-39] $E[Z^3] = E[(\frac{X-E[X]}{\sqrt{Var[X]}})^3]$

尖度 kurtosis:[pp.29-30] $E[Z^4] = E[(\frac{X-E[X]}{\sqrt{Var[X]}})^4]$

- $\mu_3 = E[(X-\mu)^3]$ (歪度を表す指標)
- $\mu_4 = E[(X-\mu)^4]$ (尖度を表す指標)

離散変数の積率母関数

Def 1.11　積率母関数 moment-generating function, MGF[pp.67]

$\exists h > 0\ \forall |t| < h$,

$$M_X(t) = E[e^{tX}]$$

が存在するとき，$M_X(t)$ を X の MGF という

- $h > 0\ \forall |t| < h$:「ある $h > 0$ がとれて (0 の近傍 neighborhood)，$|t| < h$ なるすべての t に対して」「原点を含むある開区間 $(-h, h)$ の任意の要素 t に対して」
- e^{tX} は急速に発散することがあるため，t の全域で収束するとは限らない，そのため，少なくとも 0 の近傍で有限であることを要求する
- $M_X(t)$ が存在するとは，$M_X(t) < \infty$

離散変数 X に対して，$M_X(t) = E[e^{tX}] = \sum_{k=0}^{\infty} e^{tx_k} p_k$

離散変数の積率母関数

Ex 1.14 pp.67

コーシー分布 $X \sim \text{Cauchy}(0, 1)$

$$f(x) = \frac{1}{\pi(1 + x^2)}$$

この場合，

$$M_X(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{e^{tx}}{1 + x^2}\, dx$$

は $t \neq 0$ で発散する．

- $x \to +\infty$ のとき：$\frac{e^{tx}}{1+x^2} \sim \frac{e^{tx}}{x^2}$, $t > 0$ なら指数的に発散
- $x \to -\infty$ のときも同様に，$t < 0$ なら発散
- よって $M_X(t)$ が有限なのは $t = 0$ のみ

離散変数の積率母関数

なぜ MGF が有用なのか？

1. X のすべてのモーメントを得られる –> Thm 1.4
   - モーメントは、期待値・分散を計算するための統一的な方法を与える（積分を直接行う必要を避けられる）
2. MGF が存在すれば分布を一意に決定する（確率変数の MGF を求めることは，分布を決定したことを意味する）
   - 2 つの確率変数が同じ MGF を持つならば、それらは同じ分布を持つ –> Thm 1.5
   - 特に，独立な確率変数の和を扱うときに有用 –> Thm 1.6

Supp 1.5 特性関数 characteristic function, CF

$$\varphi_X(t) = E[e^{itX}], \quad t \in \mathbb{R}$$

- 常に存在する（MGF と違い発散の心配がない）
- 分布を一意に決定する
- フーリエ変換と関係
- モーメントを取り出せる： $\varphi_X^{(k)}(0) = i^k E[X^k], i = \sqrt{-1}$

離散変数の積率母関数

Thm 1.4 pp.67-68

$\exists h > 0\ \forall |t| < h$，$M_X(t)$ が存在するとき，

$$M_X(t) = E[e^{tX}] = \sum_{k=0}^{\infty} E[X^k] \frac{t^k}{k!}$$ **(証)**

$\forall k \in \mathbb{N}$[8]

$$\left. \frac{d^k}{dt^k} M_X(t) \right|_{t=0} = \frac{d^k}{dt^k} M_X(0) = M_X^{(k)}(0) = E[X^k]$$ **(証)**

- $E[X] = M'_X(0)$
- $E[X^2] = M''_X(0) \Rightarrow Var[X] = M''_X(0) - (M'_X(0))^2$

[8] $\forall k \in \mathbb{N}$: 任意の自然数 k に対して

離散変数の積率母関数

Thm 1.5 pp.67-68

2 つの確率変数 X, Y を考える．$\exists h > 0\ \forall |t| < h$，$M_X(t), M_Y(t)$ が存在かつ一致するならば，

$$F_X(u) = F_Y(u), \quad \forall u \in \mathbb{R}$$

Ex 1.15

確率変数 X について、次が分かっている

$$M_X(t) = 2^{2-t}, \quad t \in (-2, 2).$$

上の母関数は、パラメータ $\lambda = 2$ の指数分布の母関数に一致する．
したがって、

$$X \sim \text{Exponential}(2)$$

と結論づけられる．

(後)Thm 1.7 二項分布とポアソン分布の近似

(後)Thm 2.14 正規分布と二項分布の近似

離散変数の積率母関数

Supp 1.6　テイラー展開の定義

関数 $f(x)$ が点 a で十分に滑らか（n 階まで微分可能）であるとき，その関数を多項式で近似する展開

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x)$$

- $R_n(x)$：剰余項（近似誤差）
- $a = 0$ の場合を: マクローリン展開 Maclaurin expansion

マクローリン展開: $f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots$

- 例: $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$

Thm 1.6　pp.69

$X_1, X_2, \ldots, X_n$ が独立な確率変数ならば，その和の MGF

$$M_{X_1+X_2+\cdots+X_n}(t) = M_{X_1}(t)\, M_{X_2}(t) \cdots M_{X_n}(t) \quad \text{(証)}$$

Ex 1.16

（後）二項分布とベルヌーイ分布の MGF.

代表的な離散分布

$x_k = k$ の確率関数，$P(X = k)$

- 離散一様分布 discrete uniform distribution (dist.)
- 2 **項分布** binomial dist.[pp.30]
- **ポアソン分布** Poisson dist.[pp.31]
- 幾何分布 geometric dist.[pp.33]
- 負の 2 項分布 negative binomial dist.[pp.33]
- 超幾何分布 hypergeometric dist.[pp.34](略)

積率母関数 $\Rightarrow$ 期待値・分散

離散一様分布

Def 1.12　離散一様分布

n を正の整数とする．離散変数 X が

$$P(X = k \mid n) = \frac{1}{n}, \quad k = 1, 2, \ldots, n$$

なる確率関数をもつとき，$X \sim \mathrm{DU}(n)$[9]

MGF: 等比数列の和

$$M_X(t) = \frac{e^t(1 - e^{nt})}{n(1 - e^t)} \quad \textbf{(証)}$$

$$E[X] = \frac{1+n}{2} \qquad Var[X] = \frac{(n+1)(n-1)}{12}$$

Ex 1.17

$X =$ サイコロを 1 回振ったときの出目 $\Rightarrow X \sim \mathrm{DN}(6)$

[9] ∼: に従う

2 項分布

Def 1.13 2 項分布

成功確率 p のベルヌーイ試行を n 回に繰り返す．離散変数 X が

$$P(X=k \mid n,p) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$$

なる確率関数をもつとき，$X \sim \mathrm{Bin}(n,p)$

MGF: 独立ベルヌーイ試行の和の MGF[pp.67]

$$M_X(t) = \left((1-p) + pe^t\right)^n \quad \textbf{(証)}$$

$$E[X] = np \quad Var[X] = np(1-p)$$

Supp 1.7 ベルヌーイ試行 Bernoulli trial

1 回の試行において，結果が

$$\Omega = \{\text{成功}, \text{失敗}\}$$

の 2 通りしかなく，成功する確率が p，失敗する確率が $1-p$ であるとき

Ex 1.18 n 回の試行で k 回の成功

$X =$ コインを 10 回振って H が出た回数 $\Rightarrow X \sim \mathrm{Bin}(10, \frac{1}{2})$

$X =$ サイコロを 10 回振って「1」が出た回数 $\Rightarrow X \sim \mathrm{Bin}(10, \frac{1}{6})$

ポアソン分布

Def 1.14　ポアソン分布

$\lambda > 0$ をパラメータとする．離散変数 X が

$$P(X = k \mid \lambda) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

なる確率関数をもつとき，$X \sim \mathrm{Poisson}(\lambda)$

$$M_X(t) = \exp\bigl(\lambda(e^t - 1)\bigr) \quad \textbf{(証)}$$
$$E[X] = \lambda \quad Var[X] = \lambda$$

Ex 1.19　発生を待つ waiting-for-occurence

1 時間に平均 10 件の電話がかかってくる.
$X =$ ある 1 時間にかかってくる電話の件数 $\Rightarrow X \sim \mathrm{Poisson}(10)$
λ (平均発生率 rate): 10 件時間
λ が大きいほど「事象が頻繁に発生する」ことを意味する

Thm 1.7 二項分布とポアソン分布の近似

- ポアソン分布の再帰関係 (recursion relation) $P(X=k)=\frac{k}{\lambda}P(X=k-1)$
- 二項分布の再帰関係 $P(X=k)=\frac{k}{n-k+1}\cdot\frac{1-p}{p}P(X=k-1)$

二項分布 $X\sim \mathrm{Bin}(n,p)$ に対して，

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$$

とする．n が大きく，p が小さいとき，

$$n\to\infty,\quad p\to 0,\quad \text{ただし}\ \lambda=np\ \text{は一定}$$

ならば，

$$P(X=k)\approx\frac{e^{-\lambda}\lambda^k}{k!}\quad \textbf{(証)}$$

となり，

$$X\dot{\sim}^{10}\mathrm{Poisson}(\lambda)$$

と近似できる．まれな成功を多数の試行で観測する状況．

[10]近似的に従う

幾何分布

Def 1.15　幾何分布

成功確率 p のベルヌーイ試行を繰り返す．初めて成功が出るまでの試行回数を X とすると：

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

このとき，$X \sim \mathrm{Geo}(p)$

MGF: 無限等比数列の和

$$M_X(t) = \frac{pe^t}{1-(1-p)e^t}, \quad (|e^t(1-p)| < 1) \quad \textbf{(証)}$$

$$E[X] = \frac{1}{p} \quad Var[X] = \frac{1-p}{p^2}$$

Ex 1.20　初成功を待つ

$X =$ コインを投げて初めて「1」が出る (成功) までの試行回数 $\Rightarrow X \sim \mathrm{Geo}(\frac{1}{2})$

Thm 1.8　無記憶性 memoryless property

$P(X > n+k \mid X > n) = P(X > k)$

- $P(X > n) = (1-p)^n$: n 回の試行で成功しなかった

負の二項分布 (課題)

Def 1.16　負の二項分布

成功確率 p のベルヌーイ試行を繰り返し，
r 番目の成功までの**試行回数**を X とする．

$$P(X = k \mid r, p) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, \ldots$$

r 番目の成功までの**失敗回数**を X とする．

$$P(X = k \mid r, p) = \binom{r+k-1}{k} p^r (1-p)^k, \quad k = 0, 1, \ldots$$

このとき，$X \sim \mathrm{NegBin}(r, p)$

MGF: X の定義によって少し異なる **(課題)**

超幾何分布 (略)

Def 1.17　超幾何分布

母集団 N 個のうち成功とみなすものが K 個あるとする．そこから n 個を無作為に取り出したとき，成功数を X とすると：

$$P(X = k \mid N, K, n) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}, \quad \max(0, n - (N - K)) \leq k \leq \min(n, K).$$

このとき，$X \sim \mathrm{Hyper}(N, K, n)$

MGF: 一般に閉じた形は難しい

$$M_X(t) = = \frac{1}{\binom{N}{n}} \sum_{k=\max(0,\, n-(N-K))}^{\min(n,\, K)} \binom{K}{k}\binom{N-K}{n-k} e^{tk}$$

$$E[X] = n \cdot \frac{K}{N} \quad Var[X] = n \cdot \frac{K}{N}\left(1 - \frac{K}{N}\right)\frac{N-n}{N-1}$$

Ex 1.21　品質検査

100 個中不良品 10 個，15 個抜き取ったとき，$X =$ 不良品数 $\Rightarrow X \sim \mathrm{Hyper}(100, 10, 15)$

レポート

提出期限: 次の火曜日 1:00AM

期限を過ぎた場合: メールでレポートを提出. ただし 10 点減点

課題 1.1

$\alpha \in \mathbb{R}$ に対して，$f(x) = (1+x)^\alpha$ のマクローリン展開を求めよ．

課題 1.2

$r > 0, 0 < p < 1$ とする.$X \sim \mathrm{NegBin}(r, p)$ (負の二項分布) のとき，
$M_X(t), E[X], Var[X]$ を求めよ．

BEEF+をご参照ください

2. 連続確率変数の復習とその積率母関数

2025 後期 3H606 数理科学入門 (統計系)

Oct 21, 2025

数理環境論 Room 742
Email: ZHOU YI <zhouy@people.kobe-u.ac.jp>

キーワード

連続確率変数
確率密度関数　分布関数[1]

積率母関数
代表的な分布の積率母関数

[1] Web application:
https://alain003.phs.osaka-u.ac.jp/mephas_web/1_1MFScondist_jp/

連続確率変数・分布

Def 2.1 連続確率変数 (連続変数)
確率密度関数 (密度関数) probability density function, PDF

確率変数 X が任意の $a < b$ に対して，

- $P(a \leq X \leq b) = \int_a^b f_X(x)dx$ を満たす
- $f_X(x)$ が存在する

$f_X(x)$: X の確率密度関数

Def 2.2 累積分布関数 cumulative distribution function, CDF

連続確率変数 X の分布関数を $F_X(x)$[2]で表し，

$$F_X(x) = F(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)dt \quad -\infty < x < \infty$$
$$\int_{-\infty}^{x} dF(t)$$

[2] 上: Riemann 積分 (面積); 下: Lebesgue – Stieltjes 積分 (測度)

Thm 2.3

関数 $F_X(x)$ がある確率変数の分布関数になるための**必要十分条件**

- 境界条件：$\lim_{x\to-\infty} F(x) = 0,\ \ \lim_{x\to+\infty} F(x) = 1$
- 単調増加：$x_1 \leq x_2 \ \Rightarrow\ F(x_1) \leq F(x_2)$
- 右連続：$\lim_{t\downarrow x} F(t) = F(x)$

Thm 2.4

$f_X(x)$ が連続であるような点に対して，

$$f_X(x) = \frac{d}{dx} F_X(x)$$

$$f_X(x)dx = dF_X(x)$$

Thm 2.5　上側 α%点・下側 α%点 (α 分位点)

$0 < \alpha < 1$ に対して，

- $$F_X(x_\alpha) = P(X \leq x_\alpha) = \alpha \Rightarrow x_\alpha = F_X^{-1}(x_\alpha)$$

  を満たす x_α: α 分位点 (下側 α%点)

- $$P(X > x_\alpha) = 1 - F_X(x_\alpha) = \alpha \Rightarrow x_\alpha = F_X^{-1}(1 - x_{(1-\alpha)})$$

  を満たす $x_{(1-\alpha)}$: 上側 α%点

連続変数の積率母関数

Def 2.6　期待値 expected value

連続変数 X の関数 $h(X)$ に対して，

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

X に対して，

$$E[X] = \mu_X = \int_{-\infty}^{\infty} x \cdot f_X(x)\ dx$$

Def 2.7　積率母関数 moment-generating function, MGF

$\exists h > 0\ \forall |t| < h$，連続変数 X に対して，

$$M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x)\ dx$$

[3]が存在するとする.

Thm 2.8

$\exists h > 0\ \forall |t| < h$，$M_X(t)$ が存在するとき，$\forall k \in \mathbb{N}$[4]

$$\left.\frac{d^k}{dt^k} M_X(t)\right|_{t=0} = \frac{d^k}{dt^k} M_X(0) = M_X^{(k)}(0) = E[X^k] \quad \text{(証)}$$

[3] $f_X(x)$ の Laplace 変換
[4] $\forall k \in \mathbb{N}$: 任意の自然数 k に対して

代表的な連続分布

$f_X(x)$

- **正規分布** normal dist.
    - 標準正規分布 standardized normal dist.
- 一様分布 uniform distribution (dist.)
- 指数分布 exponential dist.
- ガンマ分布　Gamma dist.
- ベータ分布　Beta dist.

積率母関数 $\Rightarrow$ 期待値・分散

正規分布

中心的な役割

- 理論的に扱いやすい分布
- 平均値と最頻値，中央値が一致する事や平均値を中心にして左右対称
- 中心極限定理: 独立な多数の因子の和として表される確率変数は正規分布に従う

Def 2.9　正規分布・ガウス　Gaussian 分布

$X \sim N(\mu, \sigma^2)$，平均 μ，分散 σ^2(パラメーター parameter)

$$f_X(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad -\infty < x < \infty$$
$$= \frac{1}{\sqrt{2\pi}\sigma} exp\left\{-\frac{1}{2}\left\{\frac{x-\mu}{\sigma}\right\}^2\right\}$$

MGF: $M_X(t) = exp(\mu t + \frac{\sigma^2 t^2}{2})$　**(証)**

Def 2.10　標準正規分布

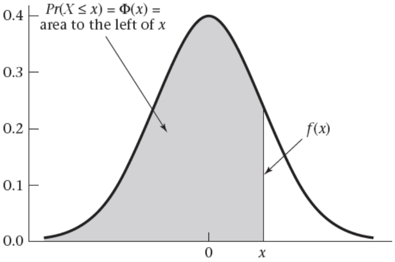
$Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

$$\phi(z) = f_Z(z) = \frac{1}{\sqrt{2\pi}} exp\left\{-\frac{z^2}{2}\right\}$$

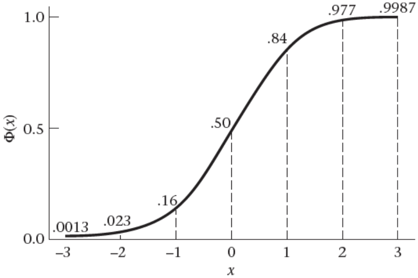
$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} exp\left\{-\frac{t^2}{2}\right\} dt$$

MGF: $M_X(t) = exp(\frac{t^2}{2})$

The cdf [$\Phi(x)$] for a standard normal distribution



The cdf for a standard normal distribution [$\Phi(x)$]



Thm 2.11

$X \sim N(\mu, \sigma^2), Z \sim N(0,1)$ とする

1. $\Phi(\infty) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} exp\left\{-\frac{t^2}{2}\right\} dt = 1$ **(証)**
2. $\int_0^{\infty} exp\left\{-\frac{t^2}{2}\right\} dt = \sqrt{\frac{\pi}{2}}$ **(証)**
3. 変曲点: $\mu \pm \sigma$

Supp 2.12 **変曲点 (inflection point) の定義**

関数 $y = f(x)$ において，点 $x = c$ が 変曲点 であるとは，

$$f''(c) = 0 \quad \text{かつ} \quad f''(x)\text{が } x = c \text{ の前後で符号を変えるとき}$$

である．

すなわち，$x = c$ において関数の 凹凸 (曲がり方) が変化する点をいう．

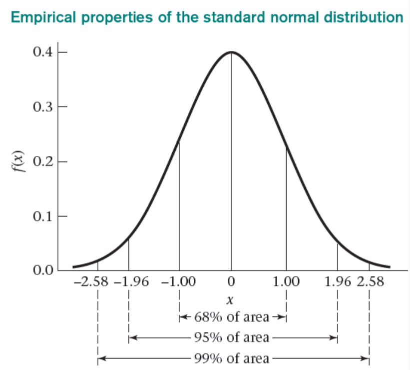
判定方法のまとめ:

- $f''(x) > 0$ のとき：上に凸 (concave up)
- $f''(x) < 0$ のとき：下に凸 (concave down)
- $f''(x)$ の符号が変わる点：変曲点

Thm 2.13

$|X-\mu| \leq \sigma, 2\sigma, 3\sigma$ の確率

- $P(|X-\mu| \leq \sigma) = P(|Z| \leq 1) = 0.6826$
- $P(|X-\mu| \leq 2\sigma) = P(|Z| \leq 2) = 0.9544$
- $P(|X-\mu| \leq 3\sigma) = P(|Z| \leq 3) = 0.9974$
- $P(|X-\mu| \leq 1.96\sigma) = P(|Z| \leq 1.96) = 0.95$



Thm 2.14　正規分布と二項分布の近似

二項分布 $X \sim \mathrm{Bin}(n, p)$ は，n が大きく p が極端でないとき，正規分布によって次のように近似できる

$$X \mathrel{\dot\sim} N(\mu = np,\, \sigma^2 = np(1-p))$$

$$Y \sim N(np, np(1-p)) \quad \textbf{(証)}$$

連続性補正 (continuity correction) を用いると，

$$P(X \leq k) \approx P(Y \leq k + 0.5)$$

$$P(X \geq k) \approx P(Y \leq k - 0.5)$$

この近似は

$$np \geq 5, \quad n(1-p) \geq 5$$

のとき良好に成り立つ．

一様分布

Def 2.15　一様分布

$X \sim \mathrm{Uniform}(a, b)$

$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{その他} \end{cases}$$

$$M_X(t) = \begin{cases} \dfrac{e^{bt} - e^{at}}{t(b-a)}, & t \neq 0, \\ 1, & t = 0. \end{cases}$$ **(証)**

$$E[X] = \frac{a+b}{2} \qquad Var[X] = \frac{(b-a)^2}{12}$$

Ex 2.16

乱数生成で「0 から 10 までの範囲で等確率に値を取る」とき，任意の 1 単位幅の区間（たとえば [2,3] や [7,8]）の確率はすべて等しくなる

1 時間に 1 本のバスが運行しているバス停に到着したとする. 直前のバスがいつ出発したかは分からないため、次のバスが到着するまでの待ち時間 $X \sim [0, 1]$(時間)

指数分布 (課題)

Def 2.17　指数分布

$X \sim \mathrm{Exp}(\lambda)$ $(\lambda > 0)$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & \text{その他} \end{cases}$$

MGF: (課題) 期待値，分散

Ex 2.18　生存時間

$X =$ 到着・故障・死亡などの「発生を待つ」時間 (waiting time)
$P(X > s)$: 時間 s を超えて生存する確率

$$P(X > s) = 1 - F(s) = e^{-\lambda s} \quad \textbf{(証)}$$

s 時間生存した条件の下で，さらに t 時間を超えて生存する確率
$P(X > s + t \,|\, X > s) = P(X > t)$　**(証)**　(無記憶性，vs 幾何分布)

Thm 2.19　ハザード関数 hazard function(生存時間解析)

X が非負連続確率変数 (故障・死亡する時間)，$X \sim f(x)$ 分布関数 $F(x)$
x まで動作 (生存) していて次の時間 $x + \Delta$ までに死亡する条件付き確率

$$P(x < X \leq x + \Delta \mid X > x) = \frac{P(x < S \leq x + \Delta, X > x)}{P(X > x)}$$
$$= \frac{P(x < X \leq x + \Delta)}{P(X > x)} = \frac{F(x + \Delta) - F(x)}{1 - F(x)}$$

$$h(x) = \lim_{\Delta x \to 0} \frac{P(x \leq X < x + \Delta \mid X \geq x)}{\Delta} = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{S(x)}$$

時刻 x まで生存していた個体が、その直後に死亡する確率・速さ（rate）
$h(x)$: ハザード関数，$S(x)$: 生存関数，$f(x)$: 死亡の速さ

- $F(x) = 1 - exp\left\{\int_0^x h(t)dt\right\}$　**(証)**
- $f(x) = h(x)exp\left\{-\int_0^x h(t)dt\right\}$　**(証)**
- $h(x) = abx^{b-1}, a > 0, b > 0$
  $\Rightarrow f(x \mid ab) = abx^{b-1}exp\{-ax^b\}, x > 0$ ウイブル分布 Weibull dist.

ガンマ分布

Def 2.20　ガンマ関数 Gamma function

実数 $\alpha > 0$ に対して，

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt$$

Thm 2.21

$$\begin{aligned}\Gamma(\alpha+1) =& \alpha\Gamma(\alpha) \quad (\alpha > 0) \quad \text{(証)}\\ \Gamma(n) =& (n-1)! \quad (n = 1, 2, 3, \dots)\\ \Gamma(1) =& 1, \quad \Gamma(\frac{1}{2}) = \sqrt{\pi}\end{aligned}$$

Def 2.22　ガンマ分布

$X \sim \mathrm{Gamma}(\alpha, \beta)$ $(\alpha > 0,\ \beta > 0)$

$$f_X(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\,\beta^{\alpha}}\, x^{\alpha-1} e^{-x/\beta}, & x \geq 0, \\ 0, & \text{その他} \end{cases}$$

- α: 形状パラメーター shape parameter
- β: 尺度パラメーター scale parameter

$$M_X(t) = (1 - \beta t)^{-\alpha} \quad (t < 1/\beta) \quad \textbf{(証)}$$
$$E[X] = \alpha\beta, \quad Var[X] = \alpha\beta^2$$

Ex 2.23

指数分布は，ガンマ分布の特別な場合である：

$$\text{Exponential}(\lambda) = \text{Gamma}(\alpha = 1,\ \beta = 1/\lambda),$$

確率密度関数は

$$f(x;\lambda) = \lambda e^{-\lambda x}, \quad x > 0.$$

ある事象が発生するまでの時間（waiting time）をモデル化する際に用いられる

Thm 2.24　ガンマ・ポアソン関係（Gamma – Poisson relationship）

$X \sim \mathrm{Gamma}(\alpha, \beta), \alpha \in \mathbb{Z}, \forall x,$

$$P(X \leq x) = P(Y \geq \alpha), \quad Y \sim \mathrm{Poisson}\left(\lambda = \frac{x}{\beta}\right)$$

（証）

ガンマ分布の CDF が、ポアソン分布の上側確率（右側確率）として表される.

Ex 2.25　ガンマ分布とその特別な場合（指数分布・カイ二乗分布）

- 指数分布 ($\lambda = 1/\beta$):

$$f(x; \beta) = \frac{1}{\beta} e^{-x/\beta}$$

  $\alpha = 1$ のときのガンマ分布

- カイ二乗分布:

$$f(x; p) = \frac{1}{2^{p/2}\Gamma(p/2)} x^{p/2-1} e^{-x/2}$$

  $\alpha = \frac{p}{2}, \beta = 2$ のときのガンマ分布

ベータ分布

Def 2.26 ベータ関数

実数 $\alpha > 0,\ \beta > 0$ に対して，

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\, dt$$

Thm 2.27 ガンマ関数との関係

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

Def 2.28 ベータ分布

$X \sim \mathrm{Beta}(\alpha, \beta)\ (\alpha > 0,\ \beta > 0)$

$$f_X(x) = \begin{cases} \dfrac{1}{B(\alpha,\beta)}\, x^{\alpha-1}(1-x)^{\beta-1}, & 0 \leq x \leq 1, \\ 0, & \text{その他} \end{cases}$$

- 対称性：$B(\alpha,\beta) = B(\beta,\alpha)$
- $\alpha > 1, \beta = 1$: 狭義単調増加 (strictly increasing $x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$)
- $\alpha < 1, \beta < 1$: U 字型
- $\alpha > 1, \beta > 1$: 単峰型
- $\alpha = \beta$: $x = \frac{1}{2}$ に対して対称
- $\alpha = \beta = 1$: 一様分布 $U(0,1)$

$$M_X(t) = \frac{1}{B(\alpha,\beta)} \int_0^1 e^{tx} x^{\alpha-1}(1-x)^{\beta-1}\,dx \quad \text{(証)}$$

$$E[X] = \frac{\alpha}{\alpha+\beta}, \quad Var[X] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Ex 2.29　確率・比率のモデル

確率・割合・成功率などを表す
ベイズ推定における事前分布

Thm 2.30　ベータ分布の n 次モーメント

$X \sim \mathrm{Beta}(\alpha, \beta)$ に対して，

$$E[X^n] = \frac{1}{B(\alpha,\beta)} \int_0^1 x^{n+\alpha-1}(1-x)^{\beta-1}\,dx$$
$$= \frac{B(\alpha+n,\beta)}{B(\alpha,\beta)} = \frac{\Gamma(\alpha+n)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\alpha+\beta+n)}.$$ (証)

整数の場合，

$$E[X^n] = \frac{(\alpha)_n}{(\alpha+\beta)_n}, \quad (a)_n = a(a+1)\cdots(a+n-1).$$

レポート

提出期限: 次の火曜日 1:00AM

課題 2.31

指数分布の MGF，期待値，分散を求めよ．

課題 2.32

$\Gamma(\frac{1}{2}) = \sqrt{\pi}$

BEEF+をご参照ください

3. 二変量離散確率変数の復習と多項分布

2025 後期 3H606 数理科学入門 (統計系)

Nov 4, 2025

数理環境論 Room 742
Email: ZHOU YI <zhouy@people.kobe-u.ac.jp>

キーワード

2 変量離散確率変数
確率 (質量) 関数 (PMF): 同時 PMF　周辺 PMF　条件付き PMF

期待値: 同時期待値 周辺期待値 条件付き期待値
条件付き分散

代表的な分布: 多項分布

Def 3.1 2 変量確率変数 Bivariate random variable

2 つ確率変数の組 (X, Y) を考える，$x \in \mathbb{R}, y \in \mathbb{R}$

(X, Y): 2 次元平面 $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ の上に値を持つ 2 変量 (次元) 確率変数

- (X, Y) の確率分布は $\mathbb{R}^2$ 上に分布する
- $X \in \mathcal{X} = \{x_0, x_1, x_2, \ldots\}$, $Y \in \mathcal{Y} = \{y_0, y_1, y_2, \ldots\}$ 2 変量離散確率変数 (X, Y)

Ex 3.2 コインを 2 枚投げる確率

Coin1: $P(H^1) = P(T^1) = \frac{1}{2}$, $X =$Coin1 の結果,
$X = \{X(H^1) = x_1 = 0, X(T^1) = x_2 = 1\}$, $X \in \mathcal{X} = \{0, 1\}$

Coin2: $P(H^2) = P(T^2) = \frac{1}{2}$, $Y =$Coin2 の結果,
$Y = \{Y(H^2) = y_1 = 0, Y(T^2) = y_2 = 1\}$, $Y \in \mathcal{Y} = \{0, 1\}$ とする

$(X, Y) \in \mathcal{X} \times \mathcal{Y} = \{\{0, 0\}, \{0, 1\}, \{1, 0\}, \{1, 1\}\}$: 2 変量 (次元) 確率変数

同時分布

$X = x$ かつ $Y = y$ である確率

$$P(\{X = x\} \cap \{Y = y\}) = P(X = x,\, Y = y) = f_{X,Y}(x, y), \quad (x, y) \in \mathcal{X} \times \mathcal{Y}$$

(X, Y) が集合 (事象)C に入る確率

$$P((X, Y) \in C) = \sum_{(x,y)\in C} f_{X,Y}(x, y)$$

$P((X, Y) \in C)$: 同時分布

Ex 3.3 **コインを 2 枚投げる確率 (つづき)**

$\{\{T^1, H^2\}, \{H^1, T^2\}\}$: 一つの T が出ること
$P((X, Y) \in C) = P(X = 1, Y = 0) + P(X = 0, Y = 1)$

Def 3.4 **同時確率質量関数** Joint probability mass function (Joint PMF)

$f_{X,Y}(x, y) \geq 0$ であり, s.t.

$$\sum_{(x,y)\in\mathcal{X}\times\mathcal{Y}} f_{X,Y}(x, y) = \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} f_{X,Y}(x, y) = 1$$

$f_{X,Y}(x, y)$: 同時確率関数

周辺分布

$\mathcal{X}$ 上の集合 A に対して，事象 $\{X \in A\} \cap \{Y \in \mathcal{Y}\}$ の確率

$$\begin{aligned} P(X \in A) &= P(\{X \in A\} \cap \{Y \in \mathcal{Y}\}) = P((X, Y) \in A \times \mathcal{Y}) \\ &= \sum_{(x,y) \in A \times \mathcal{Y}} f_{X,Y}(x, y) = \sum_{x \in A} \sum_{y \in \mathcal{Y}} f_{X,Y}(x, y) \end{aligned}$$

$P(X \in A)$: X の周辺分布

Def 3.5　X の周辺確率質量関数　Marginal PMF of X

$$f_X(x) = P(X = x) = \sum_{y \in \mathcal{Y}} f_{X,Y}(x, y)$$

$f_X(x)$: X の周辺確率質量関数

Ex 3.6　コインを 2 枚投げる確率（つづき）

X の周辺確率質量関数 $P(X = x) = P(X = x, Y = 0) + P(X = x, Y = 1)$

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

$\mathcal{Y}$ 上の集合 A に対して，事象 $\{X \in \mathcal{X}\} \cap \{Y \in A\}$ の確率

$$
\begin{aligned}
P(Y \in A) &= P(\{X \in \mathcal{X}\} \cap \{Y \in A\}) = P((X, Y) \in \mathcal{X} \times A) \\
&= \sum_{(x,y) \in \mathcal{X} \times A} f_{X,Y}(x, y) = \sum_{x \in \mathcal{X}} \sum_{y \in A} f_{X,Y}(x, y)
\end{aligned}
$$

$P(Y \in A)$: Y の周辺分布

Def 3.7 **Y の周辺確率質量関数** Marginal PMF of Y

$$
f_Y(y) = \sum_{x \in \mathcal{X}} f_{X,Y}(x, y)
$$

$f_Y(y)$: Y の周辺確率質量関数

Ex 3.8 **コインを 2 枚投げる確率 (つづき)**

Y の周辺確率質量関数 $P(Y = y) = P(X = 0, Y = y) + P(X = 1, Y = y)$

分割表

Def 3.9 分割表 contingency table

2 つ以上の質的変数（カテゴリーデータ）の関係を整理・表示するための表

行と列にそれぞれのカテゴリを配置し、各セルにはそれぞれのカテゴリの組み合わせに該当する確率 (観測頻度・割合など) を示す。

Ex 3.10 コインを 2 枚投げる確率 (つづき)

Table 1: 2 コイン投げにおける確率の分割表

	Coin2: H^2 ($Y=0$)	Coin2: T^2 ($Y=1$)	X の周辺確率
Coin1: H^1 ($X=0$)	0.25	0.25	0.50
Coin1: T^1 ($X=1$)	0.25	0.25	0.50
Y の周辺確率	0.50	0.50	1.00

期待値

Def 3.11 期待値 expected value

関数 $g(X, Y)$ の同時確率関数 $f_{X,Y}(x, y)$ に関する期待値

$$E[g(X, Y)] = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} g(x, y) f_{X,Y}(x, y)$$

X の周辺確率関数 $f_X(x)$ に関する期待値

$$\begin{aligned} E[g(X)] &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} g(x) f_{X,Y}(x, y) \\ &= \sum_{x \in \mathcal{X}} g(x) \sum_{y \in \mathcal{Y}} f_{X,Y}(x, y) \\ &= \sum_{x \in \mathcal{X}} g(x) f_X(x) \end{aligned}$$

条件付き分布

Def 3.12　条件付き確率関数 conditional probability function

$f_X(x) \neq 0$ なる x に対して，$X = x$ を与えたときの $Y = y$ の条件付き[1]確率関数

$$f_{Y|X}(y|x) = P(Y = y|X = x) = \frac{f_{X,Y}(x, y)}{f_X(x)}$$

Def 3.13　条件付き期待値 conditional expectation

$X = x$ を与えたときの $Y = y$ の条件付き期待値

$$E[Y|X = x] = \sum_{y \in \mathcal{Y}} y \cdot f_{Y|X}(y|x) = \frac{\sum_{y \in \mathcal{Y}} y \cdot f_{X,Y}(x, y)}{f_X(x)}$$

- $E[g(X, Y)] = E[E[g(X, Y)|X]] = E[E[g(X, Y)|Y]]$ **（証）**
- 条件付き分散 conditional variance:
  $Var[Y|X = x] = E[(Y - E[Y|X - x])^2|X = x] = E[Y^2|X = x] - (E[Y|X = x])^2$

[1] $Y \mid X$: Y given X

独立性

Def 3.14　独立 independent

(X, Y) の同時 PMF を $f_{X,Y}(x, y)$，X と Y の周辺 PMF を $f_X(x), f_Y(y)$ とする．
$\forall x \in \mathcal{X}, y \in \mathcal{Y}$

$$f_{X,Y}(x, y) = f_X(x) f_Y(y)$$

X と Y が独立である

$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$　**(証)**

多項分布

Def 3.15　多項分布 multinomial distribution

$(X_1,\dots,X_k)=(x_1,\dots,x_k)$ における同時確率関数が

$$f_{1,\dots,k}(x_1,\dots,x_k;\ n,p_1,\dots,p_{k-1})=\frac{n!}{x_1!\cdots x_k!}p_1^{x_1}\cdots p_k^{x_k}$$

s.t. $p_1+\cdots+p_k=1,\ \ x_1+\cdots+x_k=n$

$(X_1,\dots,X_k)\sim Multin(n,p_1,\dots,p_k)$

Ex 3.16

サイコロを n 回投げて 1 から 6 の面が出る回数をそれぞれ $(X_1,\dots,X_6)$ とするとき

$$(X_1,\dots,X_6)\sim Multin(n,p_1,\dots,p_6)$$

2 個面からなるサイコロ n 回投げて 1 から 2 の面が出る回数をそれぞれ (X_1,X_2) とするとき

$$(X_1,X_2)\sim Multin(n,p_1,p_2)$$

$\Leftrightarrow$ コインを n 回投げて 1 の面が出る回数を X とするとき

$$X\sim Bin(n,p_1)$$

Thm 3.17　多項定理 multinomial theorem

$$(p_1 + \cdots + p_k)^n = \sum_{(x_1,\ldots,x_k)\in\mathcal{X}} \frac{n!}{x_1!\cdots x_k!} p_1^{x_1}\cdots p_k^{x_k}$$

$\mathcal{X} = \{(x_1,\ldots,x_k) | x_1 + \cdots + x_k = n, 0 \le x_i \le n, x_i \in \mathbb{Z}\}$

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

期待値

$$E[\mathbf{X}] = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_k) \end{bmatrix} = \begin{bmatrix} np_1 \\ np_2 \\ \vdots \\ np_k \end{bmatrix} = n\mathbf{p}$$

$$\Leftrightarrow E(X_i) = np_i$$

分散

$$Var[X_i] = np_i(1 - p_i)$$

共分散

$$Cov[X_i, X_j] = -np_i p_j$$

$$\Leftrightarrow Cov[\mathbf{X}] = \begin{pmatrix} \mathrm{Var}(X_1) & \mathrm{Cov}(X_1, X_2) & \cdots & \mathrm{Cov}(X_1, X_k) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Var}(X_2) & \cdots & \mathrm{Cov}(X_2, X_k) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{Cov}(X_k, X_1) & \mathrm{Cov}(X_k, X_2) & \cdots & \mathrm{Var}(X_k) \end{pmatrix}$$

X_i の周辺分布は $Bin(n, p_i)$

共分散

2 つの確率変数 X と Y の間にある関係性（傾向）を示す指標

Def 3.18　共分散 Covariance

2 つの確率変数 X と Y の共分散 $Cov[X, Y]$

$$Cov[X, Y] = E[(X - E[X])(Y - E[Y])]$$

$Cov[X, Y] = E[XY] - E[X]E[Y]$

$Var[X] = Cov[X, X] = E[XX] - E[X]E[X]$

レポート

課題 3.19

$(X_1, \ldots, X_k) \sim Multin(n, p_1, \ldots, p_k)$ とする

$X_k = x_k$ を与えたとき

$$f_{1,\ldots,k-1|k}(x_1, \ldots, x_{k-1}|x_k)$$

条件付き確率関数を求めよ

BEEF+をご参照ください

4. 二変量連続確率変数の復習と多変量正規分布

2025 後期 3H606 数理科学入門 (統計系)–#6

Nov 11, 2025

数理環境論 Room 742
Email: ZHOU YI <zhouy@people.kobe-u.ac.jp>

キーワード

２変量連続確率変数
確率密度関数 (PDF): 同時 PDF　周辺 PDF　条件付き PDF

期待値: 同時期待値 周辺期待値 条件付き期待値
条件付き分散

代表的な分布: ２変量・多変量正規分布

2 変量連続確率変数

Def 4.1　2 変量確率変数 Bivariate random variable

2 つ確率変数の組 (X, Y) を考える，$x \in \mathbb{R}, y \in \mathbb{R}$

(X, Y): 2 次元平面 $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ の上に値を持つ 2 変量 (次元) 確率変数

- (X, Y) の確率分布は $\mathbb{R}^2$ 上に分布する
- $X \in \mathcal{X} = (-\infty, \infty),\ \ Y \in \mathcal{Y} = (-\infty, \infty)$ 2 変量連続確率変数 (X, Y)

同時分布

2 変量連続確率変数 (X, Y) が $\mathbb{R}^2$ 上の集合 (事象)C に対して確率

$$P((X, Y) \in C) = \int\int_{(x,y)\in C} f_{X,Y}(x, y)\ dxdy$$

$P((X, Y) \in C)$: 同時分布

Def 4.2　同時確率密度関数　Joint probability density function, Joint PDF

$f_{X,Y}(x, y) \geq 0$ であり, s.t.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)\ dxdy = 1$$

$f_{X,Y}(x, y)$: 同時確率密度関数

同時累積分布関数 Joint cumulative distribution function, Joint CDF:

$$F_{X,Y}(x,y) = P(X \leq x,\, Y \leq y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(s,t)\ dtds$$

$$\Rightarrow f_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y)$$

周辺分布

Def 4.3 X・Y **の周辺確率密度関数** Marginal PDF

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\ dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\ dx$$

期待値

Def 4.4　期待値 expected value

関数 $g(X, Y)$ の同時確率関数 $f_{X,Y}(x, y)$ に関する期待値

$$E[g(X, Y)] = \int_{x\in\mathbb{R}} \int_{y\in\mathbb{R}} g(x, y) f_{X,Y}(x, y)\ dxdy$$

X の周辺確率関数 $f_X(x)$ に関する期待値

$$E[g(X)] = \int_{x\in\mathbb{R}} g(x) f_X(x)\ dx$$

条件付き期待値・分散

Def 4.5　条件付き確率密度関数 conditional probability density function

$f_X(x) > 0$ なる x に対して，$X = x$ を与えたときの Y の条件付き[1]PDF

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x, y)}{f_X(x)}$$

Def 4.6　条件付き期待値 conditional expectation

$X = x$ を与えたとき，Y の関数 $g(x, Y)$ の条件付き期待値

$$E[g(x, Y)|X = x] = \int_{y\in\mathbb{R}} g(x, y) \cdot f_{Y|X}(y|x)\ dy = \frac{\int_{y\in\mathbb{R}} g(x, y) \cdot f_{X,Y}(x, y)\ dy}{f_X(x)}$$

Y の条件付き期待値

$$E[Y|X = x] = \int_{y\in\mathbb{R}} y \cdot f_{Y|X}(y|x)\ dy = \frac{\int_{y\in\mathbb{R}} y \cdot f_{X,Y}(x, y)\ dy}{f_X(x)}$$

[1] $Y \mid X$: Y given X

Y の条件付き期待値：$\mu_{Y|X} = E[Y|X = x]$

Y の条件付き分散：$Var[Y|X = x]$

- $Var[Y|X = x] = E[(Y - \mu_{Y|X})^2|X = x]$
- $Var[Y|X = x] = E[Y^2|X = x] - (E[Y|X = x])^2$

Thm 4.7

- $E[X] = E[E[X|Y]]$　**（証）**
- $Var[X] = E[Var[X|Y]] + Var[E[X|Y]]$　**（証）**

独立性

Def 4.8　独立 independent

(X, Y) の同時 PDF を $f_{X,Y}(x,y)$，X と Y の周辺 PDF を $f_X(x), f_Y(y)$ とする．
$\forall x \in \mathbb{R}, y \in \mathbb{R}$

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

$\Leftrightarrow X \perp Y, X \perp\!\!\!\perp Y$

$\perp, \perp\!\!\!\perp$[2]

$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$　**(証)**

[2] X と Y が独立である

共分散

2 つの確率変数 X と Y の間にある線形関係性・傾向（linear relationship）を示す指標

Def 4.9　共分散 Covariance

2 つの確率変数 X と Y の共分散

$$\sigma_{XY} = Cov[X, Y] = E[(X - E[X])(Y - E[Y])]$$

- $Cov[X, Y] = E[XY] - E[X]E[Y]$
- $Var[X] = Cov[X, X] = E[XX] - E[X]E[X]$
- $Cov[aX, bY] = abCov[X, Y]$
- $Var[aX + bY] = a^2 Var[X] + b^2 Var[Y] + 2abCov[X, Y]$　**(証)**

相関係数

Def 4.10　相関係数 correlation coefficient

2 つの確率変数 X と Y の相関係数

$$\rho_{XY} = Corr(X, Y) = \frac{Cov[X, Y]}{\sqrt{Var[X]\,Var[Y]}}$$

- $|\rho_{XY}| \leq 1$; 正の線形相関: $\rho_{XY} > 0$; 負の線形相関: $\rho_{XY} < 0$; 無相関: $\rho_{XY} = 0$
- $\{Cov[X, Y]\}^2 \leq Var[X]\,Var[Y]$
- $\{E[(X-\mu_X)(Y-\mu_Y)]\}^2 \leq E[(X-\mu_X)^2]E[(Y-\mu_Y)^2]$ (コーシー・シュバルツの不等式 Cauchy-Schwarz inequality)

Supp 4.11 コーシー・シュバルツの不等式 Cauchy-Schwarz inequality

実数列 $a_1, \ldots, a_n$ と $b_1, \ldots, b_n$ に対して

$$\left|\sum_{i=1}^{n} a_i b_i\right| \leq \sqrt{\sum_{i=1}^{n} a_i^2}\,\sqrt{\sum_{i=1}^{n} b_i^2}.$$

$\exists\lambda \in \mathbb{R},\, b_i = \lambda a_i\ (i = 1, \ldots, n)$，等号が成り立つ[3]

$\forall X,\, Y$(任意の２つ確率変数)

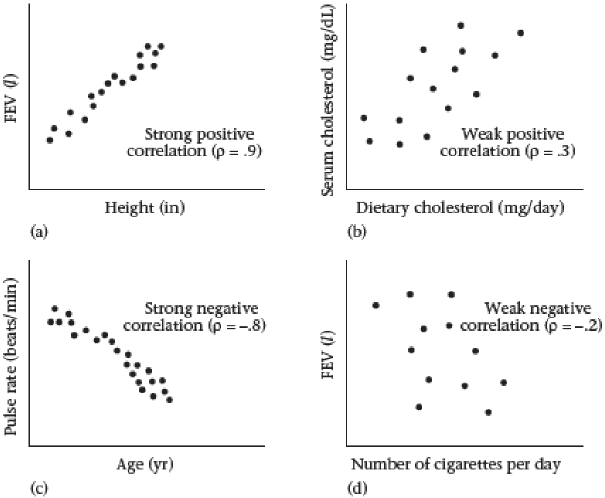
$$|E[XY]| \leq E[|XY|] \leq \sqrt{E[|X|^2]}\sqrt{E[|Y|^2]}$$

共分散 vs 相関係数

- 共分散: 2 つの変数がどの方向に一緒に変化するか（増える・減る）を表す
  - 方向（正か負か）
  - 大きさは X, Y の単位に依存する
- 相関係数: 共分散を標準化して、単位の影響を取り除いたもの
  - 方向＋強さ（どのくらい一緒に動くか）

[3]∃: exist. 実数 λ が存在

Interpretation of various degrees of correlation



Def 4.12 多変量正規分布 multivariate normal distribution

連続確率変数 $\boldsymbol{X}$ の $\boldsymbol{x} = (x_1, \ldots, x_k)^\top$ における同時密度関数が

$$f_{\boldsymbol{X}}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{k/2}} \frac{1}{|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})\right\}$$

$\boldsymbol{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

- $\int_{\mathbb{R}^k} f_{\boldsymbol{X}}(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})\ d\boldsymbol{x} = 1$
- $\boldsymbol{X} = (X_1, \ldots, X_k)$
- 平均：$\boldsymbol{\mu}^\top = (\mu_1, \ldots, \mu_k) = (E[X_1], \ldots, E[X_k])$

- 分散共分散行列 (共分散行列 covariance matrix): $\mathbf{\Sigma}$

$$\mathbf{\Sigma} = \begin{pmatrix} Var[X_1] & Cov[X_1, X_2] & \cdots & Cov[X_1, X_k] \\ Cov[X_2, X_1] & Var[X_2] & \cdots & Cov[X_2, X_k] \\ \vdots & \vdots & \ddots & \vdots \\ Cov[X_k, X_1] & Cov[X_k, X_2] & \cdots & Var[X_k] \end{pmatrix}$$

$$= \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{21} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_{kk} \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1k} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{21} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{k1} & \sigma_{k2} & \cdots & \sigma_k^2 \end{pmatrix}$$

- $\mathbf{\Sigma}$: 正定値・対称行列 $\Rightarrow |\mathbf{\Sigma}| > 0, \mathbf{\Sigma}^{-1}$ が存在する
  - $\mathbf{\Sigma}^{-1}$: $\mathbf{\Sigma}$ の逆行列 inverse matrix
  - $|\mathbf{\Sigma}| = det(\mathbf{\Sigma})$: $\mathbf{\Sigma}$ の行列式 determinant

Supp 4.13　**正定値・準正定値・負定値行列**

1. 正定値行列（positive definite matrix）: 実対称行列 A に対して、任意の零でない列ベクトル $\boldsymbol{x} \in \mathbb{R}^n$ に対して
$$\boldsymbol{x}^\top A\boldsymbol{x} > 0$$
が成り立つとき、行列 A を **正定値行列** と呼ぶ
2. 準正定値行列（positive semidefinite matrix）: 実対称行列 A に対して、任意の列ベクトル $\boldsymbol{x} \in \mathbb{R}^n$ に対して
$$\boldsymbol{x}^\top A\boldsymbol{x} \geq 0$$
が成り立つとき、行列 A を **準正定値行列** と呼ぶ
3. 負定値行列（negative definite matrix）: 実対称行列 A に対して、任意の零でない列ベクトル $\boldsymbol{x} \in \mathbb{R}^n$ に対して
$$\boldsymbol{x}^\top A\boldsymbol{x} < 0$$
が成り立つとき、行列 A を **負定値行列** と呼ぶ

これらの行列はすべて $A = A^\top$ を満たすとき，対称行列になる

対応する単一値の場合

- 実数 s に対して，$\exists c \in \mathbb{R}, c \neq 0, c \cdot s \cdot c > 0$ が成り立つ
- s は分散となる条件を満たす

Supp 4.14 **逆行列（inverse matrix）**

n 次正方行列 A に対して、同じサイズの行列 B が存在し、

$$AB = BA = I_n$$

が成り立つとき、この行列 B を A の 逆行列と呼び、

$$B = A^{-1}$$

と表す。

- 逆行列が存在する行列 A は 正則行列（invertible matrix）と呼ばれる
- 逆行列が存在しない ($det(A) = 0$ の場合) は 特異行列（singular matrix）と呼ばれる

2 変量正規分布

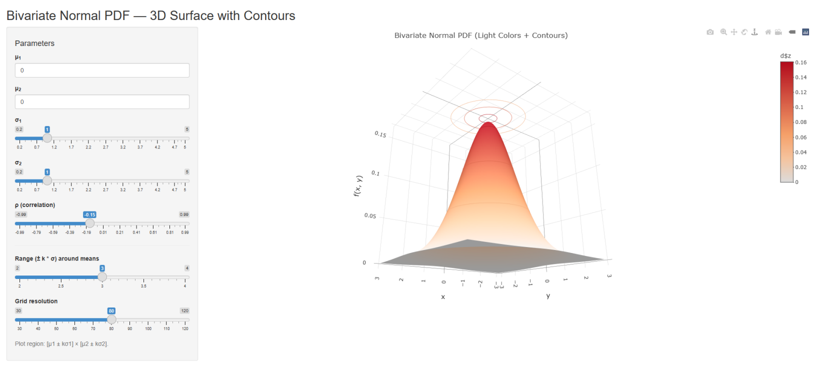
$$\boldsymbol{X} = (X,\,Y) \sim \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} \right)$$ とする

- $(X,\,Y)$ の同時分布　**(証)**

$$\begin{aligned} f_{X,Y}(x,y) = &\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \times \\ &\exp\left[-\frac{1}{2(1-\rho^2)} \left\{ \frac{(x-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2} \right\} \right] \end{aligned}$$

- 周辺分布: $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$　**(証)**
- $Y|X = x$ の条件付き分布: $Y|X \sim N_2\left(\mu_2 + \rho\sigma_2^2 \frac{x-\mu_1}{\sigma_1}, (1-\rho^2)\sigma_2^2 \right)$　**(証)**
- $E[Y|X=x] = \mu_2 + \rho\sigma_2^2 \frac{x-\mu_1}{\sigma_1}$
- $Var[Y|X=x] = (1-\rho^2)\sigma_2^2$
- $X \perp Y \Rightarrow \rho = 0$

https://yi-apr.shinyapps.io/BNormal/



レポート

課題 4.15

BEEF+ Report5 をご参照ください

5. 多変量の積率母関数・不等式

2025 後期 3H606 数理科学入門 (統計系)–#7

Nov 18, 2025

数理環境論 Room 742
Email: ZHOU YI <zhouy@people.kobe-u.ac.jp>

キーワード

復習

- １・２・多変量確率変数
  - 離散: PMF→CDF
  - 連続: PDF→CDF
  - 代表的な分布と応用

積率

- 積率母関数 MGF
- 性質と応用

不等式

- 確率不等式
- 数値不等式
- 関数不等式

1 変量確率変数

確率空間 $(\Omega, \mathcal{B}, P)$

- 標本空間 Ω，Ω の部分集合 σ-代数 $\mathcal{B}$，3 つの条件（非負性・全事象の確率・加法性）を満たす関数 P
  - P は確率
  - P の定義域は $\mathcal{B}$

確率変数（Random Variable, RV）：試行の各結果に実数を対応させる関数

$$X : \Omega \to \mathbb{R}$$

- 離散型（discrete）：取り得る値が可算個（有限または可算無限）
- 連続型（continuous）：取り得る値が非可算無限（ある区間上の全ての実数）

	離散変数	連続変数
PMF[1]/PDF[2]	$f_X(x) = P(X = x)$	$f_X(x)$
$P(a \leq X \leq b)$	$\sum_{x:a \leq x \leq b} f_X(x)$	$\int_a^b f_X(x)\, dx$
CDF[3]	$F_X(x) = \sum_{t \leq x} f(t)$	$F_X(x) = \int_{-\infty}^{x} f(t)\, dt$
確率の総和	$\sum_x f_X(x) = 1$	$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$
単一点の確率	$P(X = x) \geq 0$	$P(X = x) \equiv 0$ [4]

[1] Probability mass function 確率（質量）関数
[2] Probability density function 確率密度関数
[3] Cumulative distribution function 累積分布関数
[4] $\equiv$: equivalent 同等・常に成立

確率分布

X の確率分布

$$\forall B \subset \mathbb{R}, \quad P(X \in B) = P(\{\omega \in \Omega \mid X(\omega) \in B\})$$

で与えられる関数

$$X \sim f_X(x;\theta)$$

[5]

パラメータの値が与えられた場合、確率分布は事象の確率を予測する

[5] 或いは $f_X(x|\theta)$ で表す．θ: parameter. x: 実現値・観測値

代表的な離散型確率分布

- 離散一様分布 $\mathrm{DU}(1, N)$
  - 有限個の値を等確率でとる乱数の確率分布
    - サイコロを 1 回振ったときの出目の確率
- ベルヌーイ分布 $\mathrm{Ber}(p)$
  - 成功／失敗など二値事象の確率分布
    - 1 回コイン投げで表が出る確率
- 二項分布 $\mathrm{Bin}(n, p)$
  - n 回独立試行の成功回数の確率分布
    - 10 回コインを投げて表が出る回数 (0,1,2...) の確率
- ポアソン分布 $\mathrm{Poi}(\lambda)$
  - 単位時間・単位空間あたりの事象発生回数
    - 1 時間あたりのバス停に来るバスの本数 (0,1,2,...) の確率
- 幾何分布 $\mathrm{Geo}(p)$
  - 初めて成功するまでの試行回数の確率分布
    - 最初に表が出るまでコインを投げる回数 (1,2,...) の確率
- 負の二項分布 $\mathrm{NB}(r, p)$
  - 複数回の成功までに必要な試行回数の確率分布
    - 3 回目の成功までに必要なコイン投げの回数 (1,2...) の確率
- 超幾何分布 $\mathrm{Hyp}(N, K, n)$
  - 標本抽出における成功数のの確率分布
    - ある工場で製造された 100 個の製品の中に不良品が 10 個あるとする、抜き取った 20 個の中に不良品が 0,1,2,... 個含まれる確率

代表的な連続型確率分布

- 一様分布 $\mathrm{U}(a, b)$
    - 区間上で一様に発生する値の確率分布
        - 0〜1 の間で無作為に生成される乱数の確率分布
- 正規分布 $\mathrm{N}(\mu, \sigma^2)$
    - 自然現象や測定値の分布、平均付近に集中
        - 大学のテスト点数が平均点付近に集中する分布
- 指数分布 $\mathrm{Exp}(\lambda)$
    - 待ち時間や寿命の確率分布
        - 電球が切れるまでの寿命 (1,1.7,2,2.35... 年) の確率
- ガンマ分布 $\mathrm{Gam}(k, \lambda)$
    - 複数の独立した待ち時間の合計の分布
        - 機械の累積稼働時間（故障までの合計時間）の確率
- ベータ分布 $\mathrm{Beta}(\alpha, \beta)$
    - 確率や割合の事前分布として利用
        - ウェブサイトで広告をクリックする割合の確率分布

多変量確率変数

	離散変数	連続変数
同時 PMF/PDF	$f_{X,Y}(x,y) = P(X=x,\, Y=y)$	$f_{X,Y}(x,y)$
周辺分布 $f_X(x)$	$\sum_y f_{X,Y}(x,y)$	$\int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dy$
条件分布 $f_{X\|Y}(x\|y)$	$\frac{f_{X,Y}(x,y)}{f_Y(y)}$ $(f_Y(y) > 0)$	$\frac{f_{X,Y}(x,y)}{f_Y(y)}$ $(f_Y(y) > 0)$
確率の総和	$\sum_x \sum_y f_{X,Y}(x,y) = 1$	$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y)\, dx\, dy = 1$
CDF $F_{X,Y}(x,y)$	$\sum_{s \le x} \sum_{t \le y} f_{X,Y}(s,t)$	$\int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(s,t)\, ds\, dt$
単一点の確率	$P(X=x,\, Y=y) \ge 0$	$P(X=x,\, Y=y) \equiv 0$

- 多項分布 $\mathrm{Mult}(n, p_1, \ldots, p_k)$
    - n 回の独立試行で、各試行が k 個のカテゴリのいずれかに属する確率分布
    - 複数カテゴリにわたる成功回数の分布
        - 10 回サイコロを振ったときに、各目が出る回数の分布
        - アンケート調査で回答が複数の選択肢に分類される場合
- 多変量正規分布 $\mathrm{N}_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
    - n 次元の確率変数 $\mathbf{X} = (X_1, \ldots, X_n)^\top$ が従う連続分布で、平均ベクトル $\boldsymbol{\mu}$ と共分散行列 $\boldsymbol{\Sigma}$ によって完全に決定される分布
    - 複数連続変数間の線形関係や同時分布のモデル化
        - 身長・体重・血圧などの複数の生理的測定値の統計モデリング

積率・積率母関数（復習）

Def 5.1　モーメント・積率 moment

$X \sim F_X(x)$ とする．$n = 1, 2, \ldots$ に対して，

- $\mu'_n = E[X^n]$：X（または $F_X(x)$）の（原点まわりの）n 次モーメント the nth moment
- $\mu_n = E[(X-\mu)^n]$：X（または $F_X(x)$）の平均まわりの n 次モーメント the nth central moment

Def 5.2　積率母関数 moment-generating function, MGF

$X \sim F_X(x)$ とする．$\exists h > 0, \forall |t| < h,$

$$M_X(t) = E[e^{tX}]$$

が存在するとき，$M_X(t)$ を X の MGF という

積率母関数の性質

Thm 5.3　MGF の一意性

$X \sim F_X(x), Y \sim F_Y(y)$ 且つモーメントが存在存在するとする．

- $\exists h > 0\ \forall |t| < h$，$M_X(t) = M_Y(t) \Leftrightarrow F_X(u) = F_Y(u), \quad \forall u \in \mathbb{R}$
- X, Y が同じ有界区間 $[a, b]$ に入るとする、
  $E[X^n] = E[Y^n], \quad \forall n = 0, 1, 2, \cdots \Leftrightarrow F_X(u) = F_Y(u), \quad \forall u \in \mathbb{R}$

Ex 5.4　分布の近似

- 二項分布とポアソン分布の近似の証明 (Thm 1.7, Stat-note1 pp.7)
- 正規分布と二項分布の近似の証明 (Thm 2.14, Stat-note2 pp.7)

Thm 5.5　MGF の収束

多変量確率変数 $\{X_i, i = 1, 2, \ldots\}$ を考え，各 X_i が $M_{X_i}(t)$ を持つとする

$$\exists h > 0\ \forall |t| < h,\ \lim_{i\to\infty} M_{X_i}(t) = M_X(t)$$

$$\Leftrightarrow \lim_{i\to\infty} F_{X_i}(x) = F_X(x), \quad \forall F_X(x)$$

ここで，$F_X(x)$ はモーメント母関数 $M_X(t)$ を持つ分布関数である

Thm 5.6

確率変数 X, Y $(X \perp Y)$ を考える．それぞれが $M_X(t), M_Y(t)$ を持つとする．
$Z = X + Y$ の MGF

$$M_Z(t) = M_X(t)M_Y(t) \quad \text{(証)}$$

Ex 5.7　**正規確率変数の和の** MGF

$X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2), X \perp Y$ とする．$Z = X + Y$ の分布

$$Z \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2) \qquad \text{(証)}$$

Thm 5.8　Thm 5.6 の拡張

互いに独立な確率変数 $\{X_i, i = 1, \ldots, n\}$ を考える．それぞれが $M_{X_1}(t), \ldots, M_{X_n}(t)$ を持つとする．$Z = X_1 + \cdots + X_n$ の MGF

$$M_Z(t) = M_{X_1}(t) \cdots M_{X_n}(t)$$

$X_i, \ldots, X_n$ が同じ確率分布に従う (iid) 場合

$$M_Z(t) = (M_X(t))^n$$

Ex 5.9　ガンマ確率変数の和の MGF

$\{X_i, i = 1, \ldots, n\} \sim \ Gamma(\alpha_i, \beta), X_i \perp X_j (i \neq j)$ とする．$Z = \sum_{i=1}^{n} X_i$ の分布

$$Z \sim Gamma(\alpha_1 + \cdots + \alpha_n, \beta) \qquad \text{(証)}$$

Thm 5.10

$\forall$ 定数 a, b, $aX + b$ の MGF

$$M_{aX+b}(t) = e^{bt} M_X(at) \quad \text{(証)}$$

Thm 5.11　Thm 5.10 の拡張

互いに独立な確率変数 $\{X_i, i = 1, \ldots, n\}$ を考える．それぞれが $M_{X_1}(t), \ldots, M_{X_n}(t)$ を持つとする．$\forall$ 定数 $a_1, \ldots, a_n; b_1, \ldots, b_n$ $Z = (a_1 X_1 + b_1) + \cdots + (a_n X_n + b_n)$ の MGF

$$M_Z(t) = (e^{t \sum b_i}) M_{X_1}(a_1 t) \cdots M_{X_n}(a_n t) \quad \text{(証)}$$

Ex 5.12　正規確率変数の和の MGF–II

$\{X_i, i = 1, \ldots, n\} \sim N(\mu_i, \sigma_i^2)$, $X_i \perp X_j (i \neq j)$ とする．
$\forall$ 定数 $\{a_i, i = 1, \ldots, n\}, \{b_i, i = 1, \ldots, n\}, Z = \sum_{i=1}^n (a_i X_i + b_i)$ の分布

$$Z \sim N\left(\sum_{i=1}^n (a_i \mu_i + b_i), \sum_{i=1}^n a_i^2 \sigma_i^2\right)$$

確率不等式

Thm 5.13 チェビシェフの不等式 Chebyshev's inequality

X が確率変数且つ $g(x) \geq 0$ のとき，

$$\forall r > 0, \quad P(g(X) \geq r) \leq \frac{E[g(X)]}{r}$$ **(証)**

チェビシェフの不等式は、平均 $E[g(X)]$ から，確率変数の非負関数 $g(X)$ が何らかの正の定数以上である確率の上限（upper bound）を与える

Ex 5.14

$g(x) = \frac{(x-\mu)^2}{\sigma^2}, \mu = E[X], \sigma^2 = Var[X], r = t^2$ とする（分布不明）

$$P\left(\frac{(X-\mu)^2}{\sigma^2} \geq t^2\right) \leq \frac{1}{t^2} E[\frac{(X-\mu)^2}{\sigma^2}] = \frac{1}{t^2}$$

$\Rightarrow P(|X-\mu| \leq t\sigma) \geq 1 - \frac{1}{t^2}$

- $t = 2, P(|X-\mu| \leq 2\sigma) \geq 1 - \frac{1}{2^2} = 0.75$

Thm 5.15　マルコフの不等式 Markov's inequality

$P(Y \geq 0) = 1$ 且つ $P(Y = 0) < 1$ のとき，

$$\forall r > 0, \quad P(Y \geq r) \leq \frac{E[Y]}{r}$$

$P(Y = r) = \frac{E[Y]}{r} \Leftrightarrow P(Y = r) = p = 1 - P(Y = 0), \forall 0 < p \leq 1$

Ex 5.16　Ex5.14 を続ける

チェビシェフの不等式は保守的である．$Z \sim N(0, 1)$

$$P(|Z| \geq t) \leq \sqrt{\frac{2}{\pi}} \frac{e^{-t^2/2}}{t}, \quad \forall t > 0$$ **(証)**

- $t = 2, P(|Z| \geq 2) \leq 0.054 \Rightarrow P(|Z| \leq 2) \geq 0.946$
- 標準正規表 $\rightarrow P(|Z| \leq 2) \approx 0.9545$
- Ex5.14$\rightarrow$ $P(|Z| \geq 2) \leq 0.25$ (保守的)

$P(X \geq a) \leq e^{-at} M_X(t)$ **(証)**

数値不等式

Thm 5.17　ヘルダーの不等式 Holder's inequality (確率測度)

X, Y が確率変数，$\forall p > 0, q > 0$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$ のとき，

$$|E[XY]| \leq E[|XY|] \leq (E[|X|^p])^{1/p}(E[|Y|^q])^{1/q}$$ **(証)**

- Thm1.2-5.(Supp 1.4) $\to |E[XY]| \leq E[|XY|]$

Ex 5.18

- $Y \equiv 1 \Rightarrow E[|X|] \leq \{E[|X|^p]\}^{1/p}, \quad 1 < p < \infty$
- $|X| = |X|^r \Rightarrow E[|X|^r] \leq \{E[|X|^{pr}]\}^{1/p}, \quad 1 < r < p < \infty$
- $\{E[|X|^r]\}^{1/r} \leq \{E[|X|^s]\}^{1/s}, \quad 1 < r < s < \infty$ (Lyapounov's inequality)

Supp 5.19

$\forall a > 0, b > 0, p > 0, q > 0$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$ のとき，

$$\frac{1}{p}a^p + \frac{1}{q}b^q \geq ab$$

$\frac{1}{p}a^p + \frac{1}{q}b^q = ab \Leftrightarrow a^p = b^q$

Supp 5.20 **ヘルダーの不等式** Holder's inequality (**計数測度**)

実数列 $\{a_i, i = 1, \ldots, n\}, \{b_i, i = 1, \ldots, n\}, \forall p > 0, q > 0$ s.t. $\frac{1}{p} + \frac{1}{q} = 1$ のとき，

$$\sum_{i=1}^{n} |a_i b_i| \leq \left\{\sum_{i=1}^{n} a_i^p\right\}^{1/p} \left\{\sum_{i=1}^{n} b_i^q\right\}^{1/q}$$

- $b_i \equiv 1, p = q = 2 \Rightarrow \frac{1}{n}\left(\sum_{i=1}^{n} |a_i|\right)^2 \leq \sum_{i=1}^{n} a_i^2$

ヘルダーの不等式で，$p = q = 2$ のとき，

Thm 5.21　コーシー・シュワルツの不等式 Cauchy-Schwarz inequality

X, Y が確率変数，

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

Ex 5.22

- $\{Cov[X, Y]\}^2 \leq Var[X]Var[Y]$
- $\{E[(X - \mu_X)(Y - \mu_Y)]\}^2 \leq E[(X - \mu_X)^2]E[(Y - \mu_Y)^2]$

Thm 5.23　ミンコフスキーの不等式 Minkowski's inequality

X, Y が確率変数，

$$\forall 1 \leq p < \infty, \quad [E[|X+Y|^p]]^{1/p} \leq (E[|X|^p])^{1/p} + (E[|Y|^p])^{1/p}$$

(証)

関数不等式

Thm 5.24 イェンセンの不等式 Jensen's inequality

X が確率変数，$g(x)$ が凸関数（「下に凸な関数」convex）のとき，

$$E[g(X)] \geq g(E[X])$$ **（証）**

Supp 5.25 凸関数（convex function）

集合 $C \subset \mathbb{R}^n$ 上で定義された関数が 凸関数であるとは、任意の $x_1, x_2 \in C$ と任意の $\lambda \in [0,1]$ に対して次の不等式が成り立つときにいう：

$$f\big(\lambda x_1 + (1-\lambda)x_2\big) \leq \lambda f(x_1) + (1-\lambda)f(x_2).$$

Ex 5.26

- 算術平均, Arithmetic Mean: AM

$$\mathrm{AM} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- 幾何平均, Geometric Mean: GM, 比率の変化や成長率の平均

$$\mathrm{GM} = \sqrt[n]{x_1 x_2 \ldots x_n} = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}$$

- 調和平均, Harmonic Mean: HM, 速度や密度など、「単位あたり」の値の平均

$$\mathrm{HM} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \cdots + \frac{1}{x_n}} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

$$AM \geq GM \geq HM$$

(課題)

レポート

課題 5.27

Ex 5.26 の結論：

$$AM \geq GM \geq HM$$

を示せ

See BEEF+ Report 6.

6. 確率変数の変換

2025 後期 3H606 数理科学入門 (統計系)–#8,#9

Nov 25, Dec 9, 2025

数理環境論 Room 742
Email: ZHOU YI <zhouy@people.kobe-u.ac.jp>

キーワード

- 逆関数 Inverse function, inverse transformation
- 変数変換 Variable transformation
    - 変換された変数の分布の例
- ２変量変数変換
    - 変換された変数の分布の例

単変量の変数変換

$$g : \mathcal{X} \to \mathcal{Y}$$

$$Y = g(X)$$

- $\mathcal{X} = \{x : f_X(x) > 0\}$
- $\mathcal{Y} = \{y : y = g(x), \exists x \in \mathcal{X}\}$

逆関数

$X \sim F_X(x)$ とする，$\forall g(.), g(X)$ は確率変数

$Y = g(X)$ と定義すると，Y は X の関数となる

$\forall$集合A，Y の分布

$$P(Y \in A) = P(g(X) \in A)$$

は関数 $F_X(x), g(.)$ に依存する

$\mathcal{X}, \mathcal{Y}$ は X, Y の標本空間，関数 $y = g(x)$ とする，

$$\text{元関数・変換} \quad g : \mathcal{X} \to \mathcal{Y}$$

$$\text{逆関数・変換} \quad g^{-1} : \mathcal{Y} \to \mathcal{X}$$

Ex 6.1　逆関数

- $g(x) = x^2 \quad (x \geq 0) \Leftrightarrow g^{-1}(y) = \sqrt{y} \quad (y \geq 0)$
- $g(x) = e^x \quad$ (定義域: $\mathbb{R}$, 値域: $y > 0$) $\Leftrightarrow g^{-1}(y) = \ln(y) \quad$ (定義域: $y > 0$, 値域: $\mathbb{R}$)

変数変換

- $\forall$集合$A, g^{-1}(A) = \{x \in \mathcal{X} : g(x) \in A\}$
- $\forall$点集合$A = \{y\}, g^{-1}(\{y\}) = \{x \in \mathcal{X} : g(x) = y\}$
  - $g(x) = y$ となる x が 1 つだけ存在する場合，$g^{-1}(y) = \{x\}, g^{-1}(y) = x$

$Y = g(X)$ とする．$\forall$集合$A \subset \mathcal{Y}$，

$$
\begin{aligned}
P(Y \in A) &= P(g(X) \in A) \\
&= P(\{x \in \mathcal{X} : g(x) \in A\}) \\
&= P(X \in g^{-1}(A))
\end{aligned}
$$

$Y = g(X)$ の CDF

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) = P(g(X) \leq y) \\
&= P(\{x \in \mathcal{X} : g(x) \leq y\}) \\
&= \int_{x \in \mathcal{X} : g(x) \leq y} f_X(x)dx \quad \text{（連続）} \\
&= \sum_{x \in \mathcal{X} : g(x) \leq y} f_X(x) \quad \text{（離散）}
\end{aligned}
$$

- 変数変換を行う際、$\mathcal{X}$ と $\mathcal{Y}$ との対応関係（写像関係）を維持することが重要である
- 狭義単調関数は、最も簡単に扱える

$Y = g(X)$ とする．

- $\mathcal{X} = \{x : f_X(x) > 0\}$
- $\mathcal{Y} = \{y : y = g(x), \exists x \in \mathcal{X}\}$

$g : x \to y$ が狭義単調 (strict monotone) ならば，$\mathcal{X} \to \mathcal{Y}$ 全単射 (one-to-one and onto) であり，逆関数 g^{-1} が存在する

- 任意の x がちょうど 1 つの y に対応し、また、任意の y には、多くとも 1 つの x しか対応しない

Supp 6.2 **単調** (Monotone)

- 狭義単調増加 (Strictly Increasing): $x_1 < x_2 \Rightarrow g(x_1) < g(x_2)$
- 単調増加: $x_1 < x_2 \Rightarrow g(x_1) \leq g(x_2)$
- 狭義単調減少 (Strictly Decreasing): $x_1 < x_2 \Rightarrow g(x_1) > g(x_2)$
- 単調減少: $x_1 < x_2 \Rightarrow g(x_1) \geq g(x_2)$

狭義単調：異なる入力には、必ず異なる出力が対応する．関数が単射 (one-to-one) であることを保証する

$$\Rightarrow g^{-1}(y) = x \Leftrightarrow g(x) = y$$

$y = g(x)$ が狭義増加ならば，

- 標本区間

$$\{x \in \mathcal{X} : g(x) \leq y\} = \{x \in \mathcal{X} : g^{-1}(g(x)) \leq g^{-1}(y)\} = \{x \in \mathcal{X} : x \leq g^{-1}(y)\}$$

- CDF

$$F_Y(y) = \int_{x \in \mathcal{X} : x \leq g^{-1}(y)} f_X(x) dx = \int_{-\infty}^{g^{-1}(y)} f_X(x) dx = F_X(g^{-1}(y))$$

$y = g(x)$ が狭義減少ならば，

- 標本区間

$$\{x \in \mathcal{X} : g(x) \leq y\} = \{x \in \mathcal{X} : g^{-1}(g(x)) \geq g^{-1}(y)\} = \{x \in \mathcal{X} : x \geq g^{-1}(y)\}$$

- CDF

$$F_Y(y) = \int_{x \in \mathcal{X} : x \geq g^{-1}(y)} f_X(x) dx = \int_{g^{-1}(y)}^{\infty} f_X(x) dx = 1 - F_X(g^{-1}(y))$$

Thm 6.3

$X \sim F_X(x), Y = g(X)$ とする．

$\mathcal{X} = \{x : f_X(x) > 0\}, \mathcal{Y} = \{y : y = g(x), \exists x \in \mathcal{X}\}$

1. g が $\mathcal{X}$ 上の狭義増加関数 $\Rightarrow F_Y(y) = F_X(g^{-1}(y)), \quad y \in \mathcal{Y}$
2. g が $\mathcal{X}$ 上の狭義減少関数且つ X が連続変数
   $\Rightarrow F_Y(y) = 1 - F_X(g^{-1}(y)), \quad y \in \mathcal{Y}$

g が $\mathcal{X}$ 上の狭義減少関数且つ X が離散変数である場合，$\mathcal{X}, \mathcal{Y}$ は可算集合，Y は離散変数，CDF を考える

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(X \geq g^{-1}(y)) = 1 - P(X < g^{-1}(y))$$

- g^{-1} が g のグラフを直線 $y = x$ に関して対称に移動したもの
- 関数 g とその逆関数 g^{-1} は、常に同じ種類の単調性を持つ

PMF, PDF

- X が離散変数である場合，$Y = g(X)$ の PMF

$$f_Y(y) = P(Y = y) = \sum_{x \in g^{-1}(y)} P(X = x) = \sum_{x \in g^{-1}(y)} f_X(x), \quad y \in \mathcal{Y}$$

- X が連続変数である場合，$Y = g(X)$ の PDF
  - g が狭義増加関数

$$f_Y(y) = \frac{d}{dy} F_Y(y) = f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$$

  - g が狭義減少関数

$$f_Y(y) = \frac{d}{dy} F_Y(y) = -f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y)$$

Thm 6.4

$X \sim F_X(x)$, $Y = g(X)$
$\mathcal{X} = \{x : f_X(x) > 0\}, \mathcal{Y} = \{y : y = g(x), \exists x \in X\}$
$f_X(x)$ が $\mathcal{X}$ 上の連続関数，$g^{-1}(y)$ が $\mathcal{Y}$ に対して連続的な微分（continuous derivative）を持つとする．

$$f_Y(y) = \frac{d}{dy} F_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right|, \quad y \in \mathcal{Y}$$

$f_Y(y) = 0, y \notin \mathcal{Y}$

- $x = g^{-1}(y)$
- $\left| \dfrac{d}{dy} g^{-1}(y) \right|$: ヤコビアン（Jacobian）の絶対値，単変量の場合

変数変換の例

Ex 6.5　二項変換 Binomial transformation

$X \sim Bin(n, p)$

$$f_X(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n$$

$g(x) = n - x \Rightarrow Y = n - X$

$$\mathcal{X} = \{0, 1, \ldots, n\} \Rightarrow \mathcal{Y} = \{y : y = g(x), x \in \mathcal{X}\} = \{0, 1, \ldots, n\}$$

$$f_Y(y) = \binom{n}{y} (1-p)^y p^{n-y} \Leftrightarrow Y \sim Bin(n, 1-p) \quad \text{（証）}$$

Ex 6.6　一様一指数変換 Uniform-exponential transformation

$X \sim Unif(0,1)$

$$f_X(x) = 1, \quad 0 < x < 1 \Rightarrow F_X(x) = x$$

$g(x) = -log(x) \Rightarrow Y = -log(X)$

$$\mathcal{X} = \{x : 0 < x < 1\} \Rightarrow \mathcal{Y} = \{y : y > 0\}$$

$$F_Y(y) = 1 - e^{-y}, \quad y > 0$$ **(証)**

Ex 6.7 **逆ガンマ** inverted Gamma PDF

$X \sim Gamma(n, \beta)$

$$f_X(x) = \frac{1}{(n-1!)\beta^n} x^{n-1} e^{-x/\beta}, \quad 0 < x < \infty, \beta > 0, n \in \mathbb{Z}^+$$

$g(x) = 1/x \Rightarrow Y = 1/X$

$$\mathcal{X} = \{x : 0 < x < \infty\} \Rightarrow \mathcal{Y} = \{y : 0 < y < \infty\}$$

$$f_Y(y) = \frac{1}{(n-1!)\beta^n} \left(\frac{1}{y}\right)^{n+1} e^{-1/(y\beta)}, \quad y > 0$$ **(証)**

Thm6.3-6.4 が直接使用できない非単調関数の例

Ex 6.8　平方変換 Square transformation

連続変数 $X \sim f_X(x)$
$g(x) = x^2 \Rightarrow Y = X^2$

$$\mathcal{X} = \{x : -\infty < x < \infty\} \Rightarrow \mathcal{Y} = \{y : y > 0\}$$

$$F_y(y) = F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

$$f_Y(y) = \frac{1}{2\sqrt{y}}\left(f_X(\sqrt{y}) + f_X(-\sqrt{y})\right), \quad y > 0$$ **(証)**

Ex 6.9　正規ーカイ平方変換 Normal-chi square transformation

$X \sim N(0,1)$

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

$g(x) = x^2 \Rightarrow Y = X^2$

$$\mathcal{X} = \{x : -\infty < x < \infty\} \Rightarrow \mathcal{Y} = \{y : y > 0\}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{y}} e^{-y/2}, \quad y > 0$$ **(証)**

$$\Leftrightarrow Y \sim \chi^2(1)$$

Supp 6.10 非単調関数の場合

$X \sim F_X(x)$, $Y = g(X)$
$\mathcal{X} = \{x : f_X(x) > 0\}, \mathcal{Y} = \{y : y = g(x), \exists x \in \mathcal{X}\}$ $\mathcal{X}$ の分割 $A_0, A_1, \ldots, A_k$ が存在し、
$P(X \in A_0) = 0$ 且つ $f_X(x)$ が各分割で連続であるとする．$\exists g_1(x), \ldots, g_k(x)$ s.t.

1. $g(x) = g_i(x), \quad x \in A_i$
2. $g_i(x)$ が A_i 上の狭義単調関数
3. $\forall i = 1, \ldots, k, \mathcal{Y} = \{y * y = g_i(x), \exists x \in A_i\}$
4. $\forall i = 1, \ldots, k, g_i^{-1}(y)$ が $\mathcal{Y}$ に対して連続的な微分を持つ

$$f_Y(y) = \sum_{i=1}^{k} f_X(g_i^{-1}(y)) \left| \frac{d}{dy} g_i^{-1}(y) \right|, \quad y \in \mathcal{Y}$$

$f_Y(y) = 0, y \notin \mathcal{Y}$

Thm 6.11　確率積分変換 Probability Integral Transform

連続変数 $X \sim F_X(x)$, $Y = F_X(x)$

$$\Rightarrow Y \sim Unif(0,1), \quad 0 < y < 1 \quad \text{(証)}$$

$$F_Y(y) = y, \quad 0 < y < 1$$

- 任意の分布に従う乱数を生成する手法
- まず $(0,1)$ 上の一様乱数を生成し、それを目標とする分布の累積分布関数の逆関数に代入することで、目標の分布に従う乱数を得る
- シミュレーション、モンテカルロ法、統計モデリングなど、計算機上で任意の確率分布を再現するために不可欠な理論的基盤を提供する

多変量の変数変換

$$U = g_1(X, Y), V = g_2(X, Y)$$

$$g_1, g_2 : \mathcal{A} \to \mathcal{B}$$

- $\mathcal{A} = \{(x, y) : f_{X,Y}(x, y) > 0\}$
- $\mathcal{B} = \{(u, v) : u = g_1(x, y), v = g_2(x, y), \exists (x, y) \in \mathcal{A}\}$

変数変換の分布

２変量変数 $(X, Y) \sim f_{X,Y}(x, y), g_1(x, y), g_2(x, y)$ とする．
変数 (U, V) を定義し, s.t. $U = g_1(X, Y), V = g_2(X, Y)$

$\forall$集合$B \in \mathbb{R}^2$, $\Rightarrow$

1. $(U, V) \in B \Leftrightarrow (X, Y) \in A = \{(x, y) : (g_1(x, y), g_2(x, y)) \in B\}$
2. $P((U, V) \in B) = P((X, Y) \in A)$

・(U, V) の分布は (X, Y) によって完全に決定される

離散変数

(X, Y) が離散変数である場合，

- 標本空間（$\mathcal{A}$）は可算集合

$$\mathcal{A} = \{(x, y) : f_{X,Y}(x, y) = P(X = x, Y = y) > 0\}, \quad (x, y) \in \mathcal{A}$$

(U, V) も離散変数

- 標本空間（$\mathcal{B}$）も可算集合[1]

$$\mathcal{B} = \{(u, v) : u = g_1(x, y) \wedge v = g_2(x, y), \exists (x, y) \in \mathcal{A}\}, \quad (u, v) \in \mathcal{B}$$

A_{uv} を導入する

$$A_{uv} = \{(x, y) \in \mathcal{A} : g_1(x, y) = u \wedge g_2(x, y) = v\}, \quad (x, y) \in A_{uv}$$

$$\Rightarrow f_{U,V}(u, v) = P(U = u, V = v) = P((X, Y) \in A_{uv}) = \sum_{(x,y) \in A_{uv}} f_{X,Y}(x, y)$$

[1] $u \wedge v$: and 且つ (u, v) と表記される場合がある

Ex 6.12 ポアソン変数の和の分布 (証)

$X \sim Poisson(\theta),\, Y \sim Poisson(\lambda),\, X \perp Y$

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) = \frac{\theta^x e^{-\theta}}{x!} \frac{\lambda^y e^{-\lambda}}{y!}, \quad x = 0, 1, \ldots ; y = 0, 1, \ldots$$

$$\mathcal{A} = \{(x,y) : x = 0, 1, \cdots \wedge y = 0, 1, \ldots\}$$

$g_1(x,y) = x + y,\, g_2(x,y) = y \Rightarrow U = X + Y,\, V = Y$

$$\mathcal{B} = \{(u,v) : v = 0, 1, \cdots \wedge u = v, v+1, v+2, \ldots\}$$

$$f_{U,V}(u,v) = f_{X,Y}(u-v, v) = \frac{\theta^{u-v} e^{-\theta}}{(u-v)!} \frac{\lambda^v e^{-v}}{v!}, \quad v = 0, 1, \ldots ; u = v, v+1, \ldots$$

$$f_U(u) = \frac{e^{-(\theta+\lambda)}}{u!} (\theta+\lambda)^u, \quad u = 0, 1, 2, \cdots \Leftrightarrow U \sim Poisson(\theta+\lambda)$$

$$f_V(v) = f_Y(y)$$

Thm 6.13

$X \sim Poisson(\theta),\, Y \sim Poisson(\lambda),\, X \perp Y$

$$X + Y \sim Poisson(\theta + \lambda)$$

連続変数

(X, Y) が連続変数である場合，

- 標本空間

$$\mathcal{A} = \{(x, y) : f_{X,Y}(x, y) > 0\}, \quad (x, y) \in \mathcal{A}$$

(U, V) も連続変数

- 標本空間

$$\mathcal{B} = \{(u, v) : u = g_1(x, y) \wedge v = g_2(x, y), \exists (x, y) \in \mathcal{A}\}, \quad (u, v) \in \mathcal{B}$$

$$\Rightarrow f_{U,V}(u, v) > 0, \quad \forall (u, v) \in \mathcal{B}$$

最も簡単な場合として，g_1, g_2 を、$\mathcal{A} \leftrightarrow \mathcal{B}$ を定義している[2]関数とする $\Rightarrow$

- $\forall(u,v) \in \mathcal{B}, \exists!(x,y) \in \mathcal{A}$ s.t.[3] $(u,v) = (g_1(x,y), g_2(x,y))$
- 逆変換：$x = h_1(u,v), y = h_2(u,v)$ $\quad$ ($\rightarrow g^{-1}(y)$ 単変量の場合)

ヤコビアン Jacobian

$$J = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \left(\frac{\partial x}{\partial u} \cdot \frac{\partial y}{\partial v}\right) - \left(\frac{\partial x}{\partial v} \cdot \frac{\partial y}{\partial u}\right)$$

- $\frac{\partial x}{\partial u} = \frac{\partial h_1(u,v)}{\partial u}, \frac{\partial y}{\partial u} = \frac{\partial h_2(u,v)}{\partial u}, \frac{\partial x}{\partial v} = \frac{\partial h_1(u,v)}{\partial v}, \frac{\partial y}{\partial v} = \frac{\partial h_2(u,v)}{\partial v}$
- ($\rightarrow \frac{dx}{dy} = \frac{dg^{-1}(y)}{dy}$ 単変量の場合)

$$f_{U,V}(u,v) = f_{X,Y}(h_1(u,v), h_2(u,v))|J|, \quad J \neq 0, \forall(u,v) \in \mathcal{B}$$

($\rightarrow$Thm6.4 単変量の場合)

---

[2] $\mathcal{A} \leftrightarrow \mathcal{B}$ $\mathcal{A}$ から $\mathcal{B}$ への全単射

[3] $\exists!$: ただ一つだけ存在する; !: Unique/Exactly one ただ一つだけ

Ex 6.14　ベータ変数の積の分布　（証）

$X \sim Beta(\alpha, \beta)$, $Y \sim Beta(\alpha+\beta, \gamma)$, $X \perp Y$

$$\begin{aligned} f_{X,Y}(x,y) &= f_X(x) f_Y(y) \\ &= \left[\frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}\right] \cdot \left[\frac{1}{B(\alpha+\beta,\gamma)} y^{\alpha+\beta-1}(1-y)^{\gamma-1}\right] \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha+\beta)\Gamma(\gamma)} y^{\alpha+\beta-1}(1-y)^{\gamma-1} \end{aligned}$$

$$\mathcal{A} = \{(x,y) : 0 < x < 1 \wedge 0 < y < 1\}$$

$g_1(x,y) = xy, g_2(x,y) = x \Rightarrow U = XY, V = X$

$$\mathcal{B} = \{(u,v) : 0 < u < v < 1\}$$

$$g_1, g_2 : \mathcal{A} \leftrightarrow \mathcal{B} \Rightarrow x = h_1(u,v) = v, y = h_2(u,v) = u/v$$

$$f_{U,V}(u,v) = \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)} v^{\alpha-1}(1-v)^{\beta-1}(\frac{u}{v})^{\alpha+\beta-1}(1-\frac{u}{v})^{\gamma-1}(\frac{1}{v})$$

$$f_U(u) = \frac{\Gamma(\alpha+\beta+\gamma)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma)} u^{\alpha-1}(1-u)^{\beta+\gamma-1}, \quad 0 < u < 1 \Leftrightarrow U \sim Beta(\alpha, \beta+\gamma)$$

$$f_V(v) = f_X(x)$$

Ex 6.15　正規変数の和と差の分布　（証）

$X \sim N(0,1), Y \sim N(0,1), X \perp Y$

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) = \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right) \cdot \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \right)$$

$$\mathcal{A} = \mathbb{R}^2 = \{(x,y) : -\infty < x < \infty \wedge -\infty < y < \infty\}$$

$g_1(x,y) = x + y, g_2(x,y) = x - y \Rightarrow U = X + Y, V = X - Y$

$$\Rightarrow x = h_1(u,v) = \frac{u+v}{2}, y = h_2(u,v) = \frac{u-v}{2}$$

$$\mathcal{B} = \mathbb{R}^2$$

$$f_{U,V}(u,v) = \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-u^2/4} \cdot \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-v^2/4}$$

$$f_U(u) = \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-u^2/4} \Leftrightarrow U \sim N(0,2)$$

$$f_V(v) = \frac{1}{\sqrt{2\pi}\sqrt{2}} e^{-v^2/4} \Leftrightarrow V \sim N(0,2)$$

変数変換の独立性

Thm 6.16

$g(x)$ を x のみの関数，$h(y)$ を y のみの関数とする

$$\forall U = g(X), V = h(Y), \quad X \perp Y \Rightarrow U \perp V$$ **（証）**

単全射ではない例

（Ex 6.8-6.9, Supp 6.10 の拡張）

Ex 6.17　正規変数の比率の分布　（証）

$X \sim N(0,1)$, $Y \sim N(0,1)$, $X \perp Y$

$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) = \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\right) \cdot \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\right)$$

$$\mathcal{A} = \mathbb{R}^2 = \{(x,y) : -\infty < x < \infty \land -\infty < y < \infty\}$$

$g_1(x,y) = x/y$, $g_2(x,y) = |y| \Rightarrow U = X/Y$, $V = |Y|$

$$\Rightarrow \mathcal{B} = \{(u,v) : v > 0 \land -\infty < u < \infty\}$$

$$f_{U,V}(u,v) = \frac{v}{\pi} e^{-(u^2+1)v^2/2}$$

$$f_U(u) = \frac{1}{\pi(u^2+1)}, \quad -\infty < u < \infty \Leftrightarrow U \sim Cauchy(0,1)$$

Supp 6.18　コーシー分布（Cauchy distribution）

実数値の確率変数 X が、位置パラメータ $x_0 \in \mathbb{R}$、尺度パラメータ $\gamma > 0$ に対して

$$f_X(x) = \frac{1}{\pi\gamma}\frac{1}{1+\left(\frac{x-x_0}{\gamma}\right)^2}, \quad x \in \mathbb{R}$$

という確率密度関数を持つとき，$X \sim Cauchy(0,1)$

レポート

課題 6.19

BEEF+ Report7 をご参照ください

7. 標本分布

2025 後期 3H606 数理科学入門 (統計系)–#10

Dec 16, 2025

数理環境論 Room 742
Email: ZHOU YI <zhouy@people.kobe-u.ac.jp>

Statistics

- Neo-Latin **statisticum collegium** (“council of state” 国務院) and the Italian word **statista** (“statesman” or “politician” 政治家)
- German: Statistik (“description of a state, a country”)

”Statistics is both **the science of uncertainty** and **the technology of extracting information from data**.”

–International Encyclopedia of Statistical Science

キーワード

- 無作為標本 (random sample) IID 標本
- 単純無作為標本 (simple random sample)
- 統計量
- 標本分布
- 正規分布からの標本分布

無作為抽出

Ex 7.1　データ収集モデル

コインを投げる試行において，$X =$「H が出る」，$X \sim f(x)$ とする ($\{x : x \in \{0, 1\}\}$)

この試行を M 回独立に繰り返す

- 各観測で得られる結果は，確率変数 $X_1, X_2, \ldots, X_M$ としてモデル化できる
- 観察した結果は，実現値・データ $(x_1, x_2, \ldots, x_M)$ として収集される

Def 7.2　無作為標本 random sample，IID 標本

1. $X_1, \ldots, X_n$ が相互に独立した確率変数: $X_i \perp X_j$ $(i \neq j \in \{1, \ldots, n\})$
2. 各 X_i の周辺 PMF/PDF は $f(x)$ と同じ: $f_{X_i}(x) = f(x)$

$X_1, \ldots, X_n$ を母集団 $f(x)$ からのサイズ n の無作為標本，或いは PMF/PDF が $f(x)$ を持つ IID 標本（確率変数）と呼ぶ

IID: independently and identically distributed 独立同分布，互いに独立で同一分布に従う

無作為標本

母集団 $f(x)$ から要素を抽出する試行を考え，$X_1, \ldots, X_n$ を $f(x)$ から無作為に抽出したサイズ n の標本（n 回繰り返す）とする。

母集団 $f(x)$

- 無限母集団 infinite population: 要素が無限の母集団 $\{x_1, x_2, \ldots\}$
  - １回目の抽出 $X_1 = x_i$ は母集団からの無作為標本
  - $X_1 = x_i$ を削除/復元しても２回目の抽出 $X_2 = x_j$ には影響しない $\rightarrow X_2 = x_j$ も母集団からの無作為標本
  - $X_1 = x_i, X_2 = x_j$ を削除/復元しても $X_3 = x_k$ には影響しない $\rightarrow X_3 = x_k$ も母集団からの無作為標本
  - $\ldots$
  - $\cdots \rightarrow X_n = x_n$ も母集団からの無作為標本

$\Rightarrow$ Def7.2 の条件が満たされる (IID 標本)

- 有限母集団 finite population: 要素が有限の母集団 $\{x_1, x_2, \ldots, x_N\}$

1. 復元抽出 sampling with replacement：抽出した要素を戻してから再度抽出する
    - １回目の抽出 $X_1 = x_i$ の確率 $= \frac{1}{N}$
    - ２回目の抽出 $X_2 = x_j$ の確率 $= \frac{1}{N}$
    - 繰り返し，$X_n = x$ の確率 $= \frac{1}{N}$

    $\Rightarrow$ Def7.2 の条件が満たされる (IID 標本)
2. 非復元抽出 sampling without replacement：抽出した要素を戻さずに再度抽出する
    - １回目の抽出 $X_1 = x_i$ の確率 $= \frac{1}{N}$
    - ２回目の抽出 $X_2 = x_j$ の確率 $= \frac{1}{N-1}$

    $\Rightarrow$ Def7.2 の条件が満たされる？

非復元抽出 $\forall x_i, x_j (i \neq j) \in \{x_1, x_2, \ldots, x_N\}$

1. $P(X_2 = x_i | X_1 = x_i) = 0, P(X_2 = x_j | X_1 = x_i) = \frac{1}{N-1} \Rightarrow X_1 \not\perp X_2$
2. $\forall x_i, P(X_1 = x_i) = P(X_1 = x) = \frac{1}{N}$

$$
\begin{aligned}
P(X_2 = x) &= \sum_{i=1}^{N} P(X_2 = x, X_1 = x_i) \\
&= \sum_{i=1}^{N} P(X_2 = x | X_1 = x_i) P(X_1 = x_i) \\
&= (N-1) \frac{1}{N-1} \frac{1}{N} + 0 = \frac{1}{N}
\end{aligned}
$$

他の X_i にも同様

単純無作為標本

有限母集団から非復元抽出を単純無作為抽出 simple random sampling (SRS) という (SR 標本 $\neq$IID 標本)

- $N \to \infty, P(X_2 = x_j | X_1 = x_i) = \frac{1}{N-1} \approx \frac{1}{N} \Rightarrow X_2 \perp X_1$[1]
- $N \to \infty, P(X_i = x | X_1, X_2, \ldots, X_{i-1}) = \frac{1}{N-i+1} \approx \frac{1}{N} \Rightarrow X_2 \perp X_1$
- $N \to \infty$, SR 標本 $\to$IID 標本

Ex 7.3　有限母集団モデル

有限母集団 $\{1, \ldots, 1000\}, (N = 1000)$ からサイズ $n = 10$ の標本を非復元且つ等確率で抽出するとする，$P(X_1 > 200, \ldots, X_{10} > 200) =?$

1. $N = 1000 \Rightarrow X_i \perp X_j \ (i \neq j \in \{1, \ldots, n\})$ と近似する
   $P(X_1 > 200, \ldots, X_{10} > 200) = P(X_1 > 200) \cdots P(X_{10} > 200) = \left(\frac{800}{1000}\right)^{10} \approx 0.107374$
2. 正確な確率: $Y = \#\{>200\}, Y \sim Hypergeometric(N = 1000, M = 800, K = 10)$
   $P(X_1 > 200, \ldots, X_{10} > 200) = P(Y = 10) = \frac{\binom{800}{10}\binom{200}{0}}{\binom{1000}{10}} \approx 0.106164$

[1] $\to \infty$: 無限大に近づくとき（有限であるが、その極限として $\to \infty$ 考える）

統計量

Def 7.4 統計量 Statistic

$X_1, \ldots, X_n$ がサイズ n の無作為標本，$T(x_1, \ldots, x_n)$ が実数値関数 (T の領域は $X_1, \ldots, X_n$ の標本空間を含む) とする

- 確率変数 $Y = T(X_1, \ldots, X_n)$ を統計量という
- Y の確率分布を標本分布 sampling distribution という

唯一の制限：統計量がパラメータを含まない

Ex 7.5 統計量の例

確率変数

- 標本平均 sample mean: $\bar{X} = T(X_1, \ldots, X_n) = \frac{1}{n} \sum_1^n X_i$
- 標本分散 sample variance: $S^2 = T(X_1, \ldots, X_n) = \frac{1}{n-1} \sum_1^n (X_i - \bar{X})^2$
- 標本 (標準) 偏差 sample standard deviation: $S = \sqrt{S^2}$

それぞれの観測値: $\bar{x}, s^2, s$

標本の和の分布

Supp 7.6 (証)

$\forall x_1, \ldots, x_n \in \mathbb{R}, \bar{x} = (x_1 + \cdots + x_n)/n,$

- $\min_a \sum_{i=1}^n (x_i - a)^2 = \sum_{i=1}^n (x_i - \bar{x})^2$
- $(n-1)s^2 = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n (x_i^2) - n\bar{x}$

IID 標本 $X_1, \ldots, X_n$, $\forall g(x)\ s.t. E[g(X_1)], Var[g(X_1)]$ が存在する

- $E[\sum_{i=1}^n g(X_i)] = n(E[g(X_1)])$
- $Var[\sum_{i=1}^n g(X_i)] = n(Var[g(X_1)])$

Thm 7.7 (証)

$X_1, \ldots, X_n$ を平均が μ 分散が $\sigma^2 < \infty$[2]を持つ母集団からの IID 標本 (無作為標本) とする

1. $E[\bar{X}] = \mu$ (不偏性)
2. $Var[\bar{X}] = \frac{\sigma^2}{n}$
3. $E[S^2] = \sigma^2$(不偏性)

[2] $\sigma^2 < \infty$: 分散 σ^2 が存在する

Thm 7.8 **(証)**

$X_1, \ldots, X_n$ を MGF が $M_X(t)$ を持つ母集団からの IID 標本とする

$$M_{\bar{X}}(t) = \{M_X(t/n)\}^n$$

Ex 7.9 **(証)**

- $X_1, \ldots, X_n$ を $N(\mu, \sigma^2)$ 母集団からの IID 標本とする

$$M_{\bar{X}}(t) = exp(\mu t + \frac{(\sigma^2/n)t^2}{2}) \Leftrightarrow \bar{X} \sim N(\mu, \frac{\sigma^2}{n})$$

- $X_1, \ldots, X_n$ を $Gamma(\alpha, \beta)$ 母集団からの IID 標本とする

$$M_{\bar{X}}(t) = (\frac{1}{1-(\beta/n)t})^{n\alpha} \Leftrightarrow \bar{X} \sim Gamma(n\alpha, \beta/n)$$

正規分布からの標本分布

Thm 7.10　(証)

$X_1, \ldots, X_n$ を $N(\mu, \sigma^2)$ 母集団からの IID 標本とする
$\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i, S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$

- $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ (Ex.7.9)
- $\bar{X} \perp S^2$
- $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{p=n-1} = \chi^2_{n-1}$

$$f(x;p) = \frac{1}{2^{p/2}\Gamma(p/2)} x^{p/2-1} e^{-x/2}, \quad 0 < x < \infty, p: \text{自由度 (Ex2.25)}$$

- $\chi^2_p = Gamma(p/2, 2)$
- $Z \sim N(0,1) \Rightarrow Z^2 \sim \chi^2_1$ (Ex 6.9)
- 独立である $X_1, \ldots, X_n, X_i \sim \chi^2_{p_i} \Rightarrow \sum_{i=1}^n X_i \sim \chi^2_{\sum_{i=1}^n p_i}$

Thm 7.11　**自由度 p の T 分布** Student's t distribution　**(証)**

$X_1, \ldots, X_n$ が $N(\mu, \sigma^2)$ 母集団からの IID 標本とする

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{p=n-1}$$

$$f_T(t;p) = \frac{\Gamma(\frac{p+1}{2})}{\Gamma(\frac{p}{2})} \frac{1}{(p\pi)^{1/2}} \frac{1}{(1+t^2/p)^{(p+1)/2}} - \infty < x < \infty$$

- $\frac{\bar{X}-\mu}{S/\sqrt{n}} = \frac{(\bar{X}-\mu)/(\sigma/\sqrt{n})}{\sqrt{S/\sigma^2}} = \frac{U}{\sqrt{V/p}}, \quad U \sim N(0,1),\, V \sim \chi^2_p,\, U \perp V$
- $p = 1,\, T \sim Cauchy$
- $E[T_p] = 0\ (p > 1),\, Var[T_p] = \frac{p}{p-2}\ (p > 2)$

Thm 7.12 **自由度 p,q の F 分布** Snedecor's F distribution

$X_1, \ldots, X_n$ が $N(\mu_X, \sigma_X^2)$ 母集団からの IID 標本，
$Y_1, \ldots, Y_m$ が $N(\mu_Y, \sigma_Y^2)$ 母集団からの IID 標本とする

$$F = \frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} \sim F_{p=n-1, q=m-1}$$

$$f_F(x;p,q) = \frac{\Gamma(\frac{p+q}{2})}{\Gamma(\frac{p}{2})\Gamma(\frac{q}{2})} (\frac{p}{q})^{p/2} \frac{x^{(p/2)-1}}{(1+(p/q)x)^{(p+1)/2}}, \quad 0 < x < \infty$$

- $\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} = \frac{U/p}{V/p}, \quad U \sim \chi_p^2, V \sim \chi_q^2, U \perp V$
- $p = 1, F \sim Cauchy$
- $E[F_{n-1,m-1}] = \frac{m-1}{m-3}\ (m > 3); \quad m \to \infty, F \approx 1$
- $X \sim F_{p,q} \Rightarrow 1/X \sim F_{q,p}$
- $X \sim t_p \Rightarrow X^2 \sim F_{1,q}$
- $X \sim F_{p,q} \Rightarrow \frac{(p/q)X}{1+(p/q)X} \sim Beta(p/2, q/2)$

レポート

課題 7.13

BEEF+ Report8 をご参照ください

8. 標本の収束とその応用

2025 後期 3H606 数理科学入門 (統計系)–#11

Dec 23, 2025

数理環境論 Room 742
Email: ZHOU YI <zhouy@people.kobe-u.ac.jp>

キーワード

- 確率収束 Convergence in Probability
    - 大数の弱法則 Weak Law of Large Number (WLLN)
- 概収束 Almost Sure convergence
    - 大数の強法則 Strong Law of Large Number (SLLN)
- 分布収束 Convergence in Distribution
    - 中心極限定理 Central Limit Theorem (CLT)

復習：数列の極限・収束

Supp 8.1　数列の極限（Limit of a sequence）

数例（実数の列）$\{a_n\}_{n=1}^{\infty}$ と $a \in \mathbb{R}$ に対して，

$$\lim_{n\to\infty} a_n = a$$

$$(a_n \to a \ (n \to \infty))$$

$$(|a_n - a| \to 0 \ (n \to \infty))$$

であるとき，数列 $\{a_n\}$ が実数 a に収束するという[1]

$\epsilon - N$ (イプシロン-エヌ) 論法[2]

$$\forall\epsilon > 0, \exists N \in \mathbb{N}, n \geq N \Rightarrow |a_n - a| < \epsilon \quad (a_n \in (a-\epsilon, a+\epsilon))$$

[1] $\mathbb{N}$：自然数

[2] $\forall\epsilon > 0$：任意の正の数 ϵ（どれほど小さくてもよい）に対して；$\exists N \in \mathbb{N}$：ある自然数 N（N 番目の項）が存在する；$n \geq N$：N 番目以降の項 n に対して

確率収束

Def 8.2　確率収束 convergence in probability

確率変数[3]の列 $X_1, X_2, \ldots$ が確率変数 X に確率収束するとは

$$\forall \epsilon > 0, \quad \lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0 \Leftrightarrow \lim_{n\to\infty} P(|X_n - X| < \epsilon) = 1$$

- $X_n \xrightarrow{P} X, \quad X_n \xrightarrow{p} X$
- 無限な $X_1, X_2, \ldots$[4]は IID[5]と仮定されてない
- サンプルサイズ $n \to \infty$（極限過程 limiting process）
- n が変化すると、X_n の分布も変化する

[3] $X(\omega) \in \mathbb{R}$, Def 1.4
[4] やや空想的なアイデアではあるが、有限サンプルの状況においては有用な近似値を与える
[5] IID: Def7.2

大数の弱法則

Thm 8.3 大数の弱法則 Weak Law of Large Number (WLLN)

$X_1, X_2, \ldots \overset{i.i.d.}{\sim} (\mu, \sigma^2 < \infty)$ とする[6]，$\bar{X}_n = (1/n) \sum_{i=1}^{n} X_i$ は μ に確率収束する ($\bar{X}_n \xrightarrow{P} \mu$)

$$\Leftrightarrow \forall \epsilon > 0, \quad \lim_{n \to \infty} P(|\bar{X}_n - \mu| < \epsilon) = 1$$ (証)

- 有限な IID 標本 $X_1, \ldots, X_n$ の平均 $\xrightarrow{P}$ 母平均 (一致性)
- 確率変数[7]の列 $\bar{X}_1, \bar{X}_2, \ldots$ が実数 μ に確率収束する
- $\mu, \sigma^2 \in \mathbb{R}, \sigma^2 < \infty$ 分散は無限ではない
- 具体的な分布を仮定しない（任意の分布に対して成り立つ）
- IID 標本に限定される

[6] $X_1, X_2, \ldots$ が $E[X_i] = \mu$, $Var[X_i] = \sigma^2 < \infty$ の IID 確率変数とする
[7] $X_1, X_2, \ldots$ から計算によって新しい確率変数

Ex 8.4　S_n^2 の一致性

$X_1, X_2, \ldots \overset{i.i.d.}{\sim} (\mu, \sigma^2 < \infty)$, $Var[S_n^2] \to 0$

$$\Rightarrow S_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X_n})^2 \xrightarrow{P} \sigma^2$$ **(証)**

Supp 8.5

$X_1, X_2, \ldots \xrightarrow{P} X$，関数 $h(.)$ が連続関数 $\Rightarrow h(X_1), h(X_2), \ldots \xrightarrow{P} h(X)$

概収束

Def 8.6　概収束 almost sure convergence

確率変数の列 $X_1, X_2, \ldots$ が確率変数 X に概収束するとは，

$$\forall \epsilon > 0, \quad P(\lim_{n\to\infty} |X_n - X| \geq \epsilon) = 0 \Leftrightarrow P(\lim_{n\to\infty} |X_n - X| < \epsilon) = 1$$

- $X_n \xrightarrow{A.S.} X, \quad X_n \xrightarrow{a.s.} X$　ほとんど確実収束
- $\lim_{n\to\infty} X_n = X$
- $X_n \xrightarrow{A.S.} X \Rightarrow X_n \xrightarrow{P} X$
- $X_n \xrightarrow{P} X \Rightarrow \exists X_{n_j} \subset X_n,\ X_{n_j} \xrightarrow{A.S.} X$

大数の強法則

Thm 8.7　大数の強法則 Strong Law of Large Number (SLLN)

$X_1, X_2, \ldots \overset{i.i.d.}{\sim} (\mu, \sigma^2 < \infty)$ とする，$\bar{X}_n$ は μ に概収束する ($\bar{X}_n \xrightarrow{A.S.} \mu$)

$$\Leftrightarrow \forall \epsilon > 0, \quad P(\lim_{n\to\infty} |\bar{X}_n - \mu| < \epsilon) = 1$$

- 標本平均 $\xrightarrow{A.S.}$ 母平均
- $\lim_{n\to\infty} \bar{X}_n = \mu$
- $\forall \epsilon > 0, \exists N \in \mathbb{N}, n \geq N \Rightarrow |\bar{X}_n - \mu| < \epsilon$

確率収束 vs 概収束

Ex 8.8　概収束

確率空間 $(\Omega, \mathcal{B}([0,1]), P)$ を，$\Omega = [0,1]$ 上の一様分布，$P = \mathrm{Uniform}[0,1]$ とする

以下の確率変数を定義する

$$X_n(\omega) = \omega + \omega^n, \quad X(\omega) = \omega$$

- $\forall \omega \in [0,1)\ (\omega \neq 1),\ n \to \infty \Rightarrow \omega^n \to 0, X_n(\omega) \to \omega = X(\omega)$
- $\omega = 1, X_n(1) = 2 \neq X(1)$
- $P([0,1)) = 1 \Rightarrow X_n \xrightarrow{A.S.} X$

Ex 8.9　確率収束するが、概収束しない

確率空間 $(\Omega, \mathcal{B}([0,1]), P)$ を，$\Omega = [0,1]$ 上の一様分布 $P = \mathrm{Uniform}[0,1]$ とする

以下の確率変数を定義する

$$X(\omega) = \omega$$

$$X_1(\omega) = \omega + I_{[0,1]}(\omega)$$

$$X_2(\omega) = \omega + I_{[0,\frac{1}{2}]}(\omega), \quad X_3(\omega) = \omega + I_{[\frac{1}{2},1]}(\omega)$$

$$X_4(\omega) = \omega + I_{[0,\frac{1}{3}]}(\omega), \quad X_5(\omega) = \omega + I_{[\frac{1}{3},\frac{2}{3}]}(\omega), \quad X_6(\omega) = \omega + I_{[\frac{2}{3},1]}(\omega)$$

$$X_7(\omega) = \omega + I_{[0,\frac{1}{4}]}(\omega), \quad X_8(\omega) = \omega + I_{[\frac{1}{4},\frac{2}{4}]}(\omega), \quad X_9(\omega) = \omega + I_{[\frac{2}{4},\frac{3}{4}]}(\omega), \ldots$$

- $X_n(\omega) \xrightarrow{P} X(\omega) \Rightarrow n \to \infty, P(|X_n(\omega) - X(\omega)| \geq \epsilon) \to 0$
- ただし $\nexists \omega \in \Omega, X_n(\omega) \xrightarrow{A.S.} \omega = X(\omega)$

分布収束

Def 8.10 分布収束 convergence in distribution

確率変数の列 $X_1, X_2, \ldots$ が確率変数 X に分布収束するとは，

$$\forall x, \quad \lim_{n\to\infty} F_{X_n}(x) = F_X(x)$$

$F_X(x)$ は連続とする

- $X_n \xrightarrow{D} X, \quad X_n \xrightarrow{d} X$
- $\forall x, \quad \lim_{n\to\infty} P(X_n \leq x) = P(X \leq x)$
- $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{D} X \quad (X_n \xrightarrow{D} X \not\Rightarrow X_n \xrightarrow{P} X)$
- $X_n \xrightarrow{P} \mu \in \mathbb{R} \Leftrightarrow X_n \xrightarrow{D} \mu$
    - $\forall \epsilon > 0, P(|X_n - \mu| > \epsilon) \to 0 \Leftrightarrow P(X_n \leq x) \to \begin{cases} 0, & x < \mu \\ 1, & x > \mu \end{cases}$

中心極限定理

Thm 8.11　中心極限定理 Central Limit Theorem (CLT)

$X_1, X_2, \ldots \overset{i.i.d.}{\sim} (\mu, \sigma^2 < \infty)$(且つ 0 の近傍に $M_{X_i}(t)$ が存在とし)，$G_n(x)$ を $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$ の CDF とする，$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$ は $N(0,1)$ に分布収束する

$$\Leftrightarrow \forall -\infty < x < \infty, \quad \lim_{n \to \infty} G_n(x) = \Phi(x) \quad \text{(証)}$$

$$\Leftrightarrow \lim_{n \to \infty} P\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$

- $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$ は漸近的に標準正規分布に従う
    - $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \overset{\cdot}{\sim} N(0,1)$
    - $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{D} Z, Z \sim N(0,1)$
- Thm 5.5 MGF の収束

Ex 8.12　負の二項分布の近似

$X_1, \ldots, X_n \sim NegBin(r, p), \Rightarrow E[X] = \frac{r(1-p)}{p}, Var[X] = \frac{r(1-p)}{p^2}$

$$CLT \Rightarrow \frac{\sqrt{n}\bar{X}_n - E[X]}{\sqrt{Var[X]}} \dot{\sim} N(0,1)$$

$r = 10, p = 1/2, n = 30$

- 正確な確率

$$P(\bar{X}_n \leq 11) = P(\sum_{i=1}^{30} X_i \leq 330)$$
$$= \sum_{x=0}^{30} \binom{300+x-1}{x} (\frac{1}{2})^{300} (\frac{1}{2})^x = 0.8916$$

- 近似確率

$$P(\bar{X}_n \leq 11) = P(\frac{\sqrt{30}(\bar{X}_n - 10)}{\sqrt{20}} \leq \frac{\sqrt{30}(11-10)}{\sqrt{20}})$$
$$= P(Z \leq 1.2247) = 0.8888$$

Thm 8.13 **スルツキーの定理** Slustky's theorem

$X_n \xrightarrow{D} X,\, Y_n \xrightarrow{P} a \in \mathbb{R}$

- $X_n Y_n \xrightarrow{D} aX$
- $X_n + Y_n \xrightarrow{D} X + a$

Ex 8.14 **デルタ法** Delta method **(証)**

$Y_n, s.t.\sqrt{n}(Y_n - \theta) \xrightarrow{D} N(0, \sigma^2)$,

関数 g s.t. $g'(\theta) \neq 0$ 且つ存在するとき

$$\sqrt{n}(g(Y_n) - g(\theta)) \xrightarrow{D} N(0, \sigma^2[g'(\theta)]^2)$$

関数 g s.t. $g'(\theta) = 0, g''(\theta) \neq 0$ 且つ存在するとき

$$n(g(Y_n) - g(\theta)) \xrightarrow{D} \sigma^2 \frac{g''(\theta)}{2} \chi_1^2$$

レポート

課題 8.15

BEEF+ Report9 デルタ法の応用

9. 点推定量とその評価

2025 後期 3H606 数理科学入門 (統計系)–#12,#13

Jan 6, 13, 2026

数理環境論 Room 742
Email: ZHOU YI <zhouy@people.kobe-u.ac.jp>

キーワード

- 点推定量の導出方法
  - モーメント法 Method of moments
  - 最尤法 Method of maximum likelihood
  - ベイズ法 Bayesian method
- 推定量の評価法
  - 有限標本における評価指標
    - 平均２乗誤差 Mean squared error（精度）
    - 偏り Bias（正確性）
    - 最良不偏推定量 Best unbiased estimator
  - 無限標本における評価指標（漸近的評価）
    - 一致性 Consistency（漸近的正確性）
    - 有効性 Efficiency（漸近的精度）

点推定 Point estimation

Def 9.1　点推定量 Point estimator

標本の関数 $W(X_1, \ldots, X_n)$ として表されるすべての関数を指す

- あらゆる統計量 (Statistic) [1]は、点推定量である
- 推定値 (Estimate)：具体的に実現された値 $W(x_1, \ldots, x_n)$
- 2 つの課題
    1. 推定量の導出方法
    2. 推定方法の違いによって得られる推定量が異なる場合、どれを採用すべきか

[1] Def 7.4

モーメント法 Method of moments[3]

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} f(x|\theta_1, \ldots, \theta_k)$ とする

標本モーメント	母モーメント
$m_1 = \frac{1}{n}\sum_{i=1}^n X_i$	$\mu'_1 = E[X] = \mu'_1(\theta_1, \ldots, \theta_k)$
$m_2 = \frac{1}{n}\sum_{i=1}^n X_i^2$	$\mu'_2 = E[X^2] = \mu'_2(\theta_1, \ldots, \theta_k)$
$\vdots$	$\vdots$
$m_k = \frac{1}{n}\sum_{i=1}^n X_i^k$	$\mu'_k = E[X^k] = \mu'_3(\theta_1, \ldots, \theta_k)$

大数の法則より[2]，$m_k \xrightarrow{P} \mu'_k$

$$\begin{cases} m_1 = \mu'_1(\theta_1, \ldots, \theta_k) \\ m_2 = \mu'_2(\theta_1, \ldots, \theta_k) \\ \qquad \vdots \\ m_k = \mu'_k(\theta_1, \ldots, \theta_k) \end{cases} \Rightarrow (\theta_1, \ldots, \theta_k) = (\tilde{\theta}_1, \ldots, \tilde{\theta}_k)$$

$(\tilde{\theta}_1, \ldots, \tilde{\theta}_k)$: モーメント推定量（moment estimator）

[2]大数の法則と言う場合，通常は大数の弱法則 WLLN（Thm8.3）を指すことが多い
[3]最も古い歴史を持つ推定手法。1800 年代後半、Karl Pearson によって提唱された。

Ex 9.2　Bernoulli 分布

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} Bernoulli(p)$ とする（$\theta_1 = p$）

$$m_1 = \bar{X} \quad \mu_1' = p \Rightarrow \bar{X} = p \Rightarrow \tilde{p} = \bar{X}$$

Ex 9.3　正規分布

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ とする（$\theta_1 = \mu, \theta_2 = \sigma^2$）

$$\begin{array}{cc} m_1 = \bar{X} & \mu_1' = \mu \\ m_2 = \frac{1}{n}\sum X_i^2 & \mu_2' = \mu^2 + \sigma^2 \end{array}$$

$$\Rightarrow \begin{cases} \bar{X} = \mu \\ \frac{1}{n}\sum X_i^2 = \mu^2 + \sigma^2 \end{cases} \Rightarrow \tilde{\mu} = \bar{X}, \tilde{\sigma}^2 = \frac{1}{n}\sum (X_i - \bar{X})^2$$

Ex 9.4　二項分布（課題）

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} Binomial(n, p)$ とする（$\theta_1 = n, \theta_2 = p$）

最尤法 Method of maximum likelihood

Def 9.5　尤度関数 Likelihood function

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} f(x|\theta_1, \ldots, \theta_k)$ とする．尤度関数は以下のように定義される

$$L(\theta) = L(\theta_1, \ldots, \theta_k \mid x_1, \ldots, x_n) = \prod_{i=1}^{n} f(x_i \mid \theta)$$

対数尤度関数（log-likelihood function）：$\ell(\theta) = \log L(\theta) = \sum_{i=1}^{n} \log f(x_i \mid \theta)$

Def 9.6　最尤推定量 Maximum likelihood estimator, MLE

各固定された標本点 $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ において，尤度関数もしくは対数尤度関数を最大にする θ の解 $\hat{\theta}(\mathbf{x})$

Thm 9.7　MLE の不変性

$\hat{\theta}$ が θ の MLE であるとする．θ の任意の関数 $\tau(\theta)$ に対して、その関数の MLE は $\tau(\hat{\theta})$ となる

MLE の導出：全域的最大値（Global Maximum）を見つけ出す

1. 微分による候補の選定：尤度関数が（θ_i について）微分可能である場合、MLE の候補は、以下の連立方程式を解くことで得られる値 $(\theta_1, \ldots, \theta_k)$ となる

$$\frac{\partial}{\partial \theta_i} L(\theta|\mathbf{x}) = 0, \quad i = 1, \ldots, k$$

2. 必要条件としての微分：上式の解は、あくまで MLE の候補であることに注意が必要である。
   - 必要条件: 1 次微分が 0 であることは、最大値をとるための必要条件であって、十分条件ではない
   - 極値の性質: 1 次微分が 0 となる点は、極大値・最大値だけでなく、局所的/全域的最小値、あるいは変曲点である可能性がある
3. 境界値の確認：1 次微分が 0 となる点は、関数の定義域の内部にある極値のみを特定する
   - 境界での極値: もし最大値が定義域の境界（端点）で発生する場合、その点での 1 次微分は 0 になるとは限らない
   - 必須の作業: 境界値については別途確認を行い、全域的な最大値を見つけ出す必要がある

Supp 9.8　**極大・小値 / 最大・小値の定義**

- 極値 (Local Extrema) 関数のある限定された範囲（近傍）において、値が最も大きい、または小さい状態を指す
  - 極大・小値 (Local Maximum・Minimum)：周辺のどの点よりも高い・低い値をとる点
  - 極大値: $f'(x) = 0$ かつ $f''(x) < 0$ のとき
  - 極小値: $f'(x) = 0$ かつ $f''(x) > 0$ のとき
  - 判定不能/変曲点: $f''(x) = 0$
- 最適値 (Global Extrema) 関数の定義域全体を通して、値が最も大きい、または小さい状態を指す
  - 最大・小値 (Global Maximum・Minimum)：定義域全体で最も高い・低い値
  - 内部の極値と境界値（定義域 $a \leq x \leq b$ における $f(a)$ と $f(b)$ の値）をすべて比較して、全域的な最大値を特定する

Ex 9.9　Bernoulli 分布

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} Bernoulli(p)$ とする

1. 尤度関数の構成：$L(p|\mathbf{x}) = \prod_{i=1}^{n} p^{x_i}(1-p)^{1-x_i} = p^y(1-p)^{n-y}, \quad y = \sum x_i$ 対数尤度関数をとることで計算が大幅に簡略化される：
$$\ell(p) = \log L(p|\mathbf{x}) = y \log p + (n-y)\log(1-p)$$
2. MLE の候補の導出 $(0 < y < n)$：$\frac{d}{dp}\ell(p) = 0 \Rightarrow \hat{p} = \frac{y}{n}$
3. 最大値の確認
    - 唯一性：$\hat{p} = \frac{y}{n}$ はこの方程式の唯一の解
    - 極大値：
    $$\left.\frac{d^2}{dp^2}\ell(p)\right|_{p=\hat{p}} < 0 \Rightarrow \text{唯一の極大値}$$
    - 境界の確認：
        - $y = 0 : \ell(p) = n\log(1-p) \quad y = n : \ell(p) = n \log p$

    どちらの場合も関数は p に対して単調関数となる。$\hat{p} = y/n (= 0, 1)$ が最大値となることが確認される

Ex 9.10　二項分布（課題）

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} Binomial(n, p)$ とする．p の MLE

Ex 9.11　正規分布

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\theta, 1)$ とする

1. 尤度関数の構成

$$L(\theta|\mathbf{x}) = \prod_{i=1}^{n} \frac{1}{(2\pi)^{1/2}} e^{-(1/2)(x_i-\theta)^2} = \frac{1}{(2\pi)^{n/2}} e^{-(1/2)\sum_{i=1}^{n}(x_i-\theta)^2}$$

2. MLE の候補の導出：$\frac{d}{d\theta} L(\theta|\mathbf{x}) = 0 \Rightarrow \sum_{i=1}^{n}(x_i - \theta) = 0 \Rightarrow \hat{\theta} = \frac{1}{n}\sum x_i = \bar{x}$
3. 最大値の確認
    - 唯一性：$\hat{\theta} = \bar{x}$ はこの方程式の唯一の解
    - 極大値：

$$\left.\frac{d^2}{d\theta^2} L(\theta|\mathbf{x})\right|_{\theta=\bar{x}} < 0 \Rightarrow \text{唯一の極大値}$$

    - 境界の確認：$\theta \to \pm\infty$ において $L(\theta)$ は 0 に収束する

Ex 9.12　正規分布

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\theta, \sigma^2)$ とする

1. 尤度関数の構成

$$L(\theta, \sigma^2 \mid \mathbf{x}) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-(1/2)\sum_{i=1}^{n}(x_i-\theta)^2/\sigma^2}$$

   計算を簡略化するため、対数をとる：

$$\ell(\theta, \sigma^2) = \log L(\theta, \sigma^2 \mid \mathbf{x}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{1}{2}\sum_{i=1}^{n}(x_i-\theta)^2/\sigma^2$$

2. MLE の候補の導出：

$$\frac{\partial}{\partial\theta}\ell(\theta, \sigma^2) = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\theta) = 0 \Rightarrow \hat{\theta} = \bar{x}$$

$$\frac{\partial}{\partial\sigma^2}\ell(\theta, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\theta)^2 = 0 \Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum(x_i-\bar{x})^2$$

3. 最大値の確認
   - 唯一性：$\bar{x}$, $\frac{1}{n}\sum(x_i-\bar{x})^2$ は方程式の唯一の解

Supp 9.13　2 変数関数の極大値判定（一般論）

関数 $H(\theta_1, \theta_2)$ が点 $(\hat{\theta}_1, \hat{\theta}_2)$ で極大値を持つことを示すには、以下の 3 つの条件がすべて満たされている：

1. １次偏微分が 0 である（静止点）
$$\frac{\partial}{\partial\theta_1} H(\theta_1, \theta_2)\Big|_{\theta_1=\hat{\theta}_1, \theta_2=\hat{\theta}_2} = 0 \quad \text{および} \quad \frac{\partial}{\partial\theta_2} H(\theta_1, \theta_2)\Big|_{\theta_1=\hat{\theta}_1, \theta_2=\hat{\theta}_2} = 0$$
2. 2 次偏微分の少なくとも一方が負である
$$\frac{\partial^2}{\partial\theta_1^2} H(\theta_1, \theta_2)\Big|_{\theta_1=\hat{\theta}_1, \theta_2=\hat{\theta}_2} < 0 \quad \text{または} \quad \frac{\partial^2}{\partial\theta_2^2} H(\theta_1, \theta_2)\Big|_{\theta_1=\hat{\theta}_1, \theta_2=\hat{\theta}_2} < 0$$
3. 2 次偏微分のヤコビアンが正である
$$J = \begin{vmatrix} \frac{\partial^2}{\partial\theta_1^2} H & \frac{\partial^2}{\partial\theta_1 \partial\theta_2} H \\ \frac{\partial^2}{\partial\theta_1 \partial\theta_2} H & \frac{\partial^2}{\partial\theta_2^2} H \end{vmatrix}_{\theta_1=\hat{\theta}_1, \theta_2=\hat{\theta}_2} > 0$$

Ex 9.14 Eg 9.12 の続き

- 極大値の確認
  1. $\hat{\theta}, \hat{\sigma}^2$ を代入すると，１次偏微分が 0 である
  2. 2 次偏微分の少なくとも一方が負である

$$\frac{\partial^2}{\partial\theta^2}\ell(\theta,\sigma^2) = \frac{-n}{\sigma^2} < 0$$

  3. 2 次偏微分のヤコビアンが正である

$$\frac{\partial^2}{\partial(\sigma^2)^2}\ell(\theta,\sigma^2) = \frac{n}{2\sigma^4} - \frac{1}{\sigma^6}\sum(x_i-\theta)^2$$

$$\frac{\partial^2}{\partial\theta\partial\sigma^2}\log L = -\frac{1}{\sigma^4}\sum(x_i-\theta)$$

$$\begin{aligned} J(\bar{x},\hat{\sigma}^2) &= \frac{1}{\sigma^6}\left[\frac{-n^2}{2} + \frac{n}{\sigma^2}\sum(x_i-\theta)^2 - \frac{1}{\sigma^2}\left(\sum(x_i-\theta)\right)^2\right]\Bigg|_{\theta=\bar{x},\sigma^2=\hat{\sigma}^2} \\ &= \frac{1}{\hat{\sigma}^6}\frac{n^2}{2} > 0 \end{aligned}$$

- 境界の確認：$\theta \to \pm\infty, \sigma^2 \to \infty, \sigma^2 \to 0$，尤度は全体として 0 に収束します

ベイズ法

- 古典的統計（頻度論）: 母数 θ は「未知だが固定された値」とみなす
- ベイズ的統計: 母数 θ を変数とみなし、その変動を事前分布（prior distribution）$\pi(\theta|\xi)$ で表現する

$$\begin{cases} \mathbf{X}|\theta & \sim f(\mathbf{x}|\theta) \\ \theta & \sim \pi(\theta|\xi) \end{cases}$$

  - ξ は θ の事前分布の母数であり，超母数（hyperparameter）と呼ぶ
  - 観測データ $\mathbf{X} = \mathbf{x}$ を得た後、ベイズの定理を用いて事前分布を更新したものを事後分布（posterior distribution）$\pi(\theta|\mathbf{x}, \xi)$ と呼ぶ

$$\pi(\theta|\mathbf{x}, \xi) = \frac{f(\mathbf{x}|\theta)\pi(\theta|\xi)}{m(\mathbf{x}|\xi)}$$

  - $m(\mathbf{x}|\xi)$ は $\mathbf{X}$ の周辺分布

$$m(\mathbf{x}|\xi) = \int f(\mathbf{x}|\theta)\pi(\theta|\xi)d\theta$$

  - 事後分布の平均 $E[\theta|\mathbf{X}]$ をベイズ推定量（Bayes estimator）と呼ぶ

Ex 9.15　二項分布

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} Bernoulli(p)$ とする．$\Rightarrow Y = \sum X_i \sim Binomial(n, p)$
事前分布：$p \sim Beta(\alpha, \beta)$ とする

- 同時分布 $f(y, p)$

$$\begin{aligned} f(y,p) &= \left[\binom{n}{y} p^y (1-p)^{n-y}\right]\left[\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1}\right] \\ &= \binom{n}{y}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{y+\alpha-1}(1-p)^{n-y+\beta-1} \end{aligned}$$

- Y の周辺分布 (Beta-Binomial 分布)

$$f(y) = \int_0^1 f(y,p)dp = \binom{n}{y}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\frac{\Gamma(y+\alpha)\Gamma(n-y+\beta)}{\Gamma(n+\alpha+\beta)}$$ **(証)**

- p の事後分布

$$f(p|y) = \frac{f(y,p)}{f(y)} = \frac{\Gamma(n+\alpha+\beta)}{\Gamma(y+\alpha)\Gamma(n-y+\beta)} p^{y+\alpha-1}(1-p)^{n-y+\beta-1}$$

$$\Leftrightarrow p|y \sim Beta(y+\alpha, n-y+\beta)$$

$$\hat{p}_B = E[p] = \frac{y+\alpha}{\alpha+\beta+n}$$

$$\hat{p}_B = \left(\frac{n}{\alpha+\beta+n}\right)\left(\frac{y}{n}\right) + \left(\frac{\alpha+\beta}{\alpha+\beta+n}\right)\left(\frac{\alpha}{\alpha+\beta}\right)$$

Ex 9.16　正規分布　（証）

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\theta, \sigma^2)$ とし，事前分布：$\theta \sim N(\mu, \tau^2)$ とする．（σ^2, μ, τ^2 は既知）
事後分布の平均と分散

$$\hat{\theta}_B = E[\theta|\mathbf{x}] = \frac{n\tau^2}{n\tau^2 + \sigma^2}\bar{x} + \frac{\sigma^2}{n\tau^2 + \sigma^2}\mu$$

$$Var[\theta|\mathbf{x}] = \frac{\sigma^2\tau^2}{n\tau^2 + \sigma^2}$$

Def 9.17　共役事前分布 Conjugate prior distribution

事前分布 $\pi(\theta|\xi)$ とその事後分布 $\pi(\theta|\mathbf{x}, \xi)$ が同じ分布族に入るような事前分布を共役事前分布という

推定法の比較

- モーメント法
  - 簡単・粗い推定
  - 初期値や理論説明用に使われることが多い
  - 簡便だが、統計的最適性はあまり期待できない方法
- 最尤法
  - 大標本で特に強力
  - 頻度論的推定の標準的手法であり、理論的にも実用的にも最も広く使われる
- ベイズ法
  - 情報統合と不確実性評価に強い
  - 柔軟で解釈しやすいが、計算と事前分布の設定が鍵となる
  - 現代的計算環境と相性が良い

平均二乗誤差・バイアス

Def 9.18　平均二乗誤差 Mean Squared Error, MSE（精度の評価）

点推定量の平均二乗誤差：推定量と真の値 θ との差の二乗の期待値

$$E_\theta[(W-\theta)^2]$$

- $E_\theta[(W-\theta)^2] = Var_\theta[W] + (E_\theta[W]-\theta)^2 = Var_\theta[W] + (Bias_\theta[W])^2$
- 分散 $\approx$ 精度の欠如　バイアスの二乗 $\approx$ 正確性の欠如

Def 9.19　バイアス　偏り Bias（正確性の評価）

点推定量のバイアス：推定量の期待値と真の値 θ との差

$$Bias_\theta[W] = E_\theta[W] - \theta$$

- 推定量の精度を評価する
- 不偏推定量: バイアスが常に 0（$E_\theta[W]=\theta$）である推定量
- 不偏推定量の MSE：$E_\theta[(W-\theta)^2] = Var_\theta[W]$

Ex 9.20 **正規分布**

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ とする．
(μ, σ^2) の推定量（統計量）：$\bar{X} = \frac{1}{n}\sum X_i, S^2 = \frac{1}{n-1}\sum(X_i - \bar{X})^2$ は不偏推定量[4]

- $\bar{X}, S^2$ の MSE

$$E[(\bar{X}-\mu)^2] = Var[\bar{X}] = \frac{\sigma^2}{n}$$

$$E[(S^2-\sigma^2)^2] = Var[S^2] = \frac{2\sigma^4}{n-1} \quad \text{（正規性の仮定が必要）}$$

σ^2 の MLE[5]：$\hat{\sigma}^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$ バイアス

$$E[\hat{\sigma}^2] = \frac{n-1}{n}\sigma^2 \text{（偏った推定量）}$$

- $\hat{\sigma}^2$ の MSE

$$MSE(\hat{\sigma}^2) = E[(\hat{\sigma}^2-\sigma^2)^2] = \frac{2n-2+1}{n^2}\sigma^4 = \frac{2n-1}{n^2}\sigma^4 < \frac{2\sigma^4}{n-1}$$

偏った推定量 $\hat{\sigma}^2$ が、不偏分散 S^2 よりも MSE が小さくなる（真の値に近い確率が高い）

[4] Thm 7.7
[5] Eg 9.12

最良不偏推定量

Def 9.21 最良不偏推定量 best unbiased estimator、BUE・**一様最小分散不偏推定量** uniform minimum variance unbiased estimator, UMVUE

推定量 W^* が $\tau(\theta)$ の最良不偏推定量・一様最小分散不偏推定量であるとは、以下の 2 条件を満たす

1. W^* が不偏推定量 $E_\theta[W^*] = \tau(\theta)$
2. 他のいかなる不偏推定量 W と比較しても、その分散が最小である $Var_\theta[W^*] \leq Var_\theta[W]$

Ex 9.22　ポアソン分布

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} Poisson(\lambda)$ とする．
λ の推定量（統計量）：$\bar{X} = \frac{1}{n}\sum X_i, S^2 = \frac{1}{n-1}\sum(X_i - \bar{X})^2$，

- 不偏性[6]：$E_\lambda[\bar{X}] = \lambda, E_\lambda[S^2] = \lambda$
- 推定量の分散：$\mathrm{Var}_\lambda[\bar{X}] = \lambda/n, Var_\lambda[S^2] = \frac{\lambda}{n}\left(1 + \frac{2n}{n-1}\lambda\right)$

$$Var_\lambda[\bar{X}] \leq Var_\lambda[S^2]$$

- $\forall a, a\bar{X} + (1-a)S^2$ は不変推定量，

$$\forall a, Var_\lambda[\bar{X}] \leq Var_\lambda[a\bar{X} + (1-a)S^2]??$$

この条件が満たされる場合、$\bar{X}$ は UMVUE となる。
満たされない場合、$\bar{X}$ は UMVUE ではない。

任意の不偏推定量の分散がそれ以下にはなり得ない下限値（下界）（$B(\theta)$，クラメール・ラオの下限 Cramér-Rao Lower Bound）を特定する

[6] Thm 7.7

Thm 9.23　クラメール・ラオの不等式 Cramér-Rao Inequality

$X_1, \ldots, X_n \sim f(\mathbf{x}|\theta)$ とき，∀推定量 $W(\mathbf{X})$ s.t.

- 微分可能性: $\frac{d}{d\theta} E_\theta W(\mathbf{X}) = \int_{\mathcal{X}} \frac{\partial}{\partial\theta} [W(\mathbf{x}) f(\mathbf{x}|\theta)] d\mathbf{x}$ が成立
- 分散の有限性: $Var_\theta[W(\mathbf{X})] < \infty$

$$Var_\theta[W(\mathbf{X})] \geq \frac{\left(\frac{d}{d\theta} E_\theta[W(\mathbf{X})]\right)^2}{E_\theta\left[\left(\frac{\partial}{\partial\theta} \log f(\mathbf{X}|\theta)\right)^2\right]}$$ **(証)**

- フィッシャー情報量（Fisher Information）：

$$I(\theta) = E_\theta\left[\left(\frac{\partial}{\partial\theta} \log f(\mathbf{X}|\theta)\right)^2\right] = E[S(\theta, \mathbf{X})^2]$$

- スコア関数（score function）：$S(\theta, \mathbf{X}) = \frac{\partial}{\partial\theta} \log f(\mathbf{X}|\theta)$
- $W(\mathbf{X})$ が不変推定量であるとき，
  $E_\theta[W(\mathbf{X})] = \tau(\theta) \Rightarrow \frac{d}{d\theta} E_\theta[W(\mathbf{X})] = \tau'(\theta)$

Thm 9.24 クラメール・ラオの不等式 Cramér-Rao Inequality (IID 標本)

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} f(\mathbf{x}|\theta)$ とき，∀推定量 $W(\mathbf{X})$ s.t.

- 微分可能性: $\frac{d}{d\theta} E_\theta W(\mathbf{X}) = \int_{\mathcal{X}} \frac{\partial}{\partial\theta} [W(\mathbf{x}) f(\mathbf{x}|\theta)] d\mathbf{x}$ が成立
- 分散の有限性: $Var_\theta[W(\mathbf{X})] < \infty$

$$Var_\theta[W(\mathbf{X})] \geq \frac{\left(\frac{d}{d\theta} E_\theta[W(\mathbf{X})]\right)^2}{nE_\theta\left[\left(\frac{\partial}{\partial\theta} \log f(X|\theta)\right)^2\right]}$$

(証)

Supp 9.25 **(証)**

$f(x|\theta)$ が以下の条件を満たす場合

$$\frac{d}{d\theta}E_\theta\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right) = \int \frac{\partial}{\partial\theta}\left[\left(\frac{\partial}{\partial\theta}\log f(x|\theta)\right)f(x|\theta)\right]dx$$

$$\Rightarrow I(\theta) = E_\theta\left(\left(\frac{\partial}{\partial\theta}\log f(X|\theta)\right)^2\right) = -E_\theta\left(\frac{\partial^2}{\partial\theta^2}\log f(X|\theta)\right)$$

Ex 9.26 **Eg 9.22 の続き**

$\tau(\lambda) = \lambda, \Rightarrow \tau'(\lambda) = 1$

$$E_\lambda\left(\left(\frac{\partial}{\partial\lambda}\log\prod_{i=1}^{n}f(X_i|\lambda)\right)^2\right) = -nE_\lambda\left(\frac{\partial^2}{\partial\lambda^2}\log\frac{e^{-\lambda}\lambda^X}{X!}\right) = \frac{n}{\lambda}$$ **(証)**

$\forall$不偏推定量 W, $Var_\lambda W \geq \frac{1}{n/\lambda} = \frac{\lambda}{n}$

$\because$ $Var_\lambda \bar{X} = \frac{\lambda}{n}, \therefore$ $\bar{X}$ は UMVUE

Ex 9.27 **正規分布（課題）**

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ とする．σ^2 の不変推定量のクラメール・ラオの下限

一致性

Def 9.28　一致推定量 consistent estimator[7]（漸近的正確性）

推定量の列 $W_n = W_n(X_1, \ldots, W_n)$ が母数 θ の一致推定量の列であるとは，

$$\forall \epsilon > 0, \theta \in \Theta, \quad \lim_{n\to\infty} P_\theta(|W_n - \theta| < \epsilon) = 1 \Leftrightarrow \lim_{n\to\infty} P_\theta(|W_n - \theta| \geq \epsilon) = 0$$

Ex 9.29　正規分布

$X_1, \ldots \overset{i.i.d.}{\sim} N(\theta, 1)$ とする．$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i \sim N(\theta, \frac{1}{n})$

$$P_\theta(|\bar{X}_n - \theta| < \epsilon) \to 1 (n \to \infty)$$ **（証）**

[7] Def 8.2 確率収束

Thm 9.30

推定量の列 W_n が母数 θ の推定量である且つ以下の 2 つの条件を同時に満たすとき，$\forall\theta \in \Theta$, W_n は θ の一致推定量となる

- $\lim_{n\to\infty} Var_\theta W_n = 0$
- $\lim_{n\to\infty} Bias_\theta W_n = 0$

Thm 9.31

W_n が母数 θ の一致推定量であるとき，
$\forall a_1, a_2, \ldots, b_1, b_2, \ldots$（実数の列）s.t.,

- $\lim_{n\to\infty} a_n = 1$
- $\lim_{n\to\infty} b_n = 0$

$\Rightarrow U_n = a_n W_n + b_n$ も θ の一致推定量となる

Thm 9.32　MLE の一致性

$X_1, X_2, \ldots \overset{i.i.d.}{\sim} f(x|\theta), L(\theta|x) = \prod_{i=1}^{n} f(x_1|\theta)$ が尤度関数，$\hat{\theta}$ が θ の MLE，$\tau(\theta)$ が θ の連続関数であるとする．$f(x|\theta), L(\theta|x)$ が正則条件 (regularity conditions)[8]を満たす下で

$$\forall \epsilon > 0, \theta \in \Theta, \quad \lim_{n \to \infty} P_\theta(|\tau(\hat{\theta}) - \tau(\theta)| \geq \epsilon) = 0$$

$\tau(\hat{\theta})$ は $\tau(\theta)$ の一致推定量となる

- MLE：$\tau(\hat{\theta}) \xrightarrow{P} \tau(\theta) \Rightarrow \tau(\hat{\theta})$ は $\tau(\theta)$ の一致推定量となる

[8]正則性 regularity conditions とは、一般的に応用の範囲が広く、理論的な展開を容易にする条件を指す

Supp 9.33　正則条件の各項目

1. IID：$X_i \overset{i.i.d.}{\sim} f(x|\theta)$
2. パラメータの識別可能性 (Identifiability)：$\theta \neq \theta' \Rightarrow f(x|\theta) \neq f(x|\theta')$
3. $f(x|\theta)$ は、すべての θ に対して同じサポートを持つ且つ微分可能
4. 真の値 θ は Θ の内点です
5. f は θ に対して２次微分可能
6. $\int f(x|\theta)\,dx$ は θ に対して２次微分可能であり，微分は積分との交換が可能
7. f は θ に対して３次微分可能であり，定数 $c > 0$ と関数 $M(x)$ が存在する（即ち，対数尤度関数の３次微分が、真の値 θ_0 の近傍で発散しないこと）

$$|\frac{\partial^3}{\partial\theta^3}\log f(x;\theta)| \leq M(x), \quad \forall\theta \in (\theta_0 - c, \theta_0 + c), E_{\theta_0}[M(X)] < \infty$$

Def 9.34 漸近分散 asymptotic variance（漸近的精度）

推定量 T_n に対して、$k_n(T_n - \tau(\theta)) \xrightarrow{D} N(0, \sigma^2)$（分布収束）が成り立つとする。$\sigma^2$ を漸近分散・T_n の極限分布の分散と呼ぶ

Supp 9.35 極限分散 Limiting variance

推定量 T_n と定数列 $\{k_n\}$ に対して、以下の極限が存在し有限であるとき

$$\lim_{n \to \infty} k_n Var[T_n] = \tau^2 < \infty$$

τ^2 を、極限分散・分散の極限と呼びます。

例：$X_1, X_2, \ldots \overset{i.i.d.}{\sim} (\mu, \sigma^2)$, $T_n = \bar{X}_n$ とする，$\lim \sqrt{n} Var[\bar{X}_n] = \sigma^2/n$ は T_n の極限分散となる

$\frac{1}{\mu}$ を推定するために $T_n = \frac{1}{\bar{X}_n}$ を用いると問題が生じる

- 厳密な計算: $Var[T_n] = \infty$ となり、極限分散も無限大に発散する
- 近似的な計算（デルタ法）：$Var[T_n] \approx \frac{\sigma^2}{n\mu^4} < \infty$

Def 9.36　漸近有効 asymptotic efficient

推定量の列 W_n が以下の条件を満たすとき，

$$\sqrt{n}[W_n - \tau(\theta)] \xrightarrow{D} N(0, v(\theta)), \quad v(\theta) = \frac{[\tau'(\theta)]^2}{I(\theta)} = \frac{[\tau'(\theta)]^2}{E_\theta\left[\left(\frac{\partial}{\partial\theta}\log f(\mathbf{X}|\theta)\right)^2\right]}$$

W_n は $\tau(\theta)$ に対して漸近有効であるという

- $\tau(\theta) = \theta, \sqrt{n}[W_n - \theta] \xrightarrow{D} N(0, 1/I(\theta))$[9]
- W_n の漸近分散 $v(\theta)$ はクラメール・ラオの下限に達する

[9] Thm 9.23-3

Thm 9.37　MLE の漸近有効性

$X_1, X_2, \ldots \overset{i.i.d.}{\sim} f(x|\theta), L(\theta|x) = \prod_{i=1}^n f(x_1|\theta)$ が尤度関数，$\hat{\theta}$ が θ の MLE，$\tau(\theta)$ が θ の連続関数であるとする．$f(x|\theta), L(\theta|x)$ が正則条件を満たす下で

$$\sqrt{n}[\tau(\hat{\theta}) - \tau(\theta)] \xrightarrow{D} N(0, v(\theta)), \quad v(\theta) = \text{クラメール・ラオの下限}$$ **(証)**

- MLE：$\tau(\hat{\theta})$ は $\tau(\theta)$ の一致且つ漸近有効性推定量となる

Ex 9.38　漸近正規性

$\sqrt{n}\frac{W_n - \mu}{\sigma} \xrightarrow{D} Z, Z \sim N(0,1)$ スラツキーの定理[10]より

$$W_n - \mu = \left(\frac{\sigma}{\sqrt{n}}\right)\left(\sqrt{n}\frac{W_n - \mu}{\sigma}\right) \xrightarrow{D} \lim_{n\to\infty}\left(\frac{\sigma}{\sqrt{n}}\right)Z = 0$$

$\Rightarrow W_n - \mu \xrightarrow{D/P} 0$[11]

- 漸近正規性 $\Rightarrow$ 一致性

[10] Thm 8.13
[11] Def 8.10 点 4

レポート

課題 9.39

BEEF+ Report10-11

10. 仮説検定とその評価

2025 後期 3H606 数理科学入門 (統計系)–#14

Jan 20, 2026

数理環境論 Room 742
Email: ZHOU YI <zhouy@people.kobe-u.ac.jp>

キーワード

- 検定の導出方法
    - 尤度比検定 Likelihood Ratio Test
    - ワルド検定 Wald Test
    - スコア検定 Score Test
- 評価法
    - サイズと検出力 Size, Power
    - 一様最強力検定 Uniformly Most Powerful test, UMP test
    - 不偏検定 Unbiased test
    - P 値 P value

仮説検定 Hypothesis testing

Def 10.1　仮説 Hypothesis

仮設検定における仮説とは母集団のパラメーターに関する記述である

- 帰無仮説 null hypothesis：H_0
- 対立仮説 alternative hypothesis：H_1

Ex 10.2　仮説

母集団の PDF/PMF と $f(x|\theta)$ とし．$\theta \in \Theta$（母数空間 parameter space）とする．

$\Theta = \Theta_0 \cup \Theta_1, \Theta_0 \cap \Theta_1 = \emptyset$ と分割されるとき，$\theta \in \Theta_0? \in \Theta_1?$

$$H_0 : \theta \in \Theta_0 \quad vs \quad H_1 : \theta \in \Theta_1 (= \Theta_0^c)$$

$\Theta_0 = \{\theta_0\}$

$$H_0 : \theta = \theta_0 \quad vs \quad H_1 : \theta \neq \theta_0$$

仮説検定方式 Hypothesis testing procedure

- 仮説を棄却する (reject)：仮説を否定すること
- 仮説を受容する (accept)：仮説を受け入れること

Def 10.3　仮説検定方式 Hypothesis testing procedure

仮説検定方式とは，標本区間 $\mathcal{X}$ を $R = \{x \in \mathcal{X} | H_0\text{を棄却する}\}$,
$A = \{x \in \mathcal{X} | H_0\text{を受容する}\}$ に分割するルールのことである．
$(\mathcal{X} = R \cup A, R \cap A = \emptyset)$

- R：棄却域 rejection region
- A：受容域 acceptance region

標本 $X_1, \ldots, X_n$ に基づいた統計量 $T(X_1, \ldots, X_n)$ によって R と A 定める

- $T(X_1, \ldots, X_n)$：検定統計量 test statistics

検定統計量の導出方法

Def 10.4　尤度比検定 Likelihood Ratio Test, LRT

$H_0 : \theta \in \Theta_0$ 対 $H_1 : \theta \in \Theta_0^c$ を検定するための尤度比検定統計量 (likelihood ratio test statistic) は、以下で与えられる。

$$\lambda(\mathbf{x}) = \frac{\sup_{\Theta_0} L(\theta|\mathbf{x})}{\sup_{\Theta} L(\theta|\mathbf{x})}$$

尤度比検定とは、棄却域が $R = \{\mathbf{x} : \lambda(\mathbf{x}) \leq c\}(0 \leq c \leq 1)$[1] という形を持つあらゆる検定のことである。[2]

Ex 10.5　正規 LRT

$X_1, \ldots, X_n \sim N(\theta, 1)$ とする．$H_0 : \theta = \theta_0 \;\; vs \;\; H_1 : \theta \neq \theta_0$

- H_0 によって指定される θ の値は 1 つだけである．$\lambda(\mathbf{x})$ の分子は $L(\theta_0|\mathbf{x})$
- θ の（制約なしの）最尤推定量（MLE）は標本平均 $\bar{X}$ である．$\lambda(\mathbf{x})$ 分母は $L(\bar{x}|\mathbf{x})$

LRT 統計量：$\lambda(\mathbf{x}) = \exp\{-n(\bar{x} - \theta_0)^2/2\}$ **（証）**
棄却域：$R = \{\mathbf{x} : |\bar{x} - \theta_0| \geq \sqrt{-2\log(c)/n}\}$ **（証）**

[1] $R = \{\mathbf{x} : \lambda(\mathbf{x}) \leq c\}$ または $\lambda(\mathbf{X}) \leq c$
[2] sup: 上限

検定の評価

$H_0 : \theta \in \Theta_0$ 対 $H_1 : \theta \in \Theta_0^c$ においては、2 種類の誤りを犯す可能性がある

- 第 1 種の誤り・過誤（Type I Error）: $\theta \in \Theta_0$ である（帰無仮説が正しい）にもかかわらず、検定が誤って H_0 を棄却すると判断した場合
- 第 2 種の誤り・過誤（Type II Error）: $\theta \in \Theta_0^c$ である（対立仮説が正しい）にもかかわらず、検定が誤って H_0 を受容すると判断した場合

Table 1: 統計的仮説検定における意思決定と誤りの分類

真の状態 / 検定の結果	H_0 を受容 $(A = R^c)$	H_0 を棄却 (R)
$H_0 : \theta \in \Theta_0$ が正しい	正しい判断 $(1-\alpha)$	**第 1 種の過誤** (α)
$H_1 : \theta \in \Theta_0^c$ が正しい	**第 2 種の過誤** (β)	正しい判断 (**検出力**: $1-\beta$)

検定の棄却域を R とする

- 第 1 種の誤りの確率: $P_\theta(\mathbf{X} \in R)$ $(\theta \in \Theta_0)$
- 第 2 種の誤りの確率: $P_\theta(\mathbf{X} \in R^c)$ $(\theta \in \Theta_0^c)$

$$P_\theta(\mathbf{X} \in R) = \begin{cases} \text{第 1 種の誤りの確率} & (\theta \in \Theta_0 \text{ のとき}) \\ 1 - (\text{第 2 種の誤りの確率}) & (\theta \in \Theta_0^c \text{ のとき}) \end{cases}$$

Def 10.6 **検出力関数** power function

棄却域を R とすると、検出力関数 $\beta(\theta)$ は以下のように定義される

$$\beta(\theta) = P_\theta(\mathbf{X} \in R)$$

- $\theta \in \Theta_0$ のとき：$\beta(\theta) =$ 第 1 種の誤りの確率
- $\theta \in \Theta_0^c$ のとき：$1 - \beta(\theta) =$ 第 2 種の誤りの確率

Ex 10.7 **正規検出力関数**

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\theta, \sigma^2)$($\sigma$ が既知) とする．$H_0 : \theta \leq \theta_0$ 対 $H_1 : \theta > \theta_0$ を検定する場合を考える

LRT：$R = \{\mathbf{x} : (\bar{x} - \theta_0)/(\sigma/\sqrt{n}) > c\}$　（課題）

検出力関数：$\beta(\theta) = P\left(Z > c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right)$　**（証）**

Def 10.8　サイズと水準 size, level

- サイズ α の検定：$\sup_{\theta\in\Theta_0}\beta(\theta)=\alpha$ を満たすとき
- （有意）水準（レベル）α の検定：$\sup_{\theta\in\Theta_0}\beta(\theta)\le\alpha$ を満たすとき

Ex 10.9　LRT のサイズ

一般に、サイズ α の LRT を構成するには、$\sup_{\theta\in\Theta_0}P_\theta(\lambda(\mathbf{X})\le c)=\alpha$ となるように定数 c を選択する。

- Ex 10.5 では、棄却域を $|\bar{X}-\theta_0|\ge z_{\alpha/2}/\sqrt{n}$ とすることに対応する　**（証）**

Def 10.10　不偏検定

検出力関数 $\beta(\theta)$ を持つ検定が不偏（unbiased）であるとは、すべての $\theta' \in \Theta_0^c$（対立仮説の領域）と $\theta'' \in \Theta_0$（帰無仮説の領域）に対して、以下の不等式が成り立つことをいう。

$$\beta(\theta') \geq \beta(\theta'')$$

- $\beta(\theta) \leq \alpha$

Ex 10.11　正規検出力関数

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\theta, \sigma^2)$($\sigma$ が既知) とする．
$H_0 : \theta \leq \theta_0 \;\; vs \;\; H_1 : \theta > \theta_0$ を検定する場合を考える．検出力関数：

$$\beta(\theta) = P\left(Z > c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}}\right)$$

$\beta(\theta) : \theta$ に関して単調増加する関数である

$$\beta(\theta) > \beta(\theta_0) = \max_{t \leq \theta_0} \beta(t) \quad (\forall \theta > \theta_0)$$

この検定は不偏であると結論付けられる

Def 10.12 一様最強力検定 Uniformly Most Powerful Test, UMP 検定

$\mathcal{C}$ をある検定のクラスとする。クラス $\mathcal{C}$ に属する検定が一様最強力であるとは、その検出力関数 $\beta(\theta)$ が、同じクラスに属する他のあらゆる検定の検出力関数 $\beta'(\theta)$ に対して、対立仮説の領域 $(\theta \in \Theta_0^c)$ 全体で以下を満たすことをいう。

$$\beta(\theta) \geq \beta'(\theta)$$

- $\forall \theta \in \Theta_0^c, \beta(\theta) \geq \beta'(\theta)$
- UMP 検定は常に存在するわけではない

UMP 検定が存在するかどうかを判断するための最も基本的な道具

Thm 10.13　ネイマン・ピアソンの補題 Neyman-Pearson Lemma

$H_0 : \theta = \theta_0 \;\; vs \;\; H_1 : \theta = \theta_1$ の検定において、PMF/PDF をそれぞれ $f(\mathbf{x}|\theta_0)$、$f(\mathbf{x}|\theta_1)$ とする。
$\exists k \geq 0$[3]，以下の条件を満たす棄却域 R を持つ検定は、水準 α において最強力 (most powerful) となる。

1. $\mathbf{x} \in R$（棄却）となるのは $f(\mathbf{x}|\theta_1) > kf(\mathbf{x}|\theta_0)$ のとき
2. $\mathbf{x} \in R^c$（受容）となるのは $f(\mathbf{x}|\theta_1) < kf(\mathbf{x}|\theta_0)$ のとき
3. $\alpha = P_{\theta_0}(\mathbf{X} \in R)$（サイズ α の検定）

- 十分性 (Sufficiency)：上記の条件を満たす任意の検定は水準 α の UMP 検定
- 必要性 (Necessity)：$k > 0$ において上記の条件を満たす検定が存在する場合，任意の水準 α の UMP 検定は
  - 必ずサイズ α の検定（3 を満たす）
  - 確率が 0 でないすべての標本点 A（$P_{\theta_0}(\mathbf{X} \in A) \neq 0$）において、1 と 2 を満たす

[3] k はある定数

Ex 10.14　UMP 二項検定

$X \sim Binomial(2,\theta), H_0 : \theta = \frac{1}{2}\ \ vs\ \ H_1 : \theta = \frac{3}{4}$ を検定する．($x \in \{0,1,2\}$)

$$\frac{f(0|\theta=3/4)}{f(0|\theta=1/2)} = \frac{1}{4}, \quad \frac{f(1|\theta=3/4)}{f(1|\theta=1/2)} = \frac{3}{4}, \quad \frac{f(2|\theta=3/4)}{f(2|\theta=1/2)} = \frac{9}{4}$$

ネイマン・ピアソンの補題:

- $\frac{3}{4} < k < \frac{9}{4}$: 棄却域 $R = \{x = 2\}$
  レベル $\alpha = P(X = 2|\theta = \frac{1}{2}) = \frac{1}{4}$ の UMP 検定となる
- $\frac{1}{4} < k < \frac{3}{4}$: 棄却域 $R = \{x \in \{1,2\}\}$
  レベル $\alpha = P(X \in \{1,2\}|\theta = \frac{1}{2}) = \frac{3}{4}$ の UMP 検定となる
- $k > \frac{9}{4}, k < \frac{1}{4}$: レベル $\alpha = 0, 1$ の UMP 検定となる

Def 10.15　P 値

P 値 $p(\mathbf{X})$ とは，すべての標本点 $\mathbf{x}$ に対して $0 \leq p(\mathbf{x}) \leq 1$ を満たす検定統計量のことである。

- $\forall\theta \in \Theta_0$（帰無仮説の領域），$0 \leq \alpha \leq 1$ に対して、

$$P_\theta(p(\mathbf{X}) \leq \alpha) \leq \alpha$$

  $p(\mathbf{X}) \leq \alpha \Rightarrow H_0$ を棄却する，第 1 種の誤りの確率が α 以下に抑えられる
- $p(\mathbf{X})$ の値が小さいほど、対立仮説 H_1 が真であるという証拠が強いと判断される。

LRT の漸近分布

Thm 10.16　LRT の漸近分布

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} f(x|\theta), \hat{\theta}$ が θ の MLE，θ を１次元とする．仮説検定
$H_0 : \theta = \theta_0 \ \ vs \ \ H_1 : \theta \neq \theta_0$ における LRT 統計量を $\lambda(\mathbf{x})$ とすると，H_0 のもとで，

$$-2\log(\lambda(\mathbf{x})) \xrightarrow{D} \chi_1^2$$ **(証)**

- 水準 α の LRT の棄却域 $R = \{\mathbf{x} : -2\log(\lambda(\mathbf{x})) > \chi^2_{1,1-\alpha}\}$(上側 100 α%点)

Supp 10.17

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} f(x|\boldsymbol{\theta}), \hat{\boldsymbol{\theta}}$ が $\boldsymbol{\theta}$ の MLE，$\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^k$ とする．仮説検定
$H_0 : \boldsymbol{\theta} \in \Theta_0 \subseteq \mathbb{R}^r (r < k) \ \ vs \ \ H_1 : \boldsymbol{\theta} \notin \Theta_0$ における LRT 統計量を $\lambda(\mathbf{x})$ とすると，H_0 のもとで，

$$-2\log(\lambda(\mathbf{x})) \xrightarrow{D} \chi^2_{k-r}$$

棄却域 $R = \{\mathbf{x} : -2\log(\lambda(\mathbf{x})) > \chi^2_{k-r,1-\alpha}\}$

Ex 10.18 Poisson LRT

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} Poisson(\lambda), H_0 : \lambda = \lambda_0 \;\; vs \;\; H_1 : \lambda \neq \lambda_0$ を検定する．分子に H_0（λ_0）を、分母に H_1（MLE の $\hat{\lambda}$）を代入する

$$-2 \log \lambda(\mathbf{x}) = -2 \log \left(\frac{e^{-n\lambda_0} \lambda_0^{\sum x_i}}{e^{-n\hat{\lambda}} \hat{\lambda}^{\sum x_i}} \right) = 2n \left[(\lambda_0 - \hat{\lambda}) - \hat{\lambda} \log(\lambda_0 / \hat{\lambda}) \right]$$

棄却域: $-2 \log \lambda(\mathbf{X}) > \chi^2_{1,\alpha}$

他の漸近的検定

$H_0 : \theta = \theta_0 \ \ vs \ \ H_1 : \theta \neq \theta_0$ を検定する．

ワルド検定 (Wald Test)

推定量 W_n と、H_0 における θ_0 の距離に基づく検定

- W_n を θ の推定量で，$Var[W_n] = \sigma_n$ の推定量を S_n^2 とする．

$$Z_n = \frac{W_n - \theta}{S_n} \xrightarrow{D} N(0,1)$$

- $H_0 : \theta = \theta_0$ の下で

$$Z_n = \frac{W_n - \theta_0}{S_n} \xrightarrow{D} N(0,1)$$

棄却域：$R = \{\mathbf{x} : |W_n - \theta_0|/S_n \geq z_{1-\alpha/2}\}$

- $W_n = \hat{\theta}_n$ を MLE とする．

$$\sqrt{nI_1(\theta_n)}(\hat{\theta}_n - \theta) \xrightarrow{D} N(0,1)$$

$H_0 : \theta = \theta_0$ の下で，棄却域：$R = \{\mathbf{x} : \sqrt{nI_1(\hat{\theta}_n)}|\hat{\theta}_n - \theta_0| \geq z_{1-\alpha/2}\}$

スコア検定（Score Test / Lagrange Multiplier Test）

H_0 における尤度関数の傾き（スコア）の大きさに基づく検定

- スコア関数：$S(\theta) = \frac{\partial}{\partial\theta}\log f(\mathbf{X}|\theta) = \frac{\partial}{\partial\theta}\log L(\theta|\mathbf{X})$
- $E_\theta[S(\theta)] = 0$, $Var_\theta[S(\theta)] = I_n(\theta)$

$$Z_S = \frac{S(\theta)}{\sqrt{I_n(\theta)}} \xrightarrow{D} N(0,1)$$

- $H_0 : \theta = \theta_0$ の下で

$$Z_S = \frac{S(\theta_0)}{\sqrt{I_n(\theta_0)}} \xrightarrow{D} N(0,1)$$

IID 標本の場合

$$Z_S = \frac{S(\theta_0)}{\sqrt{nI_1(\theta_0)}} \xrightarrow{D} N(0,1)$$

棄却域：$R = \{\mathbf{x} : |S(\theta_0)|/\sqrt{nI_1(\theta_0)} \geq z_{1-\alpha/2}\}$

Ex 10.19　二項検定

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} Bernoulli(p), H_0 : p = p_0 \;\; vs \;\; H_1 : p \neq p_0$ を検定する．

- ワルド検定
  p の MLE：$\hat{p}_n = \sum_i X_i / n$

$$\frac{\hat{p}_n - p}{\sigma_n} \xrightarrow{D} N(0,1) \;\; (CLT), \quad \sigma_n = \sqrt{p(1-p)/n}$$

  - σ_n の推定量：$S_n = \sqrt{\hat{p}_n(1-\hat{p}_n)/n}$

$$\frac{\hat{p}_n - p}{\sqrt{\hat{p}_n(1-\hat{p}_n)/n}} \xrightarrow{D} N(0,1)$$

  H_0 の下で，

$$Z_n = \frac{\hat{p}_n - p_0}{\sqrt{\hat{p}_n(1-\hat{p}_n)/n}} \xrightarrow{D} N(0,1)$$

  棄却域：$R = \{\mathbf{x} : |Z_n| \geq z_{1-\alpha/2}\}$

Ex 10.20　二項検定

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} Bernoulli(p), H_0 : p = p_0 \;\; vs \;\; H_1 : p \neq p_0$ を検定する．

- スコア検定

$$S(p) = \frac{\hat{p}_n - p}{p(1-p)/n}, \quad I_n(p) = \frac{n}{p(1-p)}$$

  - H_0 の下で，

$$Z_S = \frac{S(p)}{\sqrt{I_n(p)}} = \frac{\hat{p}_n - p}{\sqrt{p_0(1-p_0)/n}} \xrightarrow{D} N(0,1)$$

棄却域：$R = \{\mathbf{x} : |Z_S| \geq z_{1-\alpha/2}\}$

レポート

課題 10.21

BEEF+ Report12

11. 区間推定とその評価

2025 後期 3H606 数理科学入門 (統計系)–#15

Jan 27, 2026

数理環境論 Room 742
Email: ZHOU YI <zhouy@people.kobe-u.ac.jp>

キーワード

- 区間推定の構成方法
  - 反転法
  - 枢軸変数法
- 評価法
  - カバレージ確率 Coverage Probability

区間推定

Def 11.1 区間推定 Interval estimation

実数値パラメータ θ の区間推定とは、標本に基づく 2 つの統計量 $L(x_1, \ldots, x_n)$ と $U(x_1, \ldots, x_n)$ のペアであり、すべての $\mathbf{x} \in \mathcal{X}$ に対して $L(\mathbf{x}) \leq U(\mathbf{x})$ を満たすものである。

$\mathbf{X} = \mathbf{x}$ 観測値が得られたとき，$L(\mathbf{x}) \leq \theta \leq U(\mathbf{x})$ という推論が行われる．このランダムな区間 $[L(\mathbf{X}), U(\mathbf{X})]$ を区間推定量 (interval estimator) と呼ぶ

Ex 11.2 区間推定量

正規分布 $n(\mu, 1)$ からの標本 X_1, X_2, X_3, X_4 において，μ の区間推定量の一つは $[\bar{X} - 1, \bar{X} + 1]$ である

これは、μ がこの区間内にあると主張する

Ex 11.3　Ex 11.2 の続き

μ を $\bar{X}$ で推定するとき，$P(\mu = \bar{X}) = 0$ である。

区間推定法を用いれば、正解する確率は正の値になる。μ が区間 $[\bar{X}-1, \bar{X}+1]$ にカバーされる確率は、次のように計算できる。

$$
\begin{aligned}
P(\mu \in [\bar{X}-1, \bar{X}+1]) &= P(\bar{X}-1 \le \mu \le \bar{X}+1) \\
&= P(-1 \le \bar{X}-\mu \le 1) \\
&= P\left(-2 \le \frac{\bar{X}-\mu}{\sqrt{1/4}} \le 2\right) \\
&= P(-2 \le Z \le 2) \quad (Z \sim (0,1)) \\
&= .9544
\end{aligned}
$$

区間推定法を使えば、未知の μ をカバーできる確率は 95% 以上あることになる。

区間推定の評価

Def 11.4

θ の区間推定量 $[L(\mathbf{X}), U(\mathbf{X})]$ において、その被覆確率（coverage probability，カバレージ確率）とは θ の値の関数であり、次のように定義される

$$P_\theta(\theta \in [L(\mathbf{X}), U(\mathbf{X})])$$

$[L(\mathbf{X}), U(\mathbf{X})]$ の信頼係数（confidence coefficient）とは、被覆確率の下限（infimum）で定義される

$$\inf_\theta P_\theta(\theta \in [L(\mathbf{X}), U(\mathbf{X})])$$

- $P_\theta(\theta \in [L(\mathbf{X}), U(\mathbf{X})]) \approx \frac{\#(\theta \in [L(\mathbf{X}), U(\mathbf{X})])}{\text{模擬試験の繰り返す回数}}$
- $P_\theta(\theta \in [L(\mathbf{X}), U(\mathbf{X})]) \geq 1-\alpha$ とき，$[L(\mathbf{X}), U(\mathbf{X})]$ を信頼係数 $1-\alpha$ の信頼区間という

Ex 11.5　被覆確率

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ とする（σ^2 が既知）.
$L(\bar{X}) = \bar{X} - \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}$, $U(\bar{X}) = \bar{X} + \frac{\sigma}{\sqrt{n}} z_{1-\alpha/2}$，($z_{1-\alpha/2}$ 上側分位点) とする

$$\begin{aligned}
P_\mu(\mu \in [L(\mathbf{X}), U(\mathbf{X})]) &= P_\mu\left(\bar{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \\
&= P_\mu\left(-z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu - \bar{X} \leq z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) \\
&= P_\mu\left(-z_{1-\alpha/2} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z_{1-\alpha/2}\right) \\
&= P(z_{\alpha/2} \leq Z \leq z_{1-\alpha/2}) \quad Z \sim N(0,1) \\
&= (1 - \alpha/2) - (\alpha/2) \\
&= 1 - \alpha
\end{aligned}$$

区間推定の構成法

Ex 11.6　検定方式の反転

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2), H_0 : \mu = \mu_0 \;\; vs \;\; H_1 : \mu \neq \mu_0$ を考える．有意水準 α において、一様最強不偏（UMP）検定の棄却域：[1]

$$\{\mathbf{x} : |\bar{x} - \mu_0| > z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\}$$

受容域：

$$\{\mathbf{x} : |\bar{x} - \mu_0| \leq z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\} = \bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu_0 \leq \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$\Rightarrow P(H_0 \text{ を受容する} | \mu = \mu_0) = 1 - \alpha$$

$$P\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu_0 \leq \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \middle| \mu = \mu_0\right) = 1 - \alpha, \quad \forall \mu_0 \in \Theta_0$$

$$P_\mu\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

水準 α の検定の受容域を反転させることで得られた区間 $[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}]$ は、$1 - \alpha$ 信頼区間である

[1] Ex 10.9

なぜ「反転」と呼ぶ

- 検定の視点: μ_0 を固定し、データ $\bar{X}$ が「受容域」に入る確率を考える
- 反転のステップ: 不等式の中身を整理して、中央に μ_0 を持ってくる

信頼区間：μ_0 を真値だと仮定したときに、今持っているデータが「棄却」とは言えないような μ_0 のリスト（範囲、区間）

「統計的検定」と「信頼区間」が実は表裏一体である

- 検定の「受容域」から「信頼区間」へ
  受容域 $A(\mu_0)$: ある特定の μ_0 を真だと仮定したとき、その μ_0 が棄却されないようなデータの集合

$$A(\mu_0) = \left\{ (x_1, \ldots, x_n) : \mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \bar{x} \le \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right\}$$

  信頼区間 $C(x_1, \ldots, x_n)$: ある特定のデータを得たとき、そのデータによって棄却されないようなパラメータ μ の集合

$$C(x_1, \ldots, x_n) = \left\{ \mu : \bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \right\}$$

- 反転の論理：トートロジー（同義反復）

$$(x_1, \ldots, x_n) \in A(\mu_0) \iff \mu_0 \in C(x_1, \ldots, x_n)$$

  「データがに入っていること」と「パラメータが信頼区間に入っていること」は数学的に全く同じ不等式を別の視点から見ている

Thm 11.7　検定の反転法

- 検定から信頼集合へ
  $H_0 : \theta = \theta_0 (\forall \theta_0)$ のサイズ α の検定の受容域を $A(\theta_0)$ とする．データ $\mathbf{x}$ に対して次のように定義される集合 $C(\mathbf{x})$ は、信頼係数 $1-\alpha$ の信頼集合となる

$$C(\mathbf{x}) = \{\theta_0 : \mathbf{x} \in A(\theta_0)\}$$

- 信頼集合から検定へ
  $C(\mathbf{X})$ が信頼係数 $1-\alpha$ の信頼集合であるとき，

$$A(\theta_0) = \{\mathbf{x} : \theta_0 \in C(\mathbf{x})\} \quad \forall \theta_0$$

$A(\theta_0)$ は $H_0 : \theta = \theta_0$ に対する有意水準 α の検定の受容域となる

Ex 11.8 枢軸変数

$X_1, \ldots, X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2), H_0 : \mu = \mu_0 \;\; vs \;\; H_1 : \mu \neq \mu_0$ を考える．

- 母分散 σ^2 が既知の場合: $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$ (z 枢軸)

$$\forall a, \quad P\left(-a \leq \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \leq a\right) = P(-a \leq Z \leq a)$$

  μ について解くと、以下の信頼区間が得られます：

$$\left\{\mu : \bar{x} - a\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + a\frac{\sigma}{\sqrt{n}}\right\}$$

- 母分散 σ^2 が未知の場合：$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$ (t 枢軸)

$$P\left(-a \leq \frac{\bar{X}-\mu}{S/\sqrt{n}} \leq a\right) = P(-a \leq T_{n-1} \leq a)$$

  μ について解くと、以下の信頼区間が得られます：

$$\left\{\mu : \bar{x} - t_{n-1,\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{n-1,\alpha/2}\frac{s}{\sqrt{n}}\right\}$$

信頼係数 $1-\alpha$ とする． $a = z_{1-\alpha/2}, t_{n-1,1-\alpha/2}$

Def 11.9　枢軸変数（Pivot）

変数 $Q(\mathbf{X}, \theta)$ の分布が θ に依存しないとき，$Q(\mathbf{X}, \theta)$ を枢軸変数・枢軸量という

Thm 11.10　枢軸変数法

- 分布の範囲を固定: Q の分布において、確率が $1 - \alpha$ となるような定数 a, b を選ぶ
$$P(a \leq Q(\mathbf{X}, \theta) \leq b) = 1 - \alpha$$
- 不等式の変形: 不等式の中身を、θ について解く形に変形・反転する
$$P(L(\mathbf{X}) \leq \theta \leq U(\mathbf{X})) = 1 - \alpha$$
- 結論: $[L(\mathbf{X}), U(\mathbf{X})]$ が、信頼係数 $1 - \alpha$ の信頼区間

Ex 11.11　母分散 σ^2 の枢軸変数

母平均 μ が未知の場合でも、母分散 σ^2 を推定するために以下の枢軸変数を利用できる

不偏分散 S^2 と母分散 σ^2 の間には次の関係が成り立つ

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

未知の σ^2 を含んでいますが、その分布自体は σ^2 には依存しない．したがって、これは枢軸変数となる

確率が $1-\alpha$ となるような範囲 $[a, b]$ を選ぶ

$$P\left(a \leq \frac{(n-1)S^2}{\sigma^2} \leq b\right) = 1-\alpha$$

σ^2 について解く

$$\left\{\sigma^2 : \frac{(n-1)s^2}{b} \leq \sigma^2 \leq \frac{(n-1)s^2}{a}\right\} \Leftrightarrow \left\{\sigma : \sqrt{\frac{(n-1)s^2}{b}} \leq \sigma \leq \sqrt{\frac{(n-1)s^2}{a}}\right\}$$

最終試験

注意事項

- 時間：2026-02-03 15:10–16:40　（1.5 時間）
- A4 用紙 1 枚のチートシート（手書きノート、両面）は使用可
- 鉛筆と消しゴムを用意してください
- 電子機器使用は禁止
- 評価：最終成績の 50%を占める
- 形式：全 5 問（確率論 2 問 ×2 セット、統計推測 3 問 ×2 セット）
  5 問中 4 問を選択して解答してください
- 各問では、確率質量関数（PMF）および確率密度関数（PDF）の数式を与える
- 未提出のレポートは、1 月 31 日 (11:59 PM) までに提出してください