

目次

第 1 章	分布の扱い	5
1.1	分布関数と密度関数	5
1.1.1	1次元の場合の分布関数の定義と特徴付け	5
1.1.2	1次元の密度関数	6
1.1.3	一意性定理	6
1.1.4	多次元の場合	6
1.2	確率表現関数	6
1.2.1	確率表現関数の定義	6
1.3	共分散	7
1.3.1	1次元の共分散	7
1.3.2	独立性と共分散	7
1.3.3	多次元の共分散	8
1.3.4	半正定値行列としての特徴付け	9
1.4	積率母関数	9
1.4.1	積率母関数の定義と収束性	9
1.4.2	積率母関数の Maclaurin 展開	9
1.4.3	一意性定理	10
1.5	キュムラント母関数	10
1.5.1	キュムラント母関数	10
1.5.2	積率問題	11
1.6	離散分布の確率母関数	12
1.6.1	離散分布の特性関数と確率母関数	12
1.6.2	確率母関数の微分係数	12
1.7	特性関数	13
1.7.1	特性関数の定義	13
1.7.2	特性関数と分布の収束の対応	13
1.7.3	特性関数の滑らかさと積率の存在	14
1.7.4	特性関数の Maclaurin 展開	15
1.7.5	特性関数の減衰速度と密度関数の可積分性との対応	16
1.7.6	反転公式と一意性定理	16
1.7.7	特性関数の特徴付け	18
1.7.8	特性関数の構成	19
1.8	第 2 キュムラント母関数	19
1.8.1	第 2 キュムラント母関数	19
第 2 章	確率変数の従属性	21
2.1	確率変数の独立性と期待値	21
2.1.1	積の期待値による独立性の特徴付け	21
2.1.2	確率変数の独立性と特性関数	22

2.1.3	密度関数と独立性	22
2.2	多次元分布の従属性の扱い	22
2.2.1	同時・周辺密度関数	22
2.2.2	条件付き密度関数	23
2.3	多次元分布の特性値	23
2.3.1	積率の定義	23
2.3.2	相関係数	24
2.3.3	キュムラントの定義	24
2.4	コピュラ	24
2.4.1	定義と例	25
2.4.2	Sklar の定理	25
2.4.3	Archimedes コピュラ	25
第 3 章	分布の変換	26
3.1	確率密度関数の変換	26
3.1.1	一般の変換則	26
3.1.2	一様分布への変換	27
3.1.3	一様分布の正規分布への変換	27
3.1.4	一様分布の指数分布への変換	28
3.2	商による変換	28
3.2.1	Gamma 分布の Beta 分布への変換	28
3.2.2	正規分布の Cauchy 分布への変換	29
3.3	自乗による変換	29
3.3.1	一般論	29
3.3.2	正規分布の二乗はカイ二乗分布	30
3.4	和による変換	30
3.4.1	一般論	30
3.4.2	二項分布は Bernoulli 分布の畳み込み	30
3.4.3	Poisson 分布の再生性	30
3.4.4	指数分布の独立和	31
3.4.5	Gamma 分布の再生性	31
3.4.6	総和核の再生性	32
3.5	順序統計量	32
3.5.1	順序統計量の密度関数	32
3.5.2	結合密度関数	32
3.5.3	一様分布の順序統計量は Beta 分布に従う	33
3.5.4	指数分布の場合	34
3.5.5	順序統計量の関数である統計量	34
3.6	線型変換と正規標本統計量	35
3.6.1	標本平均と標本分散と不偏分散	35
3.6.2	正規標本の場合	36
3.6.3	線形変換と独立性	37
3.6.4	線形変換と収束	38
3.7	混合分布	38
第 4 章	確率分布の例	39
4.1	一次元純不連続分布の例	39
4.1.1	離散一様分布	39

4.1.2	二項分布	40
4.1.3	Poisson 分布	41
4.1.4	Poisson 分布の出現	42
4.1.5	負の二項分布・幾何分布	43
4.1.6	Katz 族	45
4.1.7	超幾何分布	45
4.1.8	超幾何分布の二項近似	46
4.1.9	超幾何関数	46
4.1.10	負の超幾何分布	47
4.1.11	対数分布	47
4.1.12	Ord 族	47
4.2	多次元純不連続分布の例	47
4.2.1	多項分布	47
4.2.2	2 変量 Poisson 分布	48
4.2.3	負の多項分布	48
4.3	正規分布	49
4.3.1	定義と特性値	49
4.3.2	多変量正規分布の定義	51
4.3.3	条件付き分布	51
4.3.4	線型汎函数による特徴付け	52
4.3.5	3 次以上のキュムラントが消える	52
4.3.6	正規確率変数の独立和への分解	52
4.3.7	統計量の分布	53
4.3.8	二次形式の分布	54
4.3.9	微分方程式	54
4.3.10	混合正規分布	55
4.4	一次元絶対連続分布の例	55
4.4.1	一様分布	55
4.4.2	Gamma 関数の性質	56
4.4.3	Gamma 分布	57
4.4.4	指数分布	58
4.4.5	カイ 2 乗分布	59
4.4.6	Fisher-Cochran の定理	61
4.4.7	Student 分布	62
4.4.8	合流型超幾何関数	64
4.4.9	Fisher-Snedecor 分布	65
4.4.10	Beta 関数	66
4.4.11	Beta 分布	68
4.4.12	第 2 種 Beta 分布	69
4.4.13	Cauchy 分布	69
4.4.14	Weibull 分布	70
4.4.15	対数正規分布	70
4.4.16	Zeta 関数	70
4.4.17	logistic 分布	71
4.4.18	Pareto 分布	72
4.4.19	逆正規分布	73
4.4.20	逆 Gamma 分布	73

4.4.21	Laplace 分布	74
4.4.22	Marcenko-Pastur 分布	74
4.4.23	Pearson 系	74
4.5	多次元絶対連続分布の例	75
4.5.1	Wishart 分布	75
4.5.2	Dirichlet 分布	75
4.6	Poisson 過程に付随する分布	76
4.6.1	Poisson 確率変数の和と差	76
4.6.2	Poisson 過程に付随する分布	76
4.7	分布の変換による関係	76
第 5 章	古典数理統計学	77
5.1	指数型分布族	77
5.1.1	定義と例	77
5.1.2	指数型分布族の性質	78
5.1.3	密度関数が生成する標準指数型分布族	78
5.1.4	指数型分布族の完備十分統計量	79
5.1.5	共役事前分布	79
5.2	指数分散モデル	80
5.2.1	構成	80
5.2.2	Tweedie 分布族	80
5.3	漸近関係	81
第 6 章	参考文献	82
参考文献		83
参考文献		84

第 1 章

分布の扱い

統計推測は $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ 上の確率分布に関する逆問題である。確率分布に対して使える数学的な道具をまとめる。

- (1) 積率母関数は $M(t) := E[e^{tX}]$. 定義域が左側半平面をなすが, 収束さえすれば正則性が保証される. 0 の微分係数が積率である.
- (2) 特性関数は $\varphi(u) := E[e^{iuX}]$. 定義域は全平面であるが, $\beta_n < \infty$ が成り立つ度に C^n -級であることが保証されるのみである. 0 の p 次微分係数は積率の i^p 倍である.
- (3) 確率母関数は $g(z) := E[z^X]$. 係数に質量関数の値 p_z を並べて得る母関数である.
- (4) キュムラント母関数は $C(t) := \log M(t) = \log E[e^{tX}]$.

記法 1.0.1.

- (1) 距離空間 S について, $P(S)$ でその上の確率測度の全体とする.
- (2) Euclid 空間 \mathbb{R}^n について, $\mathcal{P}(\mathbb{R}^n)$ でその上の Lebesgue 測度に関して絶対連続な確率測度の全体とする.

1.1 分布関数と密度関数

分布関数には 3 つの同値な定義がある。これらが同値である理由は, Riemann-Stieltjes 積分による。任意の単調増加関数 F に対して, 有界関数 f の積分が定まるが, 次の 3 条件を満たすことと, F が絶対連続な確率分布を定めることが同値になる。

1.1.1 1 次元の場合の分布関数の定義と特徴付け

定義 1.1.1 (distribution function).

- (1) 次の 3 条件を満たす有界関数 $F: \mathbb{R} \rightarrow [0, 1]$ を**分布関数**という。
 - (i) 有界性: $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow \infty} F(x) = 1$.
 - (ii) 右連続性: $\forall x \in \mathbb{R} \quad \lim_{y \rightarrow x-0} F(y) = F(x)$.
 - (iii) 広義単調増加性: $x \leq y \Rightarrow F(x) \leq F(y)$.
- (2) 実確率変数 X に対して, $F^X(x) := P[X \leq x]$ により定まる関数 $F^X: \mathbb{R} \rightarrow [0, 1]$ は分布関数になる。これを X の**(累積) 分布関数**という。
- (3) $(\mathbb{R}, \mathcal{B}_1)$ 上の確率測度 ν に対して, $F_\nu(x) := \nu((-\infty, x]) = \int_{-\infty}^x f(y)dy$ により定まる関数 $F^X: \mathbb{R} \rightarrow [0, 1]$ は分布関数になる。これを ν の**(累積) 分布関数**という。

補題 1.1.2 (任意の分布関数はある確率分布の累積分布関数である). $F: \mathbb{R} \rightarrow \mathbb{R}$ を有界関数とする。このとき, 次は同値:

- (1) F は分布関数である。
- (2) ある正則確率測度 $\nu \in P(\mathbb{R})$ が存在して, この累積分布関数である: $F = F_\nu$.

このときの ν を, F の **Lebesgue-Stieltjes 測度**という。

[証明].

- (1) \Rightarrow (2) 任意の Borel 集合 $(-\infty, x] \in \mathcal{G}_1$ について, $\nu((-\infty, x]) := F(x)$ とすると, $(-\infty, x]$ は任意の \mathbb{R} の開・閉区間を生成するので, \mathcal{G}_1 を生成することから, ν は \mathcal{G}_1 上の測度に延長する. さらに, $\nu(\mathbb{R}) = F(\infty) = 1$ だから, ν は確率測度である.
- (2) \Rightarrow (1) $F(x) = \nu((-\infty, x])$ であるから,
- (i) 明らか.
 - (ii) 確率測度の連続性より.
 - (iii) 確率測度の正性より.

■

1.1.2 1次元の密度関数

定義 1.1.3. 確率分布 $\nu \in P(\mathbb{R}^d)$ に対して,

- (1) ν が原子を持たないとき, **連続型確率分布**という. 実はこれは, 分布関数 F_ν が連続であることに同値.
- (2) ν が原子を持ち, さらに $\nu = \sum_{i \in \mathbb{N}} a_i \delta_{x_i}$ と表せるとき, **離散型確率分布**という.
- (3) ν が連続型で, さらに Lebesgue 測度に対して絶対連続であるとき, **絶対連続型確率分布**といい, その全体を $\mathcal{P}(\mathbb{R}^d)$ で表す.

定義 1.1.4 (mass function, density function).

- (1) 離散分布 $\nu \in P(\mathbb{R}^d)$ に対して, $f(x) := P[\{x\}]$ を**質量関数**という.
- (2) 絶対連続確率分布 $\nu \in \mathcal{P}(\mathbb{R}^d)$ に対して, $f(x) := F'(x)$ を**密度関数**という.

1.1.3 一意性定理

Riesz の表現定理 $C_c(\mathbb{R})^* \simeq BV(\mathbb{R})$ により, 実数上の Radon 測度は有界変動関数によって表現されるという事実の系である.

命題 1.1.5 (一意性定理). 任意の確率分布 $\nu_1, \nu_2 \in P(\mathbb{R}^d)$ について, 次は同値:

- (1) $\nu_1 = \nu_2$.
- (2) $F_{\nu_1} = F_{\nu_2}$.

1.1.4 多次元の場合

定義 1.1.6. $X = (X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^n$ を確率変数とする.

- (1) 同時分布関数 $F : \mathbb{R}^n \rightarrow [0, 1]$ を次で定義する:

$$F(x_1, \dots, x_n) := P[X_1 \leq x_1, \dots, X_n \leq x_n].$$

- (2) 同時密度関数と次の関係がある:

$$f(x_1, \dots, x_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F(x_1, \dots, x_n).$$

- (3) $F_i(x_i) := F(\infty, \dots, \infty, x_i, \infty, \dots, \infty) = P[X_i \leq x_i]$ を**周辺分布関数**という.

1.2 確率表現関数

1.2.1 確率表現関数の定義

定義 1.2.1 (quantile function / representing function).

- (1) 累積分布関数 F が連続かつ狭義単調増加であるとき、逆関数 F^{-1} を **分位点関数** または **確率表現関数** と定める。
 (2) 一般の場合、 $F_L^{-1}(u) := \inf \{x \in \mathbb{R} \mid F(x) \geq u\}$, $F_R^{-1}(u) := \sup \{x \in \mathbb{R} \mid P(X \geq x) \geq 1 - u\}$ と定める。

歴史 1.2.2. Moriguti (1953) の内容を Tennessee 州 Knoxville で報告した際の座長の Tukey のコメントからの命名だという [森口繁一, 1995].

補題 1.2.3.

- (1) F_L^{-1} は左連続である。
 (2) F_R^{-1} は右連続である。

定義 1.2.4. 分布関数 $F(x)$ において、 $x \rightarrow \pm\infty$ としたときの収束の速さを **裾の重さ** という。期待値や分散は積分で定義されるために、裾の重さに影響を受けやすい。

1.3 共分散

1.3.1 1次元の共分散

定義 1.3.1 (covariance, correlation coefficient). $X, Y \in L^1(\Omega)$ について、

- (1) $\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])]$ を X, Y の共分散という。
 (2) $\text{Var}[X]\text{Var}[Y] \neq 0$ のとき、 $\rho(X, Y) = \frac{\text{Cov}[X, Y]}{\sqrt{\text{Var}[X]}\sqrt{\text{Var}[Y]}}$ を **相関係数** という。

注 1.3.2. $X, Y \in L^2(\Omega)$ ならば $\text{Cov}[X, Y] < \infty$ だが、 $X, Y \in L^1(\Omega)$ でも、 X, Y が独立ならば $\text{Cov}[X, Y] < \infty$ 。

命題 1.3.3 (1次元分散の性質). $X, Y, Z \in L^2(\Omega)$ について、

- (1) 対称性: $\text{Cov}[X, Y] = \text{Cov}[Y, X]$.
 (2) 双線型性: $\text{Cov}[aX + bY, Z] = a\text{Cov}[X, Z] + b\text{Cov}[Y, Z]$.
 (3) $\text{Cov}[X, 1] = 0$. 特に、 $\text{Cov}[aX + b, Y] = a\text{Cov}[X, Y]$.
 (4) 共分散公式: $\text{Cov}[X, Y] = E[XY] - E[X]E[Y]$.
 (5) 正: $\text{Cov}[X, X] = \text{Var}[X] \geq 0$. 等号成立条件は $X = E[X]$ a.s.
 (6) Schwarz の不等式: $|\text{Cov}[X, Y]| \leq \sqrt{\text{Var}[X]}\sqrt{\text{Var}[Y]}$.

[証明]. (1),(2) は定義に戻って計算するのみ。(3) は 1 の定める σ -代数が 2 であり、これは任意の σ -代数と独立であることから分かる。 ■

命題 1.3.4 (分散の性質). $X, Y \in L^2(\Omega)$ とする。

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

1.3.2 独立性と共分散

命題 1.3.5 (和の分散). $X_1, \dots, X_n \in L^2$ について、

$$\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{(i,j): 1 \leq i < j \leq n} \text{Cov}(X_i, Y_j).$$

命題 1.3.6 (独立ならば共分散は零). $X, Y \in L^1(\Omega)$ を独立とする。このとき、 $\text{Cov}[X, Y] = 0$ である。特に、 $\rho[X, Y] = 0$ 。逆は成り立たない。

例 1.3.7. $U \sim U((0, 1))$ として、 $X := \cos 2\pi U, Y := \sin 2\pi U$ とすると、 $X^2 + Y^2 = 1$ の関係式があるため、独立ではない。例えば、 $X = 1$ の下で Y は $[-1, 1]$ 上の一様分布には従わない。しかし、 $\rho = 0$ である。 □

1.3.3 多次元の共分散

記法 1.3.8 (期待値作用素の拡張). 期待値作用素 E を行列値確率変数 $M = [M_{ij}] \in M_{ij}(\mathbb{R})$ 上に対しても $E : M_{ij}(\mathbb{R}) \rightarrow M_{ij}(\mathbb{R}); E[M] = [E[M_{ij}]]$ と拡張すると, 平均ベクトルは $E[M]$, 共分散行列は $\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])^T]$ と表せる.

定義 1.3.9 (covariance matrix, variance-covariance matrix). $X \in L^1(\Omega; \mathbb{R}^d)$ とする.

(1) X が可積分のとき, 項別積分

$$E[X] = \begin{bmatrix} E[X_1] \\ \vdots \\ E[X_d] \end{bmatrix}$$

を X の平均ベクトルという.

(2) X, Y が2乗可積分のとき,

$$\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])^T] = \begin{bmatrix} \text{Cov}[X_1, Y_1] & \cdots & \text{Cov}[X_1, Y_s] \\ \vdots & \ddots & \vdots \\ \text{Cov}[X_d, Y_1] & \cdots & \text{Cov}[X_d, Y_s] \end{bmatrix}$$

を X, Y の共分散行列という.

(3) $\text{Var}[X] := \text{Cov}[X, X]$ を X の分散共分散行列または分散行列と呼ぶ.^{†1}

(4) $\text{Corr}[X] = (\rho(X_i, X_j))_{i,j \in [n]}$ で定まる $r \times r$ 行列を相関行列という. これは, 対角要素が1に基準化された無次元量だと考えられる.

命題 1.3.10 (多次元分散の性質). $X \in L^2(\Omega; \mathbb{R}^d), Y \in L^1(\Omega; \mathbb{R}^s)$ とする.

- (1) $\text{Cov}[X, Y] = \text{Cov}[Y, X]^T$.
- (2) $\text{Cov}[X, Y] = E[XY^T] - E[X]E[Y]^T$.
- (3) 双線型性: $\text{Cov}[AX + a, BY + b] = A\text{Cov}[X, Y]B^T$.

命題 1.3.11 (確率ベクトルの2次形式の平均). $X, Y \in L^2(\Omega; \mathbb{R}^d), G \in M_d(\mathbb{R})$ とする. このとき,

$$E[X^T G Y] = \text{Tr}(G \text{Cov}[X, Y]) + E[X]^T G E[Y].$$

命題 1.3.12 (共分散行列の双線型性).

(1) 可積分確率変数 $X_1, \dots, X_m, Y_1, \dots, Y_n$ の積 $X_i Y_j$ も可積分とする. このとき,

$$\forall a_1, \dots, a_m, b_1, \dots, b_n \in \mathbb{R} \quad \text{Cov}\left(\sum_{i=1}^m a_i X_i, \sum_{j=1}^n b_j Y_j\right) = \sum_{i=1}^m \sum_{j=1}^n a_i b_j \text{Cov}(X_i, Y_j)$$

(2) 各 X_i ($i \in [m]$) を r_i 次元確率変数, 各 Y_j ($j \in [n]$) を c_j 次元確率変数とする. すべての $X_i, Y_j, X_i Y_j \in L^1$ のとき, 任意の $r \times r_i$ 定数行列 A_i と $c \times c_j$ 定数行列 B_j について

$$\text{Cov}\left(\sum_{i=1}^m A_i X_i, \sum_{j=1}^n B_j Y_j\right) = \sum_{i=1}^m \sum_{j=1}^n A_i \text{Cov}(X_i, Y_j) B_j^T$$

注 1.3.13 (行列の絶対値). 行列 M に関して, $|M| = (\text{Tr}(MM^T))^{1/2}$ とすると, $|X_i \otimes Y_j| = |X_i| |Y_j|$ だから, $|X_i| |Y_j|$ が可積分であることと, $|X_i \otimes Y_j|$ が可積分であることと, X_i, Y_j の要素のペアの積がすべて可積分であることは同値になる.

^{†1} 分散共分散行列からは, データの相関を完全に失わせるような写像を作る変換行列を作ることができる. これは, 違った見方をすれば, データを簡便に記述するのに最適な基底を取っていることになる. (分散共分散行列のその他の性質やその証明については, en:Rayleigh quotient を参照) これは, 統計学では主成分分析 (PCA) と呼ばれており, 画像処理の分野では, カルーネン・レーベ変換 (KL-transform) と呼ばれている.

1.3.4 半正定値行列としての特徴付け

分散共分散行列と半正定値行列は同一視出来る。^a

^a <https://ja.wikipedia.org/wiki/分散共分散行列>

補題 1.3.14 . 分散共分散行列 $\text{Var}[X]$ は半正定値である.

[証明]. $\forall_{u \in \mathbb{R}^d} u^\perp \text{Var}[X] u = E[(u \cdot (X - E[X]))^2]$ より. ■

注 1.3.15 (退化した多次元確率変数). $\text{Var}[X]$ が正定値でないとすると, $\exists_{u \in \mathbb{R}^d \setminus \{0\}} u^\perp \text{Var}[X] u = E[(u^\perp (X - E(X)))^2] = 0$ である. すなわち, X は確率1で超平面 $u^\perp (X - E(X)) = 0$ 上に値を取る.

1.4 積率母関数

1.4.1 積率母関数の定義と収束性

積率母関数は左側半平面を定義域に持ち, これが0を超えると, 虚軸を含むために特性関数の解析接続でもある. 関係としては $\varphi(u) = M(iu)$.

定義 1.4.1 (moment generating function). $(\mathbb{R}, \mathcal{B}_1)$ 上の確率測度 $\nu \in P(\mathbb{R})$ に対して,

$$M_\nu(t) := \int_{\mathbb{R}} e^{tx} \nu(dx)$$

により定まる関数 $\mathbb{R} \rightarrow \mathbb{R}$ を **積率母関数** という. 収束域を $\text{Dom}(M_\nu) \subset \mathbb{R}$ で表す.

定理 1.4.2 (収束域は左半平面をなす). 分布 ν は \mathbb{R}_+ 上に台を持ち, 積率母関数 M はある $x \in \mathbb{R}$ にて収束するとする. このとき,

- (1) ある $x \leq a$ が存在して, M は $\text{Dom}(M) = (-\infty, a)$ 上で収束する.
- (2) M を複素関数と見ると, 半平面 $\{z \in \mathbb{C} \mid \text{Re } z < a\}$ 上でも収束する.
- (3) さらに, 分布が Lebesgue 測度に対して絶対連続である $\nu \in \mathcal{P}(\mathbb{R}^d)$ とき, M は半平面 $\{z \in \mathbb{C} \mid \text{Re } z < a\}$ 上正則である.

観察 1.4.3. 積率母関数が0の近傍での存在が危うくなるのは, 0を中心として両側に裾が重い場合である. 離散分布で考えれば, Poisson 分布など片側の場合は全く問題外だが, $p(x) = \frac{C}{x^2}$ ($x \in \mathbb{Z} \setminus \{0\}$) という確率質量関数を考えると, これは $\text{Dom}(M) = \{0\}$ となる. □

命題 1.4.4 (指数減衰するならば原点の近傍で収束する). 分布 $\nu \in P(\mathbb{R})$ について, 次の2条件は同値:

- (1) M_ν は0の近傍で存在する.
- (2) ν は指数減衰する: $\exists_{C, t_0 > 0} \forall_{x \in \mathbb{R}} P[|X_\nu| > x] \leq C e^{-t_0 x}$.

1.4.2 確率母関数の Maclaurin 展開

積率母関数は正則だから, 0を定義域に含んだ瞬間に任意階数の積率が存在する.

定理 1.4.5 (マクローリン展開係数は積率である). $\nu = P^X$ の積率母関数 M_ν が原点の近傍で存在するとする. すると定義域上 C^∞ 級であり,

$$\forall_{n \in \mathbb{N}^+} \partial^n M_\nu(t) = E[X^n e^{tX}].$$

特に, $\forall_{n \in \mathbb{N}^+} \partial^n M_\nu(0) = \alpha_n$ で,

$$(1) \alpha_1 = \partial M_v(0) = \partial(\log M_v)(0) = \partial \mathfrak{C}_v(0).$$

$$(2) \mu_2 = \partial^2(\log \mathfrak{M}_v)(0) = \partial^2 \mathfrak{C}_v(0).$$

[証明].

$$M_v(t) = \int_{\mathbb{R}} e^{tx} v(dx)$$

の両辺を t で n 階微分すると,

$$\frac{d^n M(t)}{dt^n} = \int_{\mathbb{R}} x^n e^{tx} v(dx).$$

よって $t = 0$ を代入すると,

$$\frac{d^n M(0)}{dt^n} = \int_{\mathbb{R}} x^n v(dx) = E[X^n] = \alpha_n.$$

■

注 1.4.6 (確率母関数との関係). なお, 積率母関数は, 確率母関数 $G(z) := E[z^X]$ に対して $G[e^t] = M_X(t)$ なる関係もある.

1.4.3 一意性定理

\mathbb{R}_+ 上の分布に限っては, 特性関数と同等に Laplace 変換も扱う. コンパクト台を持つ超関数 $\mathcal{S}'(\mathbb{R}^n) \subset \mathcal{S}'(\mathbb{R}^n)$ の Fourier 変換は, 整関数に延長する. さらに繊細な消息として, Paley-Wiener の定理により, $L^2(\mathbb{R}_+)$ の元は, 上半平面上の正則関数であって指数型の増加速度を持つ関数のクラスに対応する.

定理 1.4.7 (特性関数との関係と一意性定理). $v \in P(\mathbb{R}^d)$ とする.

- (1) $\text{Dom}(M_v)$ が $0 \in \mathbb{R}^d$ の近傍ならば, ある整数 ϵ が存在して, 積率母関数 M_v は $\mathcal{R}_\epsilon := ((-\epsilon, \epsilon) \times i\mathbb{R})^d$ 上の正則関数 M_v^\dagger に解析接続され, $\varphi_v(-) = M_v^\dagger(i-)$ が成り立つ.
- (2) 正則関数 M_v^\dagger は, $z \in \mathcal{D}_\epsilon := \{\zeta \in \mathbb{C} \mid |\zeta| < \epsilon\}^d$ 上において, 絶対収束級数

$$M_v^\dagger(z) = \sum_{k=1}^{\infty} \frac{1}{k!} \int (z \cdot x)^k v(dx)$$

なる展開を持つ.

- (3) $v_1, v_2 \in P(\mathbb{R}^d)$ のとき, 原点のある開近傍 $U \subset D_{v_1} \cap D_{v_2}$ 上で $M_{v_1} = M_{v_2}$ ならば, $v_1 = v_2$ が成り立つ.

1.5 キュムラント母関数

積率母関数の対数をキュムラント母関数という. 一方で, 特性関数の対数は第2キュムラント母関数という.

1.5.1 キュムラント母関数

積率母関数の対数の展開係数をキュムラントといい, この特徴量の列も分布を特徴付ける.

定義 1.5.1 (cumulant generating function). $v \in P(\mathbb{R}^d)$ に対して, $C_v(t) := \log M_v(t)$ で定まる $\text{Dom}(M_v)$ 上の関数を, F のキュムラント母関数という.

歴史 1.5.2. キュムラントのことは, その性質を研究した Thorvald N. Thiele に因み, ティエレの半不変数 (semi-invariant) とも呼ぶ。^{†2} 積率との繋がりが深い.

^{†2} 統計学の分野で尤度に関する初期の考察を行い, 保険数学の分野で Hafnia 保険会社を設立し, 数学部長を務め, デンマーク保険統計協会を設立した。

命題 1.5.3.

$$\kappa_n = \frac{\partial^n}{\partial t^n} C_v(t) \Big|_{t=0} = \frac{\partial^n}{\partial t^n} \log M_v(t) \Big|_{t=0}$$

注 1.5.4 (affine 変換に対する半不変性). 半不変数というのは, 積率は $Y = \alpha X + \beta$ なる affine 変換に対して $E[Y^2] = \alpha^2 E[X^2] + 2\alpha\beta E[X] + \beta^2$ なるように変換されるのに対して,

$$\kappa_1^Y = \alpha \kappa_1^X + \beta, \quad \kappa_n^Y = \alpha^n \kappa_n^X \quad (n = 2, 3, \dots)$$

というように殆ど形を変えないことにちなむ.

例 1.5.5 (キュムラント母関数は極めて少ししか変わらなくても, 密度関数は見かけ上全く異なる例). □

1.5.2 積率問題

n 次の積率の列 $(\alpha_n := E[X^n])$ に対して, これを持つ確率分布は構成できるか? という問題は, 極限定理に関する Chebyshev P. L. (1874) の漸近論の研究で初めて議論された ([Hazewinkel, 1995] p. 945, Prokhorov, A.V. 担当部分). 積率法を用いた中心極限定理の証明は Markov (1898) によって達成された. 積率問題の解の存在は, Hahn-Banach の定理の応用である. ここでは確率的な文脈であるが, 物理的にも, 電荷密度 ρ に関する条件として読める.

問題 1.5.6 (moment problem). $\alpha_n = 0$ と満たす実数列 $\alpha: \mathbb{N} \rightarrow \mathbb{R}$ に対して, これを積率の列に持つ分布 P

$$\alpha_n = \int x^n P(dx).$$

が一意に定まるか?

歴史 1.5.7. 続いて Stieltjes (1894) が無限連分数に関連して, 任意の実数列 $\{\mu_n\}_{n \in \mathbb{N}}$ に対して, \mathbb{R}_+ 上の単調増加な有界関数 ψ であって

$$\int_0^\infty x^n d\psi = \mu_n, \quad (n \in \mathbb{N})$$

を求めるとい問題を定式化し, 補間理論として成長した.

命題 1.5.8. 任意の有限個への制限 $\alpha|_{[N]}$ について, これを満たす積率問題の解は存在する.

定理 1.5.9 (存在の必要十分条件). 区間 $[a, b]$ 上の分布であって, 積率 α を持つものを考える. 次の 2 条件は同値.

- (1) 積率問題は解を持つ.
- (2) $\exists c > 0 \quad \forall N \in \mathbb{N} \quad \forall \alpha_k \in \mathbb{R} \quad \left| \sum_{k=0}^N \alpha_k \alpha_k \right| \leq c \max_{a \leq x \leq b} \left| \sum_{k=0}^N \alpha_k x^k \right|.$

定理 1.5.10 (一意性の十分条件). $(\alpha_n)_{n \in \mathbb{N}}$ が次の条件のいずれかを満たすならば, α を積率とする分布はただ一つである.

- (1) $\exists t_0 > 0 \quad \sum_{n=1}^{\infty} \frac{|\alpha_n|}{n!} t_0^n < \infty.$
- (2) Carleman 条件 [Hazewinkel, 1995]: $\sum_{n=1}^{\infty} \frac{1}{\alpha_{2n}^{1/2n}} = \infty.$

例 1.5.11. 3 次以上のキュムラントがすべて 0 になる確率分布は, 正規分布に限る. □

命題 1.5.12 (極限定理・モーメント法への応用). 累積分布関数の列 (F_n) は, 任意の $k \geq 1$ 位の積率 $\alpha_k^n < \infty$ が存在するとする. 任意の $k \in \mathbb{N}$ について, 極限 $\alpha_k^n \xrightarrow{n \rightarrow \infty} \alpha_k < \infty$ が存在するならば, ある部分列 $(F_{n_j})_{j \in \mathbb{N}}$ が存在して, (α_k) を積率の列としてもつ分布 F に弱収束する. 列 (α_k) を積率とする分布がただ一つならば, $n_j = j$ と取れる.

1.6 離散分布の確率母関数

$g(z) := E[z^X]$ は, $\varphi(u) = g(e^{iu})$, $M(t) = g(e^t)$ の関係を持つ.

1.6.1 離散分布の特性関数と確率母関数

定義 1.6.1 (characteristic function, probability generating function). μ を離散分布, $(p_x)_{x \in \mathbb{Z}}$ をその質量関数とする.

(1) 次で定まる関数 $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ を, 確率関数 μ の**特性関数**または **Fourier 変換**という:

$$\varphi(u) := E[e^{iuX}] = \sum_{x \in \mathbb{Z}} e^{iux} p_x$$

(2) $0 \in \mathbb{C}$ の近傍上の関数 $g(z) := \sum_{x \in \mathbb{N}} p_x z^x$ を**確率母関数**という. これは $g(1) = 1$ を満たす.

注 1.6.2 (積率母関数の存在). 確率母関数を考えるような $\mathcal{G} = \mathbb{N}$ という場合, $X \geq 0$ であるから, 積率母関数は $\{x \in \mathbb{R} \mid x \leq 0\}$ 上で存在する. 確率母関数の収束半径は $R = \left(\limsup_{n \rightarrow \infty} |p_n|^{1/n} \right)^{-1}$ で与えられる. これが 1 よりも真に大きいとき, すなわち $\limsup_{n \rightarrow \infty} |p_n|^{1/n} < 1$ のとき, $g(z)$ は ± 1 でも収束するから, 積率母関数 $M(t) = g(e^t)$ は 0 の近傍で存在する.

要諦 1.6.3 (確率母関数の名前の由来). 名前の由来は, 各 $p(x)$ が z^x の係数として格納されているために, $p(x) = \frac{g^{(x)}(0)}{x!}$ として確率質量関数を復元出来ることからいう.

1.6.2 確率母関数の微分係数

確率母関数の微分からは階乗モーメントを得る.

補題 1.6.4 (fractional moment: 分散公式・積率は特性関数の微分・平均と分散は確率母関数の微分).

(1) 分散公式: $\sigma^2 = \alpha_2 - \mu^2 = \sum_{x \in \mathcal{G}} x^2 p_x - \left(\sum_{x \in \mathcal{G}} x p_x \right)^2$.

(2) 特性関数の積率の関係: $\beta_r < \infty$ のとき, $\alpha_r = i^{-r} \varphi^{(r)}(0)$. 積率母関数が原点の近傍で存在するならば, $\alpha_r = M^{(r)}(0)$.

(3) g の収束半径が 1 より大きい時, 項別微分が可能で,

$$g^{(r)}(1) = \sum_{x=r}^{\infty} x(x-1) \cdots (x-r+1) p_x.$$

特に, $g''(1) = \alpha_2 - \alpha_1^2$ で, $g'''(1) = \alpha_3 - 3\alpha_2 + 2\alpha_1^2$ で,

(a) $\alpha_1 = g'(1)$.

(b) $\alpha_2 = g''(1) + g'(1)^2$.

(c) $\alpha_3 = g'''(1) + 3g''(1)g'(1) + g'(1)^3$.

(4) 特に平均と分散について

(a) $\alpha_1 = g'(1)$.

(b) $\mu_2 = g''(1) + g'(1)^2 - (g'(1))^2 = \alpha_2 - (\alpha_1)^2$.

が成り立つ.

[証明].

(1)

$$\sigma^2 = \sum_{x \in \mathcal{G}} (x - \mu)^2 p_x$$

$$\begin{aligned}
&= \sum_{x \in \mathcal{X}} x^2 p_x - 2\mu \underbrace{\sum_{x \in \mathcal{X}} x p_x}_{=\mu} + \mu^2 \underbrace{\sum_{x \in \mathcal{X}} p_x}_{=1} \\
&= \alpha_2 - \mu^2.
\end{aligned}$$

(2) $\beta_r < \infty$ ならば, $\varphi^{(r)}(u) = \sum_{x \in \mathcal{X}} (ix)^r e^{iux} p_x$ は収束し (すなわち項別微分可能で), 関数 $\varphi^{(r)}: \mathbb{R} \rightarrow \mathbb{R}$ を定める. $u = 0$ とし, $\varphi^{(r)}(0) = i^r \sum_{x \in \mathcal{X}} x^r p_x = i^r \alpha_r$.

(3) 一般に, 項別微分可能ならば $g^{(r)}(z) = \sum_{x=0}^{\infty} x(x-1)\cdots(x-r+1)p_x z^{x-r}$ であるが, いま $\mathcal{X} = \mathbb{Z}$ としているから, $z = 1$ のとき, $g^{(r)}(1) = \sum_{x=r}^{\infty} x(x-1)\cdots(x-r+1)p_x$ である.

(4) また, 特に $g'(1) = \mu$, $g''(1) = \sum_{x \in \mathcal{X}} x(x-1)p_x$ であるから, (1) より, $\sigma^2 = \underbrace{g''(1) + g'(1)}_{=\sum(x(x-1)+x)p_x} - (g'(1))^2$

■

1.7 特性関数

1.7.1 特性関数の定義

定義 1.7.1 (characteristic function). $(\mathbb{R}, \mathcal{B}_1)$ 上の確率測度 ν に対して,

$$\varphi(u) := \int e^{iux} \nu(dx)$$

により定まる関数 $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ を**特性関数**という.

命題 1.7.2 (確率変数の演算との対応). X の特性関数を φ とする.

- (1) $\frac{X-\mu}{\sigma^2}$ の特性関数は, $\varphi\left(\frac{u}{\sigma^2}\right) e^{-i\mu u/\sigma^2}$.
- (2) P^X が対称であることと, φ も対称な実数値関数となることは同値.

1.7.2 特性関数と分布の収束の対応

特性関数と分布関数の間の対応は, 次の意味で各点収束位相-弱収束位相について「連続」である.

定理 1.7.3 (Glivenko の連続定理). 確率測度 ν_n の特性関数を φ_n とする. (1) \Rightarrow (2) である. 各点収束極限 φ が原点で連続であるとき (したがってある分布の特性関数であるとき), (2) \Rightarrow (1) でもある (Glivenko).

- (1) ある確率測度 ν について, $\nu_n \Rightarrow \nu$ である.
- (2) φ_n が関数 φ に各点収束する.

なお, このとき (2) の収束はコンパクト一様に起こる.

[証明].

(1) \Rightarrow (2) 各点収束は明らかだから, 広義一様収束を示す. このとき, $\{\mu_n\} \cup \{\mu\}$ は一様に緊密であるから, 任意の $\epsilon > 0$ に対して, $R > 0$ が存在して,

$$\mu(B_R(0)^c) \leq \epsilon, \quad \mu_n(B_R(0)^c) \leq \epsilon, \quad n \in \mathbb{N}.$$

そこで, $B_{R+1}(0)$ 上に台を持ち, $B_R(0)$ 上で 1 な連続関数を $\varphi \in C_c(\mathbb{R}^d; [0, 1])$ とし,

$$\varphi(z) = \int_{\mathbb{R}^d} e^{i(x|z)} \varphi(x) d\mu_n(x) + \int_{\mathbb{R}^d} e^{i(x|z)} (1 - \varphi(x)) d\mu_n(x) =: \varphi_n^1(z) + \varphi_n^2(z).$$

と分解する. すると,

$$|\varphi_n^1(z) - \varphi_n^1(y)| = \left| \int_{\mathbb{R}^d} (e^{i(x|z)} - e^{i(x|y)}) \varphi(x) d\mu_n(x) \right| \leq C_R |z - y|.$$

と評価でき, これは $n \in \mathbb{N}$ に依らない評価であり, φ_n^1, φ^1 は同程度連続であることを意味する. さらに, φ_n^1 は φ^1 に各点収束し, 特に各点有界であるから, Ascoli-Arzelà の定理より $\varphi_n^1 \rightarrow \varphi^1$ は広義一様. 一方で, 任意の $z \in \mathbb{R}^d$ について, $|\mu_n^2(z)| \leq \epsilon$ かつ $|\mu^2(z)| \leq \epsilon$.

(2) \Rightarrow (1) $\varphi_n \rightarrow \varphi$ は各点収束するため, 任意の $f \in S(\mathbb{R}^d)$ に対して,

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} \hat{f}(z) \varphi_n(z) dz = \int_{\mathbb{R}^d} \hat{f}(z) \varphi(z) dz = \int_{\mathbb{R}^d} f(x) d\mu(x).$$

よって,

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}^d} f d\mu_n = \int_{\mathbb{R}^d} f d\mu, \quad f \in S(\mathbb{R}^d).$$

これは任意の $f \in C_c(\mathbb{R}^d)$ についても成立する. ■

定理 1.7.4 (Levy の収束定理). $\varphi_n := \mathcal{F}\mu_n$ が各点で φ に収束し, かつ, ある $a > 0$ が存在して, $|x| < a$ 上一様ならば, 極限関数 φ はある分布 μ の特性関数である.

補題 1.7.5. $\mu \in P(\mathbb{R})$ の特性関数を φ とする.

$$\forall h > 0 \quad h \left(\left[-\frac{2}{h}, \frac{2}{h} \right] \right) \geq \mu(\mathbb{R}) + \frac{1}{h} \int_{-h}^h (\varphi(t) - \varphi(0)) dt.$$

例 1.7.6.

$$\nu_n(A) := \sum_{j=1}^n \frac{1}{n} \delta_{\frac{j}{n}}(A) \quad (A \in \mathcal{B}(\mathbb{R}))$$

は一様分布 $U(0,1)$ に弱収束する. 実際, 積分の定義より,

$$\int_{\mathbb{R}} g(x) \nu_n(dx) = \sum_{j=1}^n \frac{1}{n} g\left(\frac{j}{n}\right) \xrightarrow{n \rightarrow \infty} \int_0^1 g(x) dx.$$

□

1.7.3 特性関数の滑らかさと積率の存在

絶対積率が n 位まで存在するならば, 特性関数は C^n -級である. 逆も部分的に成り立つ.

定理 1.7.7 (確率変数が L^p ならば特性関数は C^p). $X \in L^p(\mathbb{R}^d)$, すなわち P^X の p 次の絶対積率が存在する $\beta_p < \infty$ ならば, 次の 2 条件が成り立つ:

- (1) $\varphi_X \in C^p(\mathbb{R}^d)$.
- (2) $\varphi_X^{(p)}$ は一様連続.
- (3) $\forall m \in \mathbb{N}^d \quad |m| = p \Rightarrow D^m \varphi_X(u) = i^p E[X^m e^{iu \cdot X}]$. 特に, $\varphi^{(p)}(0) = i^p \alpha_p$.

m は多重指数であることに注意.

[証明].

- (1) 積分と微分の交換が可能であることを示せば良い. いま,

$$\partial_{z_1}^{\alpha_1} \cdots \partial_{z_d}^{\alpha_d} e^{i(x|z)} = \left(\prod_{k=1}^d (ix_k)^{\alpha_k} \right) e^{i(x|z)}.$$

であるが,

$$|\partial_{z_1}^{\alpha_1} \cdots \partial_{z_d}^{\alpha_d} e^{i(x|z)}| \leq \prod_{k=1}^d |x_k|^{\alpha_k}.$$

より, $\beta_r < \infty$ はこの右辺が可積分になるための条件を与えている.

■

命題 1.7.8 (特性関数の滑らかさが示唆する絶対積率の存在). 分布 ν の特性関数 φ が $u = 0$ において $2r$ 階微分可能であるとする.

このとき, $|\beta_{2r}| < \infty$. 特に, $\beta_k = \frac{\phi^{(k)}(0)}{i^k}$ ($k \in [2n]$).

例 1.7.9 (奇数次のモーメントについては成り立たない [盛田健彦, 2004]). Borel 測度

$$\mu = \sum_{k=2}^{\infty} \frac{1}{k^2 \log k} (\delta_{-k} + \delta_k)$$

を考えると, 1 次のモーメントを持たないが, 原点の近傍で C^1 -級な特性関数

$$\varphi(t) = 2 \sum_{k=2}^{\infty} \frac{\cos kt}{k^2 \log k}$$

をもつ.

□

1.7.4 特性関数の Maclaurin 展開

特性関数を Taylor 展開することで, 高次の積率 α_n も求めることが出来る.

定理 1.7.10. $\nu \in P(\mathbb{R}^d)$ は $\beta_r < \infty$ を満たすとする. このとき, 次が成り立つ:

(1) $|n| \leq r$ を満たす多重指数 $n \in \mathbb{N}^d$ について,

$$\partial^n \varphi(u) = \int_{\mathbb{R}^d} (ix)^n e^{iu \cdot x} \nu(dx).$$

(2) 次の不等式が成り立つ, ただし $\beta_r(u) = \int_{\mathbb{R}^d} |u \cdot x|^r \nu(dx)$ とする.

$$\left| \varphi(u) - \sum_{|n| < r} \frac{(iu)^n}{n!} \alpha_n \right| \leq \frac{|u|^r}{r!} \beta_r(u) \quad (u \in \mathbb{R}^d).$$

[証明].

(1) Lebesgue の定理による.

(2) 次の等式による:

$$e^{it} = \sum_{j=0}^{r-1} \frac{(it)^j}{j!} + \frac{(it)^r}{(r-1)!} \int_0^1 (1-s)^{r-1} e^{its} ds.$$

■

系 1.7.11 (特性関数の Taylor 展開係数からの積率の計算). $\beta_r < \infty$ とする. このとき, φ は $u = 0$ の周りで展開

$$\varphi(u) = \sum_{n=0}^r \alpha_n \frac{(iu)^n}{n!} + o(u^r) \quad (u \rightarrow 0)$$

を持つ. すなわち, 中心積率は特性関数を用いて次のように表せる:

$$\alpha_n = \frac{1}{i^n} \frac{d^n \varphi(u)}{du^n} \Big|_{u=0}$$

系 1.7.12 (平均と分散の特性関数による特徴付け).

- (1) 平均 (1 次の積率) は $\alpha_1 = \frac{1}{i} \varphi'(0)$.
 (2) 分散 (2 次の中心積率) は $\mu_2 = -\varphi''(0) + (\varphi'(0))^2$.

1.7.5 特性関数の減衰速度と密度関数の可積分性との対応

系 1.7.13 (Levy の反転公式の系). $\mu \in P(\mathbb{R})$ の特性関数 φ が $L^1(\mathbb{R})$ ならば, μ は絶対連続であり, 密度関数が微分

$$\frac{d\mu}{dm}(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} \varphi(t) dt$$

で得られる.

命題 1.7.14 (特性関数の増加が $o(x^{-m})$ ならば, 確率密度関数は C_0^m). $\nu \in P(\mathbb{R}^d)$ の特性関数の増加速度

$$\int |u|^m |\varphi_F(u)| du < \infty$$

を満たすほど遅いとする. このとき, ν の確率密度関数

$$f(x) = \frac{1}{(2\pi)^d} \int e^{-iu \cdot x} \varphi_F(u) du$$

は

- (1) C^m -級で,
 (2) $\forall k \leq m \ f^{(k)} \in C_0(\mathbb{R}^d)$ を満たす.

[証明]. Fourier 変換の一般論による. ■

定理 1.7.15 (Marcinkiewicz). $Q(z)$ を m 次の多項式とする. $\varphi(z) = e^{Q(z)}$ が特性関数ならば, $m \leq 2$.

要諦 1.7.16. 特性関数が有限位数の零点を持たない整関数であるとき, 整関数論と Hadamard の因数分解定理から, 多項式 $Q(z)$ を用いて $\varphi(z) = e^{Q(z)}$ と表せることがわかる.

1.7.6 反転公式と一意性定理

補題 1.7.17 (Levy's inversion formula). 確率測度 $\mu \in P(\mathbb{R}^d)$ とその特性関数 φ について,

- (1) $A := \prod_{n=1}^d [a_n, b_n]$ が μ -連続集合であるならば,

$$\mu(A) = \lim_{R \rightarrow \infty} \frac{1}{(2\pi)^d} \int_{[-R, R]^d} \prod_{n=1}^d \frac{e^{-iz_n a_n} - e^{-iz_n b_n}}{iz_n} \varphi(z) dz.$$

- (2) $d = 1$ のとき, 任意の区間 $[a, b] \subset \mathbb{R}$ について,

$$\mu((a, b)) + \frac{\mu(\{a\}) + \mu(\{b\})}{2} = \lim_{R \rightarrow \infty} \frac{1}{2\pi} \int_{-R}^R \frac{e^{-iza} - e^{-izb}}{iz} \varphi(z) dz.$$

とすればよい.

[証明]. (2) を示す. まず右辺は

$$\begin{aligned} \frac{1}{2\pi} \int_{-R}^R \frac{e^{-iza} - e^{-izb}}{iz} \varphi(z) dz &= \frac{1}{2\pi} \int_{-R}^R \frac{e^{-iza} - e^{-izb}}{iz} \left(\int_{\mathbb{R}} e^{izx} \mu(dx) \right) dz \\ &= \frac{1}{\pi} \int_{\mathbb{R}} \frac{1}{2} \int_{-R}^R \frac{e^{iz(x-a)} - e^{iz(x-b)}}{iz} dz d\mu(x) \end{aligned}$$

$$= \frac{1}{\pi} \int_{\mathbb{R}} \underbrace{\int_0^R \left(\frac{\sin(x-a)z}{z} - \frac{\sin(x-b)z}{z} \right) dz}_{=: I_{a,b,R}} d\mu(x).$$

と表せるが、 $I_{a,b,R}$ は任意の $R > 0$ について可積分で

$$\lim_{R \rightarrow \infty} I_{a,b,R}(x) = \begin{cases} \pi & x \in (a, b), \\ \pi/2 & x \in \partial[a, b], \\ 0 & x \notin [a, b]. \end{cases}$$

より、Lebesgue の優収束定理から結論を得る. ■

系 1.7.18. $\varphi_\mu \in L^1(\mathbb{R}^d)$ ならば、 μ は Lebesgue 測度に関して絶対連続である.

[証明].

$$A := \prod_{i=1}^d [a_i, b_i]$$

を任意にとる. これが μ -連続集合であるとき、 $R > 0$ を十分大きくとれば Levy の反転公式より、

$$\mu(A) = \int_A \underbrace{\int_{\mathbb{R}^d} \frac{e^{-i(z|x)}}{(2\pi)^d} \varphi(z) dz}_{=: f(x)} dx$$

であり、 $f \geq 0$ a.e. 一般の A についても、これに収束する μ -連続集合の列が存在するから、Lebesgue の優収束定理から同様の式が示せる. ここで、任意の $R > 0$ について

$$\mathcal{F} := \left\{ \prod_{i=1}^d (a_i, b_i] \subset \mathbb{R}^d \mid a_i \leq b_i \subset [-R, R] \right\}$$

は乗法族で $\sigma[\mathcal{F}] = \mathcal{B}((-R, R]^d)$ を満たす. $R \rightarrow \infty$ の極限について考えて結論を得る. ■

定理 1.7.19 (一意性定理). 特性関数の全体と確率測度の全体とに標準的な全単射が存在する. すなわち、任意の $\mu_1, \mu_2 \in P(\mathbb{R}^d)$ について、次は同値.

- (1) $\mu_1 = \mu_2$.
- (2) $\varphi_{\mu_1} = \varphi_{\mu_2}$.

[証明]. $d = 1$ の場合について、2通りの方法で示す.

反転公式による証明 $\varphi_\mu = \varphi_{\tilde{\mu}}$ とする. 反転公式 1.7.17 により、任意の (端点が $F_\mu, F_{\tilde{\mu}}$ の連続点である) 开区間の測度は φ が一意に定める: $F_\mu(b) - F_\mu(a) = F_{\tilde{\mu}}(b) - F_{\tilde{\mu}}(a)$. $b \in \mathbb{R}$ が $F_\mu, F_{\tilde{\mu}}$ 両方の連続点であるとき、 $F_\mu(b) = F_{\tilde{\mu}}(b)$. 分布関数は右連続で、不連続点は高々可算個だから、これは $F_\mu = F_{\tilde{\mu}}$ を含意する. 分布関数の一意性より、 $\mu = \tilde{\mu}$.

緩増加分布上の Fourier 変換による証明 $P(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d)$ であることから直ちに従う. これは、任意の $\varphi \in \mathcal{S}$ に対して、

$$|(\mu|\varphi)| \leq \int_{\mathbb{R}^d} |\varphi(x)| d\mu(x) \leq \|\varphi\|_\infty = \|\varphi\|_0.$$

よって、 $P(\mathbb{R}^d)$ の元は階数 0 の緩増加分布である. ■

要諦 1.7.20. というのも、極めて一般的に、 \mathbb{R} 上の有限な Borel 測度と原点で連続な \mathbb{R} 上の正の定符号関数との間に全単射が存在する. 全射性は Bochner の定理からわかる.

系 1.7.21. L^2 -確率変数 $X \in \mathcal{L}^2$ がある $\alpha > 0$ について $\varphi_X(u) = \exp(-|u|^\alpha)$ ならば、 $\alpha = 2$ で、 $X \sim N(0, 2)$ である.

1.7.7 特性関数の特徴付け

一様連続な正定値関数は、ある Borel 測度の Fourier 逆変換である。従って、後は $\varphi(0) = 1$ を課せば、ある確率分布の特性関数であることがわかる。

補題 1.7.22. φ を μ の特性関数とする。

$$|\varphi(z+h) - \varphi(z)| \leq \sqrt{2|\varphi(0) - \varphi(h)|}.$$

この不等式評価は、一般の正の定符号関数 φ について成り立つ。

[証明].

$$\begin{aligned} |\varphi(z+h) - \varphi(z)| &= \left| \int_{\mathbb{R}^d} e^{i(z+h|x)} (e^{i(h|x)} - 1) d\mu(x) \right| \\ &\leq \int_{\mathbb{R}^d} |e^{i(h|x)} - 1| d\mu(x) \\ &\leq \left(\int_{\mathbb{R}^d} |(1 - \cos(h|x))^2 + \sin^2(h|x)| d\mu(x) \right)^{1/2} \\ &= \left(\int_{\mathbb{R}^d} 2(1 - \cos(h|x)) d\mu(x) \right)^{1/2}. \end{aligned}$$

一方で、

$$\begin{aligned} |\varphi(h) - \varphi(0)| &= \left| \int_{\mathbb{R}^d} e^{i(h|x)} d\mu(x) - 1 \right| \\ &= \left| \int_{\mathbb{R}^d} (1 - \cos(h|x)) d\mu(x) - i \int_{\mathbb{R}^d} \sin(h|x) d\mu(x) \right| \\ &\geq \int_{\mathbb{R}^d} (1 - \cos(h|x)) d\mu(x). \end{aligned}$$

命題 1.7.23. 関数 $\varphi : \mathbb{R}^d \rightarrow \mathbb{C}$ が特性関数であるならば、次の3条件が成り立つ。

- (1) $\varphi(0) = 1$.
- (2) 一様連続である。
- (3) 正定値である：任意の $n \in \mathbb{N}$ 個の $z_1, \dots, z_n \in \mathbb{R}^d$ と $\xi_1, \dots, \xi_n \in \mathbb{C}$ について、

$$\sum_{j,k=1}^n \varphi(z_j - z_k) \xi_j \overline{\xi_k} \geq 0.$$

[証明].

- (1) $\varphi(0) = E[e^{i0X}] = E[1] = 1$.
- (2) 次の不等式評価が成り立つ。Lebesgue の優収束定理より、 $h \rightarrow 0$ のとき $u \in \mathbb{R}$ に依らず一様に 0 に収束する。

$$|\varphi(u+h) - \varphi(u)| \leq E[|e^{iuX}| |e^{ihX} - 1|] = E[|e^{ihX} - 1|]$$

(3)

$$\begin{aligned} \sum_{j,k=1}^n \varphi(\xi_j - \xi_k) z_j \bar{z}_k &= \sum_{j,k=1}^n \int_{\mathbb{R}^d} e^{i(z_j - z_k|x)} \xi_j \overline{\xi_k} d\mu(x) \\ &= \int_{\mathbb{R}^d} \sum_{j,k=1}^n e^{i(z_j|x)} \xi_j \overline{e^{i(z_k|x)} \xi_k} d\mu(x) \\ &= \int_{\mathbb{R}^d} \left| \sum_{k=1}^n e^{i(z_k|x)} \xi_k \right|^2 d\mu(x) \geq 0. \end{aligned}$$

定理 1.7.24 (Bochner). 複素関数 $\varphi: \mathbb{R}^d \rightarrow \mathbb{C}$ が次の3条件を満たすならば、ある一次元分布 $\mu \in \mathcal{P}(\mathbb{R})$ の特性関数である。特に、命題の3条件を満たすならば、特性関数である。

- (1) $\xi = 0$ で連続.
- (2) $\varphi(0) = 1$.
- (3) 正定値である.

要諦 1.7.25. 一般の正の定符号関数 φ に関して成り立つものであり、次のように述べられる：関数 $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ が原点で連続な正の定符号関数ならば、ある有限 Borel 測度 $\mu \in M(\mathbb{R}; \mathbb{R}_+)$ が存在して、 $\varphi = \varphi_\mu$ を満たす。

1.7.8 特性関数の構成

[伊藤清, 1991] による。

定義 1.7.26. 確率空間 (A, ν) 上の可積分関数族 $\{F_a\}_{a \in A} \subset L^1(\mathfrak{E})$ について、

$$F(\xi) := \int_A F_a(\xi) \nu(da)$$

を積分的凸結合という。

定理 1.7.27. 分布の積分的凸結合は分布であり、特性関数の積分的凸結合は特性関数である。

定理 1.7.28 (Polya). $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ が次を満たすならば、特性関数である。

- (1) 非負な偶関数である.
- (2) $x > 0$ 上減少かつ凸である.
- (3) $\varphi(0) = 1$ かつ 0 は φ の連続点である.

1.8 第2キュムラント母関数

特性関数の Taylor 展開係数には、積率 α_r が登場した。キュムラント関数の Taylor 展開係数を、キュムラントと言い、これが平均・分散の一般化となっている。存在するならば積率母関数の \log を取れば、 i 倍が消えて純粋なキュムラントを扱うことが出来るが、第2キュムラント母関数には常に存在するという理論的な利点がある。

1.8.1 第2キュムラント母関数

キュムラントはキュムラント母関数＝積率母関数の対数の展開係数として得られるが、特性関数の対数からも、虚数単位 i に関する修正を施して得られる。その関係は、特性関数と積率の関係と並行である。この関数を、第2キュムラント母関数という。

命題 1.8.1 (キュムラント関数の Maclaurin 展開). $\beta_r < \infty$ とする。このとき、関数 $\psi(u) := \log \varphi(u)$ は $u = 0$ の近傍で定まり、次の形の展開を持つ：

$$\psi(u) = \sum_{j=1}^r \kappa_j \frac{(iu)^j}{j!} + o(u^r) \quad (|u| \rightarrow 0).$$

定義 1.8.2 (cumulant). 関数 $\psi(u) := \log \varphi(u)$ を第2キュムラント母関数またはキュムラント関数といい、この展開係数 $\kappa_n = \frac{1}{i^n} \frac{\partial^n}{\partial u^n} \psi(u) \Big|_{u=0}$ を、分布 ν の n 次キュムラントと呼ぶ。この $n \in \mathbb{N}$ は、多重指数 $n \in \mathbb{N}^d$ と読み替えれば、多次元の場合にも通用する。分布の高次の形態を表す特性値である。

例 1.8.3 (キュムラントの例).

- (1) 1 次のキュムラント $(\kappa_n)_{|n|=1}$ とは, 平均ベクトルである.
- (2) 2 次のキュムラント $(\kappa_n)_{|n|=2}$ とは, 分散共分散行列である.

さらに高次のキュムラントはテンソルになる 2.3.7. 正規分布ではここは消える. これが平均と分散までが大事であることに繋がるが, これは人類の認知負荷の限界と関係があるのか? □

系 1.8.4 (キュムラントの積率・中心積率による表現). 可積分性を仮定する.

- (1) $\kappa_1 = \alpha_1$.
- (2) $\kappa_2 = \alpha_2 - \alpha_1^2 = \mu_2$.
- (3) $\kappa_3 = \alpha_3 - 3\alpha_1\alpha_2 + 2\alpha_1^3 = \mu_3$.
- (4) $\kappa_4 = \alpha_4 - 3\alpha_2^2 - 4\alpha_1\alpha_3 + 12\alpha_1^2\alpha_2 - 6\alpha_1^4 = \mu_4 - 3\mu_2^2$.

系 1.8.5 (積率・中心積率のキュムラントによる表現). 可積分性を仮定する.

- (1) $\alpha_2 = \kappa_2 + \kappa_1^2, \mu_2 = \kappa_2$.
- (2) $\alpha_3 = \kappa_3 + 3\kappa_1\kappa_2 + \kappa_1^3, \mu_3 = \kappa_3$.
- (3) $\alpha_4 = \kappa_4 + 3\kappa_2^2 + 4\kappa_1\kappa_3 + 6\kappa_1^3\kappa_2 + \kappa_1^4, \mu_4 = \kappa_4 + 3\kappa_2^2$.

第2章

確率変数の従属性

2.1 確率変数の独立性と期待値

有限個の確率変数 X_1, \dots, X_n の独立性とは、これが定める (X_1, \dots, X_n) の結合分布が、各 X_i の分布 P^{X_i} が定める直積速度 $P^{X_1} \times \dots \times P^{X_n}$ に一致することをいう。本質的に「直積」とみなせるものを独立性という。

2.1.1 積の期待値による独立性の特徴付け

記法 2.1.1. $X_j : (\Omega, \mathcal{F}, P) \rightarrow (\mathcal{X}_j, \mathcal{G}_j)$ を確率変数とする。

命題 2.1.2 (独立確率変数の積は可積分). $n \geq 2$ を整数とする。可測関数 $f_j : \mathcal{X}_j \rightarrow \overline{\mathbb{R}}$ について、 $f_j(X_j) \in L^1(\Omega, \mathcal{F}, P)$ とする。このとき、 X_1, \dots, X_n が独立ならば、

- (1) 積 $\prod_{i=1}^n f_i(X_i)$ も可積分である。
- (2)

$$E \left[\prod_{i=1}^n f_i(X_i) \right] = \prod_{i=1}^n E[f_i(X_i)]$$

を満たす。

f_j が複素数値でも同様の結論が成り立つ。

[証明]. $X := f_1(X_1), Y := f_2(X_2) \geq 0$ の場合に証明する。 X, Y に下から収束する単関数の列 $X_n = \sum_{i \in \mathbb{N}} b_i 1_{B_i}, Y_n = \sum_{i \in \mathbb{N}} c_i 1_{C_i}$ が取れる。 B_i, C_i が独立であることに注意すると、

$$E[X_n Y_n] = E \left[\sum_{i,j \in \mathbb{N}} b_i c_j 1_{B_i \cap C_j} \right] = \sum_{i,j \in \mathbb{N}} b_i c_j P[B_i] P[C_j] = E[X_n] E[Y_n].$$

$n \rightarrow \infty$ の極限を考えると、 $E[XY] = E[X]E[Y]$. ■

系 2.1.3 (独立ならば無相関). $X, Y \in L^1(\Omega)$ を独立とする。このとき、 $\text{Cov}[X, Y] = 0$.

[証明]. 命題より、 $XY \in L^1(\Omega)$ 。よって、 $\text{Cov}[X, Y] < \infty$ 。これより、計算

$$\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])] = E[X - E[X]]E[Y - E[Y]] = 0$$

が正当化される。 ■

系 2.1.4 (独立性の十分条件). 任意の有界可測関数 f_i について

$$E \left[\prod_{i=1}^n f_i(X_i) \right] = \prod_{i=1}^n E[f_i(X_i)]$$

が成り立つならば、 X_1, \dots, X_n は独立である。

2.1.2 確率変数の独立性と特性関数

平均作用素 E の積に対する振る舞いによって、独立性の十分条件を与えることが出来た。特性関数によっても特徴付けることが出来る。

定理 2.1.5 (Kac). 確率変数列 $X_1, \dots, X_n \in L(\Omega)$ について、次の2条件は同値。

- (1) X_1, \dots, X_n は独立である。
- (2) $\varphi_X = \varphi_{X_1} \otimes \dots \otimes \varphi_{X_n} := \prod_{j=1}^n \varphi_{X_j}(u_j)$ である。

[証明].

(1) \Rightarrow (2) $X_1, (X_2, \dots, X_n)$ の独立性と命題 2.1.2 より帰納的に成り立つ。

(2) \Rightarrow (1) φ_X はある独立な確率変数列 $\tilde{X}_1, \dots, \tilde{X}_n$ であって、 $\varphi_{X_i} = \varphi_{\tilde{X}_i}$ を満たすものの特性関数に等しい。特性関数の一意性より、 X, \tilde{X} は同分布である。特に X_1, \dots, X_n は独立。

■

系 2.1.6. d 次元確率変数の列 X_1, \dots, X_n が独立ならば、

$$\varphi_{\sum_{j=1}^n X_j}(u) = \prod_{j=1}^n \varphi_{X_j}(u) \quad (u \in \mathbb{R}^d)$$

2.1.3 密度関数と独立性

命題 2.1.7. $X := (X_1, \dots, X_n)$ は密度関数 p を持つとする。このとき、次は同値：

- (1) X_1, \dots, X_n は独立。
- (2) $p = \otimes_{i \in [n]} p_i$ a.e.

[証明].

(2) \Rightarrow (1) 任意の $B_i \in \mathcal{B}(\mathbb{R})$ に対して、

$$\begin{aligned} P[X_1 \in B_1, \dots, X_n \in B_n] &= P^X \left[\prod_{i \in [n]} B_i \right] \\ &= \int_{\prod B_i} p dx_1 \cdots dx_n = \int_{\prod B_i} p_1 \cdots p_n dx_1 \cdots dx_n \\ &= \prod_{i \in [n]} P^{X_i}[B_i] = \prod_{i \in [n]} P[X_i \in B_i]. \end{aligned}$$

が成り立つから、これは独立性を意味する。

(1) \Rightarrow (2) 上述の等式が任意の閉矩形 $B = \prod_{i \in [n]} B_i$ 上で成り立つ。これは乗法族をなすから、この上で $\prod_{i \in [n]} p_i = p$ であることは、
 $p = \otimes_{i \in [n]} p_i$ a.s. を意味する。

■

2.2 多次元分布の従属性の扱い

2.2.1 同時・周辺密度関数

定義 2.2.1. $X \in L(\Omega; \mathbb{R}^d)$ について、

- (1) $F(x) := P[X_1 \leq x_1, \dots, X_d \leq x_d]$ を同時分布関数という.
- (2) $S \subset [d]$ について, $F_S(x) := \lim_{x_i \rightarrow \infty (i \notin S)} F(x)$ を周辺分布関数という.
- (3) 同時分布関数の適切な参照測度に関する Radon-Nikodym 微分 f を同時密度関数という.

補題 2.2.2. f の連続点 $x \in \mathbb{R}^d$ について,

$$f(x) = \frac{\partial^d F(x)}{\partial x_1 \partial x_2 \cdots \partial x_d}.$$

例 2.2.3. $X \in L(\Omega; \mathbb{R}^d)$ の密度を

$$f_{X,Y}(x, y) := \frac{1}{\pi^2} + \frac{1}{10} \cos x \cos y 1_{[0, \pi]^2}(x, y)$$

で与えると, 周辺分布はいずれも $U([0, \pi])$ である. □

2.2.2 条件付き密度関数

定義 2.2.4. (X, Y) の同時密度関数を f とする.

- (1) 条件付き密度関数を

$$f_1(x|y) := \frac{f(x, y)}{f_2(y)} = \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x, y) dx}$$

で定義する. $X \perp\!\!\!\perp Y$ のとき, $\forall y \in \mathbb{R} \ f_1(x|y) = f_1(x)$ である.

- (2) 条件付き期待値・分散を

$$E[X|Y=y] := \int_{\mathbb{R}} x f_1(x|y) dx, \quad \text{Var}[X|Y=y] := \int_{\mathbb{R}} (x - E[X|Y=y])^2 f_1(x|y) dx$$

で定める.

2.3 多次元分布の特性値

2.3.1 積率の定義

積率の次元もベクトル値 $n \in \mathbb{N}^d$ で指定できる.

記法 2.3.1. $x = (x_1, \dots, x_d) \in \mathbb{R}^d$ に対して, $\partial_i := \frac{\partial}{\partial x_i}$ とし, $n := (n_1, \dots, n_d) \in \mathbb{Z}_+^d$ に対して,

$$\begin{aligned} |n| &:= n_1 + \cdots + n_d, & n! &:= n_1! \cdots n_d!, \\ x^n &:= x_1^{n_1} \cdots x_d^{n_d}, & \partial^n &:= \partial_1^{n_1} \cdots \partial_d^{n_d}. \end{aligned}$$

ただし, $x_j^0 = 1, \partial^0 = 1$ とよむ.

定義 2.3.2 (moment, central moment). d 次元確率変数 $X = (X_1, \dots, X_d)$ に対して, 可積分性の仮定の下で,

- (1) $\alpha_n := E[X^n]$ を n 次積率という.
- (2) $\mu_n := E[(X - E[X])^n]$ を n 次中心積率という.

定義 2.3.3. $(\mathbb{R}^d, \mathcal{G}_d)$ 上の確率測度 ν に対して, 可積分性の仮定の下で,

- (1) $\mu := \left(\int_{\mathbb{R}^d} x_i \nu(dx) \right) \in \mathbb{R}^d$ を平均 (ベクトル) という.
- (2) $\alpha_n := \int_{\mathbb{R}^d} x^n \nu(dx)$ を n 次の積率という.
- (3) $\nu_n := \int_{\mathbb{R}^d} (x - \mu)^n \nu(dx)$ を n 次の中心積率という.
- (4) 特に $(\mu_n)_{|n|=2}$ を ν の分散共分散行列という.

2.3.2 相関係数

相関係数の言う「相関」とは、線型関係との近さであって、必ずしも一般的な意味での相関性 (association) を示すものではない。ここは自然言語的な誤解を極めて招きやすい。

定義 2.3.4. X, Y の基準化

$$Z_1 := \frac{X - E[X]}{\text{Var}[X]^{1/2}}, \quad Z_2 := \frac{Y - E[Y]}{\text{Var}[Y]^{1/2}}$$

の共分散

$$\rho(X, Y) := \text{Cov}[Z_1, Z_2] = \frac{\text{Cov}[X, Y]}{(\text{Var}[X]\text{Var}[Y])^{1/2}}$$

を相関係数という。

歴史 2.3.5. この流儀を Pearson の相関係数、または積率相関係数ともいう。

命題 2.3.6 (相関係数の値域). $\mu_X = E[X], \mu_Y = E[Y], \sigma_X = \sqrt{\text{Var}[X]}, \sigma_Y = \sqrt{\text{Var}[Y]}$ とする。

$$|\rho(X, Y)| \leq 1$$

等号成立条件は、 $Y = \pm \frac{\sigma_Y}{\sigma_X}(X - \mu_X) + \mu_Y$ a.s..

[証明]. $\sigma_X \sigma_Y \neq 0$ のとき、

$$0 \leq E[(\sigma_X^{-1}(X - \mu_X) \pm \sigma_Y^{-1}(Y - \mu_Y))^2] = 2(1 \pm \rho(X, Y))$$

より。 ■

2.3.3 キュムラントの定義

定義 2.3.7. 第2キュムラント母関数によるキュムラントの定義

$$\psi(u) = \sum_{j=1}^r \kappa_j \frac{(iu)^j}{j!} + o(u^r) \quad (|u| \rightarrow 0).$$

は、 n を多重指数 \mathbf{n} に読み替えることで、 $\kappa_{\mathbf{n}} := i^{-|\mathbf{n}|} \partial^{\mathbf{n}} \psi(0)$ そのまま拡張される。

記法 2.3.8. $(\partial_{u_i})_0$ を、 $u := (u_i) = 0$ における u_i -偏微分係数を表す。

議論 2.3.9 (キュムラントのテンソル表現). d 次元確率分布 ν について、 $\alpha_1, \dots, \alpha_r \in [d]$ に対して

$$\lambda^{\alpha_1 \cdots \alpha_r} := (-i)^r (\partial_{u_{\alpha_1}})_0 \cdots (\partial_{u_{\alpha_r}})_0 \log \varphi(u)$$

と定める。これも ν の r 次のキュムラントといい、行列 $(\lambda^{\alpha_1 \cdots \alpha_r})_{\alpha_1, \dots, \alpha_r \in [d]}$ は対称テンソルを定める。

要諦 2.3.10. $\mathbf{r} := (r_1, \dots, r_d) \in \mathbb{Z}_+^d$ が $r_k := |\{j \in [r] \mid \alpha_j = k\}|$ を満たすとする。このとき、

$$\lambda^{\alpha_1 \cdots \alpha_r} = \kappa_{\mathbf{r}}$$

2.4 コピュラ

相関係数は、2つの確率変数に対して、実数 $[-1, 1]$ を対応させるが、もっと表現力豊かに、変量間の従属性を、多次元分布によって直接的に表現することを考える。^a

これはなんだか、周辺分布の張り合わせとして多次元分布を表現するホモトピー論のようである。逆に、コピュラなく、直積によって周辺分布を張り合わせた場合が、独立性である。

^a この単語は元々音楽や言語学で使われていたが、統計学の用語として用いたのは、1959年にスクラー (Abe Sklar) がパリ大学統計学会誌 (the Statistical Institute of the University of Paris) で発表したのが最初である。

記法 2.4.1. $I := [0, 1]$ とする.

2.4.1 定義と例

定義 2.4.2 (copula). 次の条件を満たす関数 $C : I^2 \rightarrow I$ を接合関数という :

- (1) $C(-, 0) = C(0, -) = 0$.
- (2) $C(-, 1) = C(1, -) = \text{id}_I$.
- (3) $\forall u_1, u_2, v_1, v_2 \in I \quad u_1 \leq v_1, u_2 \leq v_2 \Rightarrow C(v_1, v_2) - C(u_1, v_2) - C(v_1, u_2) + C(u_1, u_2) \geq 0$.

例 2.4.3 (凸関数と product copula). C^2 級関数 C が $\partial_{u_1} \partial_{u_2} C(u_1, u_2) \geq 0$ をみたすとき, コピュラである. 特に, $C^\Pi(u_1, u_2) := u_1 u_2$ を積コピュラという. □

例 2.4.4 (二次元の分布関数が定めるコピュラ).

- (1) 区間 $(0, 1)$ 上の一様分布に従う2つの確率変数 $X_1, X_2 \sim U((0, 1))$ の分布関数 F はコピュラである.
- (2) □

2.4.2 Sklar の定理

多次元分布は, 周辺分布関数とコピュラによって定まる.

定理 2.4.5 (2次元の場合 (Sklar 1959)). (X_1, X_2) の結合分布関数を F , それぞれの周辺分布関数を F_1, F_2 とする. このとき, コピュラ C が存在して,

$$\forall x_1, x_2 \in \mathbb{R} \quad F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$$

と表せる. F_1, F_2 が連続のとき, C は一意である.

要諦 2.4.6. 逆関数 F_1^{-1}, F_2^{-1} が存在するときは, これを用いて $C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2))$ とコピュラを構成できる.

2.4.3 Archimedes コピュラ

第 3 章

分布の変換

3.1 確率密度関数の変換

確率密度関数とは一般化された 1-形式にほかならないから, $p^*(x)dx = p(y)dy$ という変換則を持つ. $y = T(x)$ のとき, $dy = |J_T(x)|dx$ より, Jacobian がかかる変換則 $p^*(x)dx = p(T(x))|J_T(x)|dx$ を持つ.

3.1.1 一般の変換則

命題 3.1.1 (変数変換による密度関数の変換則). $A, B \subset \mathbb{R}^d$ 上の可微分同相 $T : A \rightarrow B$ は Jacobian が A 上で消えないとする. $Y = T(X)$ が成り立つとき, Y の密度関数 p を用いて, X の密度関数 p^* は

$$p^*(x) = p(T(x))|J_T(x)| \quad (x \in A)$$

と表せる.

[証明]. 任意の有界な可測集合 $A_0 \subset A$ に対して, その上での積分は

$$\begin{aligned} P^X[A_0] &= P[X \in A_0] = \int_{A_0} p^*(x)dx \\ &= P[Y \in T(A_0)] = \int_{A_0} p(y)dy = \int_{A_0} p(T(x))|J_T(x)|dx. \end{aligned}$$

■

要諦 3.1.2 (変換の方向).

$$\begin{array}{ccc} x \in A & \xrightarrow{T} & B \ni y \\ \uparrow & & \nearrow p \\ \downarrow p^* & & \\ \Omega & & \end{array}$$

この状況において, p を T によって引き戻す際に $p^* = T^*p = p \circ T \cdot |J_T|$ が生じるのである.

命題 3.1.3. 可微分同相の組 $(T_i : A_i \rightarrow B_i)_{i \in [k]}$ と \mathbb{R}^d -値確率変数 X, Y について, 次を仮定する:

- (A1) $\{B_i\}_{i \in [k]} \subset \mathcal{O}(\mathbb{R}^d)$ は互いに素である.
- (A2) T_i は各 A_i 上 Jacobian が消えない.
- (A3) $A := \cup_{i \in [k]} A_i, B := \cup_{i \in [k]} B_i$ について, $P[X \in A] = P[Y \in B] = 1$.
- (A4) $X = \sum_{i \in [k]} T_i(Y) 1_{B_i}(Y)$ a.s.
- (A5) Y は絶対連続である. 密度を p とする.

このとき, X も絶対連続で, その密度は

$$p^*(x) = \sum_{i \in [k]} p(T_i(x))|J_T(x)|1_{A_i}(x) \quad (x \in \mathbb{R}^d).$$

例 3.1.4. 変換 $X = Y^3 - 3Y$ などの際も, $B_1 := (-\infty, -1), B_2 := (-1, 1), B_3 := (1, \infty)$ として対応できる.

□

3.1.2 一様分布への変換

命題 3.1.5. X は分布関数 F を持つ絶対連続分布に従うとする。このとき、 $F(X) \sim U([0, 1])$ 。

[証明]. F は連続で狭義単調増加，すなわち可逆とする。この場合は，

$$P[F(X) \leq a] = P[X \leq F^{-1}(a)] = F(F^{-1}(a)) = a.$$

より判る。 F が高々可算個の不連続点を持つ場合も同様。 ■

3.1.3 一様分布の正規分布への変換

命題 3.1.6 (Box-Müller transform). $Y_1, Y_2 \sim U(0, 1)$ を独立とする。

$$\begin{cases} X_1 = \sqrt{-2 \log Y_1} \cos(2\pi Y_2) \\ X_2 = \sqrt{-2 \log Y_2} \sin(2\pi Y_2) \end{cases}$$

と定める。

(1) 逆に解くと，次の関係を得る：

$$\begin{cases} y_1 = e^{-\frac{1}{2}(x_1^2 + x_2^2)} \\ y_2 = \frac{1}{2\pi} \tan^{-1} \left(\frac{x_2}{x_1} \right) \end{cases}$$

(2) $(X_1, X_2) \sim N(0, I_2)$ 。すなわち， X_1, X_2 は独立な標準正規確率変数となる。

[証明].

逆に解く

$$\begin{aligned} & \begin{cases} x_1 = \sqrt{-2 \log y_1} \cos(2\pi y_2) \\ x_2 = \sqrt{-2 \log y_2} \sin(2\pi y_2) \end{cases} \\ \Leftrightarrow & \begin{cases} x_1^2 + x_2^2 = -2 \log y_1 \\ \frac{x_2}{x_1} = \tan(2\pi y_2) \end{cases} \\ \Leftrightarrow & \begin{cases} y_1 = e^{-\frac{1}{2}(x_1^2 + x_2^2)} \\ y_2 = \frac{1}{2\pi} \tan^{-1} \left(\frac{x_2}{x_1} \right) \end{cases} \end{aligned}$$

密度の変換 こうして得た変換 $T: A \rightarrow B$ を $y = T(x)$ とおくと，

$$A := \mathbb{R}^2 \setminus \mathbb{R}_+ \times \{0\}, \quad B := (0, 1)^2$$

について $T(A) = B$ を満たす可微分同相で， $y = \tan^{-1} x$ の微分は

$$\frac{dx}{dy} = \frac{1}{\cos y} = 1 + \tan^2 y = 1 + x^2$$

であることに注意すれば，その勾配行列は

$$DT \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_1 e^{-\frac{1}{2}(x_1^2 + x_2^2)} & -x_2 e^{-\frac{1}{2}(x_1^2 + x_2^2)} \\ \frac{1}{2\pi} \frac{1}{1 + \left(\frac{x_2}{x_1}\right)^2} \left(-\frac{x_2}{x_1}\right) & \frac{1}{2\pi} \frac{1}{1 + \left(\frac{x_2}{x_1}\right)^2} \frac{1}{x_1} \end{pmatrix}$$

であるから，Jacobian は

$$J_T = e^{-\frac{1}{2}(x_1^2 + x_2^2)} \frac{1}{2\pi} \frac{1}{1 + \left(\frac{x_2}{x_1}\right)^2} \left(-1 - \frac{x_2^2}{x_1^2}\right) = -\frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2 + x_2^2)}.$$

よって， (X_1, X_2) の結合密度関数は，

$$f(x) = \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2 + x_2^2)} 1_A(x_1, x_2).$$

3.1.4 一様分布の指数分布への変換

命題 3.1.7. $T(x) := -\log x$ によって $T: (0, 1) \rightarrow (0, \infty)$ を定めると, $X \sim U(0, 1)$ に対して, $T(X) \sim \text{Exp}(1)$.

注 3.1.8. これは指数分布の, その分布関数による一様分布への変換の, 逆変換である.

3.2 商による変換

3.2.1 Gamma 分布の Beta 分布への変換

命題 3.2.1. $Y_i \sim \text{Gamma}(\alpha, \nu_i)$ ($i = 1, 2$) は独立とする. 確率変数

$$X_1 := \frac{Y_1}{Y_1 + Y_2}$$

の密度関数は

$$p^*(x_1) = \frac{1}{B(\nu_1, \nu_2)} x_1^{\nu_1-1} (1-x_1)^{\nu_2-1}, \quad (x_1 \in (0, 1)).$$

である. 特に $X_1 \sim \text{Beta}(\nu_1, \nu_2)$ である.

[証明].

逆に解く

$$\begin{cases} X_1 = \frac{Y_1}{Y_1 + Y_2} \\ X_2 = Y_2 \end{cases}$$

を逆に解くことで,

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} \frac{X_1 X_2}{1 - X_1} \\ X_2 \end{pmatrix} = T \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

を得る. 確かに, $A := (0, 1) \times (0, \infty)$, $B := (0, \infty)^2$ とすれば, $T: A \rightarrow B$ は可微分同相で, Jacobian は

$$DT = \begin{pmatrix} \frac{X_2}{(1-X_1)^2} & \frac{X_1}{1-X_1} \\ 0 & 1 \end{pmatrix}, \quad J_T = \frac{X_2}{(1-X_1)^2}$$

は消えない.

密度の変換 結合密度関数は

$$\begin{aligned} p^*(x_1, x_2) dx_1 dx_2 &= p(T(x_1, x_2)) J_T(x_1, x_2) dx_1 dx_2 \\ &= \frac{\alpha^{\nu_1}}{\Gamma(\alpha_1)} y_1^{\nu_1-1} e^{-\alpha y_1} \frac{\alpha^{\nu_2}}{\Gamma(\nu_2)} y_2^{\nu_2-1} e^{-\alpha y_2} \frac{x_2}{(1-x_1)^2} dx_1 dx_2 \\ &= \frac{\alpha^{\nu_1+\nu_2}}{\Gamma(\nu_1)\Gamma(\nu_2)} \left(\frac{x_1 x_2}{1-x_1} \right)^{\nu_1-1} e^{-\alpha \left(\frac{x_1 x_2}{1-x_1} \right)} x_2^{\nu_2-1} e^{-\alpha x_2} \frac{x_2}{(1-x_1)^2} dx_1 dx_2 \\ &= \frac{\alpha^{\nu_1+\nu_2}}{\Gamma(\nu_1)\Gamma(\nu_2)} \frac{x_1^{\nu_1-1} x_2^{\nu_1+\nu_2-1}}{(1-x_1)^{\nu_1+1}} e^{-\alpha \frac{x_2}{1-x_1}} dx_1 dx_2. \end{aligned}$$

よって, x_1 の周辺密度関数は, 積分により,

$$\begin{aligned} p^*(x_1) &= \frac{\alpha^{\nu_1+\nu_2}}{\Gamma(\nu_1)\Gamma(\nu_2)} \frac{x_1^{\nu_1-1}}{(1-x_1)^{\nu_1+1}} \int_0^\infty x_2^{\nu_1+\nu_2-1} e^{-\alpha \frac{x_2}{1-x_1}} dx_2 \\ &= \frac{\alpha^{\nu_1+\nu_2}}{\Gamma(\nu_1)\Gamma(\nu_2)} \frac{x_1^{\nu_1-1}}{(1-x_1)^{\nu_1+1}} \left(\frac{1-x_1}{\alpha} \right)^{\nu_1+\nu_2} \int_0^\infty \left(\frac{\alpha}{1-x_1} \right)^{\nu_1+\nu_2} x_2^{\nu_1+\nu_2-1} e^{-\frac{\alpha}{1-x_1} x_2} dx_2 \\ &= \frac{\Gamma(\nu_1+\nu_2)}{\Gamma(\nu_1)\Gamma(\nu_2)} x_1^{\nu_1-1} (1-x_1)^{\nu_2-1} = \frac{1}{B(\nu_1, \nu_2)} x_1^{\nu_1-1} (1-x_1)^{\nu_2-1}. \end{aligned}$$

3.2.2 正規分布の Cauchy 分布への変換

命題 3.2.2 . $Y_1, Y_2 \sim N(0, 1)$ を独立とする. 確率変数

$$X_1 := \frac{Y_1}{Y_2}$$

は $\text{Cauchy}(0, 1)$ に従う.

[証明].

変換の決定 変換

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_1 Y_2 \\ Y_2 \end{pmatrix} = T \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

を考えると,

$$DT(x_1, x_2) = \begin{pmatrix} x_2 & x_1 \\ 0 & 1 \end{pmatrix}, \quad J_T(x_1, x_2) = x_2$$

より, $\mathbb{R} \times \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R} \times \mathbb{R} \setminus \{0\}$ の Jacobian の消えない可微分同相を与える.

密度の計算 よって, 結合密度関数は

$$\begin{aligned} p^*(x_1, x_2) &= p(T(x_1, x_2)) |J_T(x_1, x_2)| \\ &= \frac{1}{2\pi} e^{-\frac{x_1^2 x_2^2}{2}} e^{-\frac{x_2^2}{2}} |x_2|. \end{aligned}$$

よって x_1 の周辺密度関数は

$$\begin{aligned} p^*(x_1) &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-\frac{x_2^2}{2}(1+x_1^2)} |x_2| dx_2 \\ &= \frac{1}{\pi} \int_0^\infty e^{-\frac{x_2^2}{2}(1+x_1^2)} x_2 dx_2 \\ &= \frac{1}{\pi} \frac{1}{1+x_1^2} \left[e^{-\frac{x_2^2}{2}(1+x_1^2)} \right]_0^\infty = \frac{1}{\pi} \frac{1}{1+x_1^2}. \end{aligned}$$

■

系 3.2.3 . 正定値対称行列 $\Sigma = (\sigma_{ij}) \in M_2(\mathbb{R})$ について, $(Y_1, Y_2) \sim N(0, \Sigma)$ とする. このとき,

$$X_1 := \frac{Y_1}{Y_2} \sim \text{Cauchy} \left(\frac{\sigma_{12}}{\sigma_{22}}, \frac{\sqrt{\det \Sigma}}{\sigma_{22}} \right).$$

3.3 自乗による変換

3.3.1 一般論

系 3.3.1 . $Y \in L(\Omega)$ が密度 p を持ち, $X := Y^2$ とする.

(1) $T_1: \mathbb{R}^+ \rightarrow \mathbb{R}^+, T_2: \mathbb{R}^+ \rightarrow \mathbb{R}^+$ を $T_1(x) := -\sqrt{x}, T_2(x) := \sqrt{x}$ で定めると, 命題 3.1.3 の条件を満たす.

(2) X の密度は次で表される:

$$p^*(x) = \frac{1}{2\sqrt{x}} \left(p(\sqrt{x}) + p(-\sqrt{x}) \right) 1_{\mathbb{R}^+}(x), \quad (x \in \mathbb{R}).$$

(3) 特に Y が対称であるとき, X の密度は次で表される:

$$f_X(x) = \frac{p(\sqrt{x})}{\sqrt{x}} 1_{\mathbb{R}^+}(x).$$

3.3.2 正規分布の二乗はカイ二乗分布

命題 3.3.2. $Z \sim N(0, 1)$ のとき, $Z^2 \sim \chi^2(1)$.

[証明]. 系より,

$$f_X(x) = \frac{\phi(\sqrt{x})}{\sqrt{x}} 1_{\mathbb{R}^+}(x) = \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{x}{2}}}{\sqrt{x}} = \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}}.$$

3.4 和による変換

3.4.1 一般論

定理 3.4.1. $X_1, X_2 \in L(\Omega; \mathbb{R}^d)$ は独立で, 密度関数 f_1, f_2 を持つとする. このとき, $X_1 + X_2$ について,

- (1) 絶対連続である.
- (2) 密度関数は

$$p(x) = \int_{\mathbb{R}^d} f_1(x - y) f_2(y) dy.$$

3.4.2 二項分布は Bernoulli 分布の畳み込み

命題 3.4.2.

- (1) $\text{Bin}(n_1, p) * \text{Bin}(n_2, p) = \text{Bin}(n_1 + n_2, p)$.
- (2) 二項分布 $\text{Bin}(n, p)$ は $\text{Ber}(p)$ の n 重の畳み込みである.

[証明].

- (1) $\text{Bin}(n, p)$ の確率母関数は

$$\begin{aligned} g(z) &= \sum_{k=0}^{\infty} p_k z^k \\ &= \sum_{k=0}^n \binom{n}{k} (pz)^k (1-p)^{n-k} = (pz + q)^n. \end{aligned}$$

よって, $\text{Bin}(n_1, p) * \text{Bin}(n_2, p) = \text{Bin}(n_1 + n_2, p)$.

- (2) $\text{Ber}(p) = \text{Bin}(1, p)$ であるから, 特別な場合である.

3.4.3 Poisson 分布の再生性

命題 3.4.3.

- (1) $\text{Pois}(\lambda_1) * \text{Pois}(\lambda_2) = \text{Pois}(\lambda_1 + \lambda_2)$.
- (2) $\text{Hermite}(\alpha_1, \beta_1) * \text{Hermite}(\alpha_2, \beta_2) = \text{Hermite}(\alpha_1 + \alpha_2, \beta_1 + \beta_2)$.
- (3) $\text{Skellam}(\alpha_1, \beta_1) * \text{Skellam}(\alpha_2, \beta_2) = \text{Skellam}(\alpha_1 + \alpha_2, \beta_1 + \beta_2)$.

[証明].

- (1) Poisson 分布の確率母関数は

$$g(z) = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} z^n = e^{-\lambda} e^{\lambda z} = e^{\lambda(z-1)}.$$

であるから, $\text{Pois}(\lambda_1) * \text{Pois}(\lambda_2)$ の確率母関数は $e^{(\lambda_1 + \lambda_2)(z-1)}$.

3.4.4 指数分布の独立和

命題 3.4.4.

- (1) $X_1, \dots, X_n \sim \text{Exp}(\lambda) = G(\lambda, 1)$ が独立同分布ならば, $X_1 + \dots + X_n \sim G(\lambda, n)$.
- (2) $X_i - X_j \sim \text{Lap}(0, \lambda^{-1})$.

[証明].

- (1) 指数分布 $\text{Exp}(\lambda) = G(\lambda, 1)$ の特性関数 $\frac{1}{1 - \frac{iu}{\lambda}}$ の n 乗は $G(\lambda, n)$ の特性関数であるから,

命題 3.4.5. $N \sim \text{Pois}(\lambda)$ として, $\text{Exp}(\mu)$ ($\mu > 0$) の独立同分布列の N 項和

$$Y := X_1 + \dots + X_N$$

について,

- (1) 特性関数は $\varphi_Y(u) = \exp\left(\frac{iu\lambda}{\mu - iu}\right)$.
- (2) 平均は $\alpha_1 = \frac{\lambda}{\mu}$.
- (3) 分散は $\mu_2 = \frac{2\lambda}{\mu^2}$.

[証明].

- (1) 繰り返し期待値の法則より,

$$\begin{aligned} E[e^{iuY}] &= E[E[e^{iu(X_1 + \dots + X_N)} | N = n]] = E\left[\left(\frac{\mu}{\mu - iu}\right)^N\right] \\ &= \sum_{n=0}^{\infty} \left(\frac{\mu}{\mu - iu}\right)^n \frac{\lambda^n}{n!} e^{-\lambda} = e^{\frac{\mu\lambda}{\mu - iu}} e^{-\lambda}. \end{aligned}$$

- (2) 積率母関数は $M(t) = e^{\frac{\lambda t}{\mu - t}}$ となる. この微分は

$$\begin{aligned} \frac{dM}{dt} &= \left(\frac{\lambda}{\mu - t} + \frac{\lambda t}{(\mu - t)^2}\right) e^{\frac{\lambda t}{\mu - t}}, \\ \frac{d^2 M}{dt^2} &= \left(\frac{2\lambda}{(\mu - t)^2} + \frac{2\lambda t}{(\mu - t)^3}\right) e^{\frac{\lambda t}{\mu - t}} + \left(\frac{\lambda}{\mu - t} + \frac{\lambda t}{(\mu - t)^2}\right)^2 e^{\frac{\lambda t}{\mu - t}}. \end{aligned}$$

よって, $\alpha_1 = M'(0) = \frac{\lambda}{\mu}$.

- (3) $\alpha_2 = M''(0) = \frac{2\lambda}{\mu^2} - \left(\frac{\lambda}{\mu}\right)^2$. $\mu_2 = \alpha_2 - \alpha_1^2 = \frac{2\lambda}{\mu^2}$.

3.4.5 Gamma 分布の再生性

命題 3.4.6. $\text{Gamma}(\alpha, \nu_1) * \text{Gamma}(\alpha, \nu_2) = \text{Gamma}(\alpha, \nu_1 + \nu_2)$.

3.4.6 総和核の再生性

命題 3.4.7.

- (1) $\text{Cauchy}(\mu_1, \sigma_1) * \text{Cauchy}(\mu_2, \sigma_2) = \text{Cauchy}(\mu_1 + \mu_2, \sigma_1 + \sigma_2)$.
- (2) $N(\mu_1, \sigma_1^2) * N(\mu_2, \sigma_2^2) = N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.
- (3) $\text{IG}(\delta_1, \gamma) * \text{IG}(\delta_2, \gamma) = \text{IG}(\delta_1 + \delta_2, \gamma)$.

3.5 順序統計量

3.5.1 順序統計量の密度関数

順序統計量の考え方は、「ボールと区切り」の議論の一般形である。

定理 3.5.1. X_1, \dots, X_n を独立同分布列, その分布関数を F とすると, r 番目の順序統計量 $X_{n:r}$ の分布関数 F_r は次のように表せる:

$$F_r(x) = \sum_{i=r}^n \binom{n}{i} F(x)^i (1 - F(x))^{n-i}.$$

[証明]. $X_{n:n+1} = \infty$ と定めると, 事象 $\{X_{n:r} \leq x\} = \{x \in [X_{n:r}, \infty)\}$ を, どの区間 $[X_{n:i}, X_{n:i+1}]$ ($i = r, \dots, n$) に入るかで場合分けして考えることより,

$$F_r(x) = P[X_{n:r} \leq x] = \sum_{i=r}^n P[X_{n:i} \leq x < X_{n:i+1}]$$

であるが,

$$\{X_{n:i} \leq x < X_{n:i+1}\} = \{x \text{ 以下の観測が } i \text{ 個}\}$$

に注意すれば, 右辺は $Y := 1_{\{X \leq x\}} \sim B(n, p)$ に関する事象と見て, その確率を計算出来る:

$$P[X_{n:i} \leq x < X_{n:i+1}] = \binom{n}{i} F(x)^i (1 - F(x))^{n-i}.$$

■

系 3.5.2. X_1 の分布は絶対連続とする. このとき, $X_{n:r}$ も絶対連続で,

$$f_r(x) = r \binom{n}{r} F(x)^{r-1} (1 - F(x))^{n-r} f(x).$$

特に, $\max_{i \in [n]} X_i = X_{n:n}, \min_{i \in [n]} X_i = X_{n:1}$ の密度関数は, それぞれ

$$f_n(x) = nF(x)^{n-1}f(x), \quad f_1(x) = n(1 - F(x))^{n-1}f(x).$$

[証明]. F_r の微分による. ■

要諦 3.5.3. より本質的な議論は次のようになる. \mathbb{R} を $(-\infty, x), (x, x + dx), (x + dx, \infty)$ に分割し, X がどこに入るかについて 3 項分布 $\text{Mult}_3(F(x), f(x)dx, 1 - F(x + dx))$ に従う. これに対して n が $(r - 1, 1, n - r)$ にわかれれば良いので, その確率密度関数は

$$f_r(x)dx = \binom{n}{r-1, 1, n-r} F(x)^{r-1} f(x)dx (1 - F(x))^{n-r}.$$

3.5.2 結合密度関数

命題 3.5.4. X_1, X_2, \dots, X_n を独立同分布, その密度を $f(x)$ とする.

(1) $(X_{n:1}, \dots, X_{n:n})$ の結合密度関数は,

$$f(x_1, \dots, x_n) = \begin{cases} n!P[X_1 = x_1] \cdots P[X_n = x_n] & x_1 \leq x_2 \leq \cdots \leq x_n, \\ 0 & \text{otherwise} \end{cases} = n!f(x_1) \cdots f(x_n) 1_{\{x_1 \leq x_2 \leq \cdots \leq x_n\}}.$$

(2) $i < j$ について, $(X_{n:i}, X_{n:j})$ の結合密度関数は,

$$f_{i,j}(u, v) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} F(u)^{i-1} (F(v) - F(u))^{j-i-1} (1 - F(v))^{n-j} f(u) f(v) 1_{\{u < v\}}.$$

3.5.3 一様分布の順序統計量は Beta 分布に従う

系 3.5.5 (一様分布の順序統計量の密度). $X_1, \dots, X_n \sim U([0, 1])$ を独立同分布とする. r 番目の順序統計量 $X_{n:r}$ について, $X_{n:r} \sim \text{Beta}(r, n - r + 1)$.

[証明]. 定理 3.5.1 より,

$$f_r(x) = r \binom{n}{r} x^{r-1} (1-x)^{n-r} 1 = \frac{n!}{(r-1)!(n-r)!} x^{r-1} (1-x)^{n-r}.$$

特に階乗部分を $\frac{\Gamma(n+1)}{\Gamma(r)\Gamma(n-r+1)}$ とみれば, これは $\text{Beta}(r, n - r + 1)$ の確率密度関数である. ■

例 3.5.6 (一様分布族の完備十分統計量). このことから, 一様分布の族 $(U(0, \theta)^{\otimes n})_{\theta > 0}$ について, 最大値を表す順序統計量 $X_{n:n}$ が完備十分統計量を与えることが判る. □

[証明].

十分性 $\tilde{1}_A := \sum_{\sigma \in S_n} \frac{1_A(\sigma(x))}{n!}$ ($A \in \mathcal{B}(\mathbb{R}^n)$) について,

$$q(A|t) := \int_{\mathbb{R}^{n-1}} \tilde{1}_A(x_1, \dots, x_{n-1}, t) \frac{1}{t^{n-1}} \sum_{j=1}^{n-1} 1_{(0,t)}(x_j) dx_1 \cdots dx_{n-1}, \quad (A \in \mathcal{B}(\mathbb{R}^n))$$

と定めると, これが T を与えた下での x の条件付き確率分布を与えることを示せば良い. $U(0, \theta)^{\otimes n}$ の密度は

$$P_\theta(dx) = \frac{1}{\theta^n} \prod_{j \in [n]} 1_{(0,\theta)}(x_j) dx, \quad (x \in \mathbb{R}^n)$$

であるから, 任意の $A \in \mathcal{B}(\mathbb{R}^n), B \in \mathcal{B}(\mathbb{R})$ に対して,

$$\begin{aligned} P_\theta[A \cap T^{-1}(B)] &= \int_{\mathbb{R}^n} \tilde{1}_A(x) 1_B(T(x)) \frac{1}{\theta^n} \prod_{j \in [n]} 1_{(0,\theta)}(x_j) dx \\ &= n \int_{\mathbb{R}^n} \tilde{1}_A(x) 1_B(x_n) \frac{1}{\theta^n} \prod_{j=1}^{n-1} 1_{(0,x_n)}(x_j) 1_{(0,\theta)}(x_n) dx \\ &= \int_B q(A|x_n) P_\theta^T(dx_n). \end{aligned}$$

完備性 任意の可測関数 $h: \mathbb{R}_+ \rightarrow \mathbb{R}$ について, $E_\theta^T[h] = 0 \Leftrightarrow E_\theta[h(T)] = 0$ とする. 一様分布 $U(0, 1)$ からの標本の順序統計量 T は $\text{Beta}(n, 1)$ に従うから, 今回の T の密度は

$$p_\theta(t) = \frac{1}{\theta^n} n t^{n-1} 1_{(0,\theta)}(t)$$

であることに注意すれば,

$$\int_0^\theta h^+(t) t^{n-1} dt = \int_0^\theta h^-(t) t^{n-1} dt \quad (\theta > 0)$$

が必要なが判る. これは $(0, \infty)$ 上殆ど至る所 $h^+ = h^-$ であることを要請しており, \mathbb{R}_+ 上殆ど至る所 $h = 0$ であることが解った. ■

3.5.4 指数分布の場合

命題 3.5.7. X_1, \dots, X_n をそれぞれ母数 $\lambda_1, \dots, \lambda_n$ を持つ指数分布に従う独立確率変数とする.

- (1) $X_{(1)} = \min \{X_1, \dots, X_n\} \sim \text{Exp}(\lambda_1 + \dots + \lambda_n)$.
 (2) 簡単のため同分布の仮定 $\lambda = \lambda_1 = \dots = \lambda_n$ をおく. このとき,

$$X_{(n)} \stackrel{d}{=} \max_{i \in [n]} X_i = X_1 + \frac{1}{2}X_2 + \dots + \frac{1}{n}X_n.$$

特に, もはや指数分布には従わない.

注 3.5.8 (パラメータ変換 T は本当に可微分同相に取れて居るか?). 次の議論の問題点を指摘できるか?

[誤った議論]. 積分を通じて, $(E_{(1)}, E_{(2)})$ の周辺密度関数は, 結合密度関数 $f(x, y, z) = 6e^{-(x+y+z)}1_{\{x \leq y \leq z\}}$ を積分することより,

$$f_{1,2}(x, y) = \int_{z=y}^{z=\infty} 6e^{-(x+y+z)}1_{\{x \leq y\}} dz = \left[-6e^{-(x+y+z)}1_{\{x \leq y\}} \right]_{z=y}^{z=\infty} = 6e^{-(x+2y)}1_{\{x \leq y\}}.$$

$2(E_{(2)} - E_{(1)})$ の密度は, $Y_1 := 2(E_{(2)} - E_{(1)}), Y_2 := E_{(2)}$ と定めると, 逆に解くことで

$$\begin{cases} E_{(1)} = Y_2 - \frac{Y_1}{2} \\ E_{(2)} = Y_2 \end{cases}$$

であるから, この線型変換の Jacobian は $-\frac{1}{2}$ で, 結合密度は

$$p(y_1, y_2) = 6e^{-(3y_2 - y_1/2)} \frac{1}{2}.$$

y_2 について積分することで, $2(E_{(2)} - E_{(1)})$ の密度は,

$$\int_0^\infty p(y_1, y_2) dy_2 = \int_0^\infty 3e^{-3y_2 + \frac{y_1}{2}} dy_2 = \left[-e^{-3y_2 + \frac{y_1}{2}} \right]_0^\infty = e^{y_1/2}.$$

命題 3.5.9. $X_1, \dots, X_n \sim \text{Exp}(\lambda)$ を独立とする. $X_{n0} := 0$ とし,

$$Y_j := (n - j + 1)(X_{nj} - X_{n,j-1}), \quad j \in [n]$$

とすると, Y_1, \dots, Y_n は再び独立に指数分布に従う.

要諦 3.5.10 (Alfred Renyi (53)). $X_1, \dots, X_n \sim \text{Exp}(\lambda)$ を独立同分布とする. このとき, 独立同分布 $Z_1, Z_2, \dots \sim \text{Exp}(1)$ について,

$$X_{(i)} \stackrel{d}{=} \frac{1}{\lambda} \sum_{j=1}^i \frac{Z_j}{n - j + 1}.$$

3.5.5 順序統計量の関数である統計量

定義 3.5.11. (1) 中央値とは,

$$X_{\text{med}} := \begin{cases} X_{n:(n+1)/2} & n \text{ が奇数のとき,} \\ \frac{1}{2}(X_{n:n/2} + X_{n:(n/2+1)}) & n \text{ が偶数のとき.} \end{cases}$$

(2) 範囲の中央とは,

$$T := \frac{X_{n:1} + X_{n:n}}{2}.$$

(3) 標本範囲とは,

$$R := X_{n:n} - X_{n:1}.$$

命題 3.5.12. R, T の結合密度は

$$f_{R,T}(r, t) = n(n-1) \left(F\left(t + \frac{r}{2}\right) - F\left(t - \frac{r}{2}\right) \right)^{n-2} f\left(t - \frac{r}{2}\right) f\left(t + \frac{r}{2}\right) 1_{R^+ \times \mathbb{R}}(r, t).$$

3.6 線型変換と正規標本統計量

Gauss 系の線型変換は Gauss 系である。Gauss 確率変数が直交することは独立性に同値である。

3.6.1 標本平均と標本分散と不偏分散

定義 3.6.1. $X_1, \dots, X_n \sim (\mu, \sigma^2)$ を独立同分布とする。

(1) 次の統計量を**標本平均**という：

$$\bar{X} := \frac{1}{n} \sum_{i \in [n]} X_i = \mathbb{P}_n[X].$$

(2) 平均母数 $\mu = \mu_0$ が所与のとき、

$$S^2(\mu_0) := \frac{1}{n} \sum_{i=1}^n (X_i - \mu_0)^2$$

を μ_0 周りの**標本分散**という。

(3) 次の統計量を**標本分散**という：

$$S^2 := \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

標本分散は標本平均の合成関数 $S^2 = S^2(\bar{X})$ である点に注意。

(4) 次の統計量を**不偏分散**という：

$$U^2 := \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

命題 3.6.2.

(1) $\bar{X} \sim (\mu, \sigma^2/n)$.

(2) 母平均が μ_0 に等しいとき、 $S^2(\mu_0) \sim \left(\sigma^2, \frac{\mu_4 - \sigma^4}{n}\right)$ 。ただし、 μ_4 は 4 次の中心積率とした。

(3) $S^2 \sim \left(\left(1 - \frac{1}{n}\right)\sigma^2, \left(1 - \frac{1}{n}\right)^2 \text{Var}[U^2]\right)$ 。

(4) $U^2 \sim \left(\sigma^2, \frac{1}{n} \left(\mu_4 - \frac{n-3}{n-1}\sigma^4\right)\right)$ 。

[証明].

(1)

$$E[\bar{X}] = \frac{1}{n} \sum_{i \in [n]} E[X_i] = \frac{n\mu}{n} = \mu.$$

$$\begin{aligned} \text{Var}[\bar{X}] &= E[(\bar{X} - \mu)^2] = E\left[\left(\frac{1}{n} \sum_{i \in [n]} X_i - \mu\right)^2\right] \\ &= E\left[\frac{1}{n^2} \left(\sum_{i \in [n]} (X_i - \mu)\right)^2\right] \\ &= \frac{1}{n^2} E\left[\sum_{i \in [n]} (X_i - \mu)^2\right] + \frac{1}{n} \sum_{i \neq j} E[(X_i - \mu)(X_j - \mu)] \\ &= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}. \end{aligned}$$

(2)

$$E[S^2(\mu_0)] = E\left[\frac{1}{n}\sum_{i=1}^n(X_i - \mu_0)^2\right] = \sigma^2$$

$$\begin{aligned}\text{Var}[S^2(\mu_0)] &= E\left[\left(\frac{1}{n}\sum_{i=1}^n(X_i - \mu_0)^2 - \sigma^2\right)^2\right] \\ &= \frac{1}{n^2}E\left[\left(\sum_{i \in [n]}(X_i - \mu_0)^2 - n\sigma^2\right)^2\right] \\ &= \frac{1}{n^2}E\left[\sum_{i \in [n]}(X_i - \mu_0)^4\right] - \frac{2\sigma^2}{n^2}E\left[\sum_{i \in [n]}(X_i - \mu_0)^2\right] + \frac{\sigma^4}{n} \\ &= \frac{\mu_4}{n} - \frac{\sigma^4}{n}.\end{aligned}$$

(3)

$$\begin{aligned}\sum_{i \in [n]}(X_i - \bar{X})^2 &= \sum_{i \in [n]}(X_i - \mu + \mu - \bar{X})^2 \\ &= \sum_{i \in [n]}(X_i - \mu)^2 + n(\bar{X} - \mu)^2 + 2(\mu - \bar{X})\sum_{i \in [n]}(X_i - \mu) \\ &= \sum_{i \in [n]}(X_i - \mu)^2 + n(\bar{X} - \mu)^2 + 2(\mu - \bar{X})n(\bar{X} - \mu) \\ &= \sum_{i \in [n]}(X_i - \mu)^2 - n(\bar{X} - \mu)^2.\end{aligned}$$

これを両辺 n で割ると,

$$S^2 = S^2(\mu) - (\bar{X} - \mu)^2$$

であるから, (1) と (2) より,

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

分散は, 関係 $S^2 = \frac{n-1}{n}U^2$ による.

(4) (3) と同様の計算により,

$$E[U^2] = \frac{n}{n-1}E[S^2(\mu)] - \frac{n}{n-1}\text{Var}[\bar{X}] = \frac{n}{n-1}\sigma^2 - \frac{n}{n-1}\frac{\sigma^2}{n} = \sigma^2.$$

分散は, $\mu = 0$ と仮定すると少し楽になるが, ものすごい計算量になる.

■

要諦 3.6.3. U^2 は母平均の不偏推定量であるが, S^2 の方が分散が小さいことに注目.

3.6.2 正規標本の場合

命題 3.6.4. $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ を独立とする.

$$(1) \frac{nS^2(\mu)}{\sigma^2} \sim \chi^2(n).$$

(2) この確率変数は,

$$\frac{nS^2(\mu)}{\sigma^2} = \frac{n-1}{\sigma^2}U^2 + \frac{n(\bar{X} - \mu)^2}{\sigma^2}$$

と独立な2つの確率変数の和に分解され, 右辺第1項は $\chi^2(n-1)$, 第2項は $\chi^2(1)$ に従う.[証明].

(1)

$$\frac{nS^2(\mu)}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \mu)^2}{\sigma^2}$$

は, $\frac{X_i - \mu}{\sigma} \sim N(0, 1)$ の二乗の独立和であるため.

(2)

$$\begin{aligned} \sum_{i \in [n]} (X_i - \mu)^2 &= \sum_{i \in [n]} (X_i - \bar{X} + \bar{X} - \mu)^2 \\ &= \sum_{i \in [n]} (X_i - \bar{X})^2 + 2(\bar{X} - \mu) \underbrace{\sum_{i \in [n]} (X_i - \bar{X})}_{=0} + n(\bar{X} - \mu)^2 \\ &= \sum_{i \in [n]} (X_i - \bar{X})^2 + n(\bar{X} - \mu)^2. \end{aligned}$$

この両辺を σ^2 で割ると表式を得る. 独立性 $S^2 \perp \bar{X}$ は次から従う.

命題 3.6.5. $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ を独立同分布列, $x := (X_1, \dots, X_n)^\top$ とする.

(1) $B \in M_{mn}(\mathbb{R}), A \in M_n(\mathbb{R})$ を対称行列とする. $BA = O$ のとき, 2つの確率変数 Bx と $x^\top Ax$ とは独立になることを示せ.

(2) 標本平均 $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ と標本分散 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ が独立であることを示せ.

[証明].

(1) A は対称行列だから, ある直交行列 $U \in O_n(\mathbb{R}), U^\top U = I_n$ を用いて, $U^\top D U = A$ と対角化出来る. よって, $y := Ux \in \mathbb{R}^n$ と定めると, これは再び独立な正規確率変数のベクトル $y \sim N(\mu U 1_n, \sigma^2 I_n)$ で,

$$x^\top Ax = (Ux)^\top D(Ux) = a_1 y_1^2 + \dots + a_r y_r^2, \quad r := \text{rank } A, D =: \text{diag}(a_1, \dots, a_r)$$

と表せる. 次に, $BA = 0$ より, $\text{Im } A \subset \text{Ker } B$ すなわち $\text{Im } B \subset \text{Ker } A$ であるから, Bx は y_{r+1}, \dots, y_n のみによって表せる確率変数のベクトルである (使わないものもあるかもしれないが). よって, Bx と $x^\top Ax$ は独立.

(2) $m = 1, B := \frac{1}{n} 1_n^\top$ と

$$A := \frac{1}{n-1} \begin{pmatrix} 1 - \frac{1}{n} & -\frac{1}{n} & \dots & -\frac{1}{n} \\ -\frac{1}{n} & 1 - \frac{1}{n} & \dots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \dots & 1 - \frac{1}{n} \end{pmatrix} =: \frac{1}{n-1} \Lambda^2$$

と定めると, $B\Lambda = O$ より, $BA = O$ であり, 同時に $\bar{X} = Bx$ かつ $S^2 = x^\top Ax$ である.

要諦 3.6.6. (1) の事実は $B := \frac{bb^\top}{\|b\|^2}$ を射影行列とし,

$$X^\top AX \perp b^\top X \Leftrightarrow X^\top AX \perp X^\top BX$$

に注意すれば, Fisher-Cochran の定理??からすぐに従う. また, Basu の定理からも従う??. (2) の独立な標本の標本平均と標本分散が独立であるという事実は, 標本が正規分布に従っているという事実を特徴付ける [Kawata and Sakamoto, 1949].

3.6.3 線形変換と独立性

定理 3.6.7 (Levy (57)). $X, Y \in L(\Omega)$ は, ある $U \perp X, V \perp Y$ を満たす $U, V \in L(\Omega)$ について, 独立確率変数の線型和

$$\begin{cases} Y = aX + U, \\ X = bY + V \end{cases}$$

と表されているとする。このとき、次の3つのいずれかが成り立つ：

- (1) $\{X, Y\}$ は Gauss 系である。
- (2) $X \perp\!\!\!\perp Y$.
- (3) X, Y 間に線型関係がある。

3.6.4 線形変換と収束

定理 3.6.8 (Cramer-Wold). $X_n, X \in L(\Omega; \mathbb{R}^d)$ について、次は同値：

- (1) $X_n \xrightarrow{d} X$.
- (2) 任意の $a \in \mathbb{R}^d$ について、 $a^\top X_n \xrightarrow{d} a^\top X$.

3.7 混合分布

第 4 章

確率分布の例

純不連続分布を離散分布ともいう。離散分布というクラスは、極限構成について閉じていないという意味で、理論的な実用性がない。極限について議論するには、絶対連続分布と実数というところに行き着く。

4.1 一次元純不連続分布の例

4.1.1 離散一様分布

定義 4.1.1 (discrete uniform distribution). 有限集合 $[n]$ ($n \in \mathbb{N}^+$) 上の定値関数 $p = \frac{1}{n}$ を確率関数として定まる分布 $U: \mathbb{N}^+ \rightarrow P(\mathbb{N})$ を離散一様分布という。

命題 4.1.2 (離散一様分布の特性値).

- (1) 平均: $\alpha_1(U(n)) = \frac{n+1}{2}$.
- (2) 分散: $\mu_2(U(n)) = \frac{n^2-1}{12}$.
- (3) 歪度と尖度: $\gamma_1 = 0, \gamma_2 = \frac{3(3n^2-7)}{5(n^2-1)}$.
- (4) 確率母関数: $g(z) = \frac{z(1-z^n)}{n(1-z)}$.

[証明].

(1)

$$E[X] = \sum_{x=1}^n x \frac{1}{n} = \frac{1}{n} \times \frac{1}{2} n(n+1) = \frac{n+1}{2}.$$

(2)

$$E[X^2] = \sum_{x=1}^n x^2 \frac{1}{n} = \frac{1}{6} (n+1)(2n+1).$$

より,

$$\text{Var}[X] = E[X^2] - E[X]^2 = \frac{2n^2+3n+1}{6} - \frac{n^2+2n+1}{4} = \frac{n^2-1}{12}.$$

(3) 引き続き, 和の公式より 4 次の積率までは求まる.

(4) 等比級数の和の公式より,

$$\begin{aligned} g(z) &= E[z^X] = \frac{1}{n} (t + t^2 + \cdots + t^n) \\ &= \frac{1}{n} \frac{z(1-z^n)}{1-z}. \end{aligned}$$

4.1.2 二項分布

$B(n, p)$ とは成功確率 p の試行を独立に n 回繰り返した際の成功回数の分布である。二項定理より、確率母関数は $g(z) = (pz + (1 - p))^n$ となる。

定義 4.1.3 (binomial distribution). 可測空間 $(n + 1, P(n + 1))$ において、

(1) 二項分布 $B(n, p) : \mathbb{N} \times [0, 1] \rightarrow P(\mathbb{N})$ は、次の確率質量関数で定義される確率分布である：

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x} \quad x \in n + 1.$$

成功確率が p で一定な試行（Bernoulli 試行という）を独立に n 回続けた際の成功回数に従う確率分布である。

(2) $\text{Ber}(p) := B(1, p)$ は $\{0, 1\}$ 上の確率分布であり、**Bernoulli 分布** という。

命題 4.1.4 (二項分布の平均と分散).

- (1) 平均： $\alpha_1(b(n, p)) = np$.
- (2) 確率母関数： $g(z) = (pz + q)^n$.
- (3) 分散： $\mu_2(b(n, p)) = npq$.

[証明].

(1)

$$\begin{aligned} \mu &= \sum_{x=0}^n x \binom{n}{x} p^x q^{n-x} \\ &= \sum_{x=0}^n x \frac{n!}{x!(n-x)!} p^x q^{n-x} \\ &= np \sum_{x=1}^n \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x} \\ &= np \sum_{y=0}^{n-1} \frac{(n-1)!}{y!(n-1-y)!} p^y q^{n-1-y} \quad y := x - 1 \\ &= np \sum_{y=0}^{n-1} b(n-1, p) = np. \end{aligned}$$

(2) 確率関数 $(b(x; n, p))_{x \in n+1}$ が定める母関数 $g : \mathbb{R} \rightarrow \mathbb{R}$ は

$$\begin{aligned} g(z) &= \sum_{x=0}^n b(x; n, p) z^x \\ &= \sum_{x=0}^n \binom{n}{x} p^x q^{n-x} z^x = (pz + q)^n \end{aligned}$$

と表せる。

(3) $g'(z) = np(pz + q)^{n-1}$, $g''(z) = n(n-1)p^2(pz + q)^{n-2}$ より、分散公式 1.6.4 から、

$$\sigma^2 = g''(1) + g'(1) - g'(1)^2 = n(n-1)p^2 + np - n^2p^2 = npq.$$

■

要諦 4.1.5. 確率母関数が、二項展開の式に一致することから、ここに予想だにしなかった抜け道がある。

4.1.3 Poisson 分布

指数関数 e^z の Maclaurin 展開

$$e^z = 1 + z + \frac{z^2}{2!} + \cdots + \frac{z^x}{x!} + \cdots$$

の速さで減少する稀現象のモデルであり、その時の $z = \lambda$ を強度という。確率母関数 $g(z) = e^{\lambda(z-1)}$ を持ち、平均も分散も λ に一致する。二項分布 $B(n, p)$ を、 $np = \lambda > 0$ を一定にしながら稀現象・大量観測になるような極限分布は、1次元統計多様体をなす。

定義 4.1.6 (Poisson distribution (1838)). 可測空間 $(\mathbb{N}, P(\mathbb{N}))$ 上の **Poisson 分布** $\text{Pois}(\lambda) : (0, \infty) \rightarrow P(\mathbb{N})$ とは、質量関数

$$p(x; \lambda) = \frac{\lambda^x}{x!} e^{-\lambda} \quad x \in \mathbb{N}$$

が定める確率分布である。母数 λ は、単位時間当たりの事象の平均発生回数などの割合と見なされる場合は、**到着率**と呼ばれる。

命題 4.1.7 (分布の特性値).

- (1) well-definedness : $\sum_{x \in \mathbb{N}} p(x; \lambda) = 1$.
- (2) 平均も分散も λ である.
- (3) 確率母関数 : $g(z) = e^{\lambda(z-1)}$.
- (4) したがって積率母関数は $M(t) = e^{\lambda(e^t-1)}$ で、その導関数は
 - (a) $M'(t) = \lambda e^t M(t)$.
 - (b) $M''(t) = (\lambda^2 e^{2t} + \lambda e^t) M(t)$.
 - (c) $M'''(t) = (\lambda^3 e^{3t} + 3\lambda^2 e^{2t} + \lambda e^t) M(t)$.
- (5) よって積率は順に
 - (a) $\alpha_1 = \lambda$.
 - (b) $\alpha_2 = \lambda^2 + \lambda$.
 - (c) $\alpha_3 = \lambda^3 + 3\lambda^2 + \lambda$.

[証明].

- (1) e^λ の Taylor 展開を考えることより、 $\sum_{n \in \mathbb{N}} \frac{\lambda^n}{n!} = e^\lambda$.

- (2) 平均は

$$\sum_{n=1}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} n = \lambda e^{-\lambda} \sum_{n=1}^{\infty} \frac{\lambda^{n-1}}{(n-1)!} = \lambda.$$

分散は、確率母関数 g について $g''(1) + \alpha_1 - (\alpha_1)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

- (3) 確率変数列 $(p(x; \lambda))_{x \in \mathbb{N}} = \left(\frac{\lambda^x}{x!} e^{-\lambda} \right)_{x \in \mathbb{N}}$ が定める母関数は、

$$\begin{aligned} g(z) &= \sum_{x=0}^{\infty} \left(\frac{\lambda^x}{x!} e^{-\lambda} \right) z^x \\ &= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda z)^x}{x!} = e^{-\lambda} e^{\lambda z} = e^{\lambda(z-1)}. \end{aligned}$$

■

歴史 4.1.8 (Poisson 分布を総事故数で条件付けると二項分布を得る (Poisson 1873)). $X_1, X_2 \sim \text{Pois}(\lambda)$ を独立とする。

- (1) Poisson 分布の再生性 : $T := X_1 + X_2 \sim \text{Pois}(2\lambda)$.
- (2) $X_1 | T \sim B(t, 1/2)$.

すなわち、総事故数 T は分布族 $\{\text{Pois}(\lambda)^{\otimes 2}\}_{\lambda \in \mathbb{R}^+}$ の十分統計量である。

[証明]. 母数を $\theta \in \mathbb{R}^+$ で表す. Poisson 分布の再生性 $\text{Pois}(\theta) * \text{Pois}(\theta) = \text{Pois}(2\theta)$ に注意すれば, $T \sim \text{Pois}(2\theta)$ であり, $P_\theta[\{T = t\}]$ は $e^{-2\theta} \frac{(2\theta)^t}{t!}$ と計算出来る. よって, 条件付き確率は素朴に

$$\begin{aligned} P_\theta[\{(x_1, x_2)\} | T = t] &= \frac{P_\theta[\{(x_1, x_2)\} \cap \{T = t\}]}{P_\theta[\{T = t\}]} \\ &= \frac{e^{-\theta} \frac{\theta^{x_1}}{x_1!} e^{-\theta} \frac{\theta^{t-x_1}}{(t-x_1)!}}{e^{-2\theta} \frac{(2\theta)^t}{t!}} \\ &= \frac{1}{2^t} \frac{t!}{x_1!(t-x_1)!} = \frac{1}{2^t} \binom{t}{x_1}. \end{aligned}$$

これは θ に依らない. 特に, $X_1 | T \sim B(t, 1/2)$ を得た. ■

歴史 4.1.9. De Moivre (1711) からすでに利用されていた. Khinchin (1933) は, Poisson 分布の, 離散ジャンプ過程の最も初等的な形としてその歴史的な役割を果たしてきたことを指摘している. [Last and Penrose, 2017]

4.1.4 Poisson 分布の出現

命題 4.1.10 (二項分布の稀現象極限分布). 減衰する確率 $p(n) := \frac{\lambda}{n}$ についての二項分布 $B(n, p(n))$ を考える.

- (1) $\lim_{n \rightarrow \infty} b(x; n, p(n)) = \frac{\lambda^x}{x!} e^{-\lambda} \ (\forall x \in \mathbb{N})$ を満たす.
- (2) $B\left(n, \frac{\lambda}{n}\right) \Rightarrow \text{Pois}(\lambda)$.

[証明].

(1)

$$\begin{aligned} b\left(x; n, \frac{\lambda}{n}\right) &= \binom{n}{x} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \frac{n(n-1) \cdots (n-x+1)}{x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \frac{\lambda^x}{x!} \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right) \left(1 + \frac{-\lambda}{n}\right)^{n/(-\lambda) \cdot (-\lambda)} \left(1 - \frac{\lambda}{n}\right)^{-x} \\ &\xrightarrow{n \rightarrow \infty} \frac{\lambda^x}{x!} e^{-\lambda} \end{aligned}$$

この収束は一様収束である.

(1) から, 任意の $B(n, p(n))$ の独立同分布列 $\{X_n\}$ について,

$$E[f(S_n)] = \sum_{k=0}^n f(k) b(x; n, p(n)), \quad f \in C_b(\mathbb{N}).$$

は $E[f(X)]$ ($X \sim \text{Pois}(\lambda)$) に収束することがわかる. ■

定義 4.1.11 (Poisson process). $\lambda > 0$ を定数として, 次の3条件を満たす確率過程 $(N_t)_{t \in \mathbb{R}_+}$ を **Poisson 過程** または **Poisson 配置** という:

- (1) 強度: $\forall t \in \mathbb{R}_+ \ \forall h > 0 \ P[N_{t+h} - N_t = 1] = \lambda(t)h + o(h) \ (h \searrow 0)$
- (2) 秩序性: $\forall t \in \mathbb{R}_+ \ \forall h > 0 \ P[N_{t+h} - N_t \geq 2] = o(h) \ (h \searrow 0)$.
- (3) $N_0 = 0$ a.s. な独立増分過程である.

$\lambda: \mathbb{R}_+ \rightarrow \mathbb{R}^+$ を強度といい, 定数関数であるとき **定常過程** であるという.

例 4.1.12. 空の星, ケーキの中の干しぶどう, 材料のヒビ, 野原に散らかった紙くずは Poisson 配置に従う. なお, Poisson 過程の共分散関数は $K_\epsilon(t, s) = \lambda(t \wedge s) \ (t, s \in \mathbb{R}_+)$ となる. □

命題 4.1.13 (Poisson 過程の計数的性質). Poisson 過程 (N_t) が定める, 時刻 t までの起こる回数 $(P_k(t) := P[N_t = k])_{k \in \mathbb{N}}$ は \mathbb{N} 上の Poisson 分布 $\text{Pois}(\lambda t)$ である.

[証明].

関係式の導出 区間 $(t, t + \Delta t]$ で点事象が 1 回起こる確率が $\lambda \Delta t + o(\Delta t)$, 2 回以上起こる確率が $o(\Delta t)$ より, 0 回起こる確率が $1 - \lambda \Delta t - o(\Delta t)$ だから,

$$P_k(t + \Delta t) = P_k(t)(1 - \lambda \Delta t - o(\Delta t)) + P_{k-1}(t)(\lambda \Delta t - o(\Delta t)) + o(\Delta t)$$

となる. ただし, $k \in \mathbb{N}, P_{-1}(t) = 0$ とした. $\Delta t \searrow 0$ を考えることで, 微分方程式

$$P'_k(t) = \lambda(P_{k-1}(t) - P_k(t)), \quad P_{-1}(t) = 0$$

を得る.

分布の導出 まず $k = 0$ とすると,

$$P'_0(t) = -\lambda P_0(t), \quad P_0(0) = 1$$

より, $P_0(t) = e^{-\lambda t}$. 次に $k = 1$ とすると,

$$P'_1(t) = \lambda(e^{-\lambda t} - P_1(t)), \quad P_1(0) = 0$$

より, $P_1(t) = \lambda t e^{-\lambda t}$. $k = 2$ とすると,

$$P'_2(t) = \lambda(\lambda t e^{-\lambda t} - P_2(t)), \quad P_2(0) = 0$$

より, $P_2(t) = \frac{(\lambda t)^2}{2} e^{-\lambda t}$. 以下帰納的に, $P_k(t) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$ を得る.

■

[別証明] . 過程 $(P_k(t))_{k \in \mathbb{N}}$ の定める確率母関数 $g(z, t) = \sum_{k=0}^{\infty} P_k(t) z^k$ が, Poisson 分布 $\text{Pois}(\lambda t)$ の確率母関数 $e^{\lambda t(z-1)}$ に一致することを示しても良い.

■

4.1.5 負の二項分布・幾何分布

独立な Bernoulli 試行が成功するまでの「待ち時間」を表す確率変数は, 幾何分布に従う. [森岡毅 and 今西聖貴, 2016] では消費者の特定ブランドの購買回数を負の二項分布でモデリングしている.

定義 4.1.14 (negative binomial distribution / Pascal distribution, geometric distribution).

- (1) 負の二項分布または **Pascal 分布** $\text{NB}(k, p) : \mathbb{N}^+ \times [0, 1] \rightarrow P(\mathbb{N})$ を, Bernoulli 試行 $\text{Ber}(p)$ が k 回成功するために必要な失敗数 $x \in \mathbb{N}$ が従う確率分布とする.
- (2) 確率関数は, 合計 $k + x$ 回の試行で最後の一回は必ず成功である必要があるから,

$$p(x; k, p) = \binom{x+k-1}{x} p^k q^x \quad (x \in \mathbb{N}^+, q := 1 - p)$$

と表せる.

- (3) 特に $k = 1$ の場合を **幾何分布** $G(p) := \text{NB}(1, p)$ といい, 確率関数は $p(x; 1, p) = p q^x$ ($x \in \mathbb{Z}_+$) と表せる. $(p(n))_{n \in \mathbb{N}}$ が初項 p で公比 q の等比級数を定めるためである.
- (4) 確率母関数を見れば判る通り, 負の二項分布 $\text{NB}(k, p)$ は $k \in \mathbb{R}^+$ 上に延長できる.

補題 4.1.15 (well-definedness).

- (1) 次の等式が $z \in \Delta, \alpha \in \mathbb{R}$ 上で成り立つ:

$$(1+z)^\alpha = 1 + \sum_{n=1}^{\infty} \frac{\alpha!}{n!(\alpha-n)!} z^n.$$

(2) $NB \in P(\mathbb{N})$ である. すなわち,

$$\sum_{k=m}^{\infty} \binom{k-1}{m-1} p^m (1-p)^{k-m}.$$

命題 4.1.16 (負の二項分布の確率母関数・平均・分散).

(1) 負の二項分布 $NB(k, p)$ の確率母関数は $g(z) = p^k (1 - qz)^{-k} = \left(\frac{p}{1 - qz} \right)^k$ で表せる.

(2) 負の二項分布 $NB(k, p)$ の平均は $\alpha_1 = \frac{kq}{p}$, 分散は $\mu_2 = \frac{kq}{p^2}$ である.

[証明].

(1) $|zq| < 1$ すなわち $|z| < \frac{1}{q}$ のとき, $(1 - qz)^{-k}$ の $z = 0$ における Taylor 展開を考えると

$$\begin{aligned} (1 - qz)^{-k} &= 1 + k(qz) + \frac{k(k+1)}{2!}(qz)^2 + \cdots = \sum_{x=0}^{\infty} \frac{(-k)(-k-1)\cdots(-k-x+1)}{x!} (-qz)^x \\ &= \sum_{x=0}^{\infty} \frac{(x+k-1)(x+k-2)\cdots(k+1)k}{x!} (qz)^x \\ &= \sum_{x=0}^{\infty} \binom{x+k-1}{x} (qz)^x. \end{aligned}$$

$$\left. \frac{(-k)(-k-1)\cdots(-k-x+1)}{x!} \right|_{x=0} = 1 \text{ としたことに注意. よって, } NB(k, p) \text{ の確率母関数は } g(z) = p^k (1 - qz)^{-k}.$$

(2) 確率母関数と平均・分散の関係 1.6.4(3) より,

$$\begin{aligned} g'(z) &= kp^k q (1 - qz)^{-k-1}, & \mu &= g'(1) = \frac{kp^k q}{p^{k+1}} = \frac{kq}{p}, \\ g''(z) &= k(k+1)p^k q^2 (1 - qz)^{-k-2}, & \sigma^2 &= g''(1) + g'(1) - (g'(1))^2 \\ & & &= \frac{k(k+1)q^2}{p^2} + \frac{kq}{p} - \frac{k^2 q^2}{p^2} \\ & & &= \frac{kq(p+q)}{p^2} = \frac{kq}{p^2}. \end{aligned}$$

命題 4.1.17 (Poisson 近似). 平均 $\lambda := \frac{kq}{p} \in (0, \infty)$ を一定にして, 成功数を表す母数を $k \rightarrow \infty$ とすると (このとき失敗率 q は極めて小さくなる), Poisson 分布 $\text{Pois}(\lambda)$ に収束する: $NB(k, p) \rightarrow \text{Pois}(\lambda)$ ($k \rightarrow \infty$).

[証明]. 確率関数が $\tilde{p}(x) = \frac{\lambda^x}{x!} e^{-\lambda}$ に収束することを見れば良い.

$$\begin{aligned} p(x; k, p) &= \binom{x+k-1}{x} p^k q^x \\ &= \frac{(x+k-1)\cdots(k+1)k}{x!} p^k q^x \\ &= \frac{(x+k-1)\cdots(k+1)k}{k^x} x! \left(\frac{kq}{p} \right)^x p^{k+x} \\ &\xrightarrow[\lambda=\text{const.}]{k \rightarrow \infty} 1 \cdot \frac{\lambda^x}{x!} e^{-\lambda}. \end{aligned}$$

実際,

$$p^{k+x} = \left(\frac{p}{p+q} \right)^{k+x} = \left(1 + \frac{\lambda}{k} \right)^{-(k+x)} \xrightarrow{k \rightarrow \infty} e^{-\lambda}.$$

注 4.1.18 (幾何分布の無記憶性). $X \sim G(\theta)$ とするとき,

$$P(X \geq m+n | X \geq m) = P(X \geq n) \quad (m, n \in \mathbb{N})$$

が成り立つ。これは、時刻 m までに成功していないことはその後の成功までの待ち時間の分布に影響しないことを意味しているとみなせる。

4.1.6 Katz 族

定義 4.1.19 (Katz family). \mathbb{N} 上の離散分布の族であって、 $a, b \in \mathbb{R}$ を用いた漸化式

$$p_0 > 0, \quad p_x = \left(a + \frac{b}{x}\right) p_{x-1}, \quad x \in \mathbb{N}$$

を満たす族を **Katz 族** という。この漸化式を満たすことを、 $(a, b, 0)$ -クラスの分布であるという。

例 4.1.20.

- (1) $\delta_0 : a + b = 0, p_0 = 1$ のとき.
- (2) $B(n, \theta) : a = -\frac{\theta}{1-\theta}, \quad b = \frac{(n+1)\theta}{1-\theta}, \quad p_0 = (1-\theta)^n.$
- (3) $\text{Pois}(\lambda) : a = 0, b = \lambda, p_0 = e^{-\lambda}.$
- (4) $\text{NB}(k, \theta) : a = 1-\theta, b = (k-1)(1-\theta), p_0 = \theta^k.$

□

命題 4.1.21 (実際の母数はかなり狭い). \mathbb{N} 上の確率分布 $p = (p_x)_{x \in \mathbb{N}}$ が Katz 族であるとする。このとき、 p は上の例の4つの場合に限られる。

- (1) $a + b = 0$ ならば、 $p = \delta_0$.
- (2) $a = 0, b > 0$ ならば、 $p = \text{Pois}(b)$.
- (3) $a + b > 0, a \in (0, 1)$ ならば、 $p = \text{NB}((a+b)/a, 1-a)$.
- (4) $a + b > 0, a \in (-\infty, 0)$ ならば、 $\exists_{n \in \mathbb{N}} a(n+1) + b = 0, p = B(n, -a/(1-a))$.

特に、 p がデルタ測度 δ_0 でなければ、 $a + b > 0$ かつ $a < 1$ であり、確率母関数は

$$g(z) = \begin{cases} \left(\frac{1-az}{1-a}\right)^{-\frac{a+b}{a}}, & a \neq 0, \\ \exp(b(z-1)), & a = 0. \end{cases}$$

と表せる。

4.1.7 超幾何分布

母集団 N のうちの n 個体が、サイズ r の標本調査を行った際に x 個見つかる確率は超幾何分布に従う。これは確率母関数が Gauss の超幾何関数を用いて表せるためである。 $N \rightarrow \infty$ の極限では、確率 $n/N =: p$ の Bernoulli 試行を r 回繰り返した場合に等しい。幾何分布には無記憶性があり、二項分布は Bernoulli 試行の独立反復であるが、初めての従属試行列を考えているとみなせる。

定義 4.1.22 (hypergeometric distribution). 成功 n 個、失敗 $N - n$ 個が入った多重集合から r 個玉を取り出したときにそのうち成功が x 個である確率

$$p(x; N, n, r) = \frac{\binom{n}{x} \binom{N-n}{r-x}}{\binom{N}{r}}$$

が、 $\Omega := \{x \in \mathbb{N} \mid \max\{0, r - N + n\} \leq x \leq \min\{r, n\}\}$ 上に定める確率分布 $\text{HG}(N, n, r)$ を **超幾何分布** という。

命題 4.1.23.

- (1) 平均は $\alpha_1 = r \frac{n}{N}$.

- (2) 分散は $\mu_2 = r \left(\frac{N-r}{N-1} \right) \left(\frac{n}{N} \right) \left(1 - \frac{n}{N} \right)$.
 (3) 確率母関数は $F(\alpha, \beta, \gamma; z)$ を超幾何関数として,

$$g(z) = \frac{\binom{N-n}{r} F(-r, -n, N-n-r+1; z)}{\binom{N}{r}}.$$

4.1.8 超幾何分布の二項近似

超幾何分布は漸近論の導入により二項分布として扱える.

例 4.1.24 (mark and recapture method における Lincoln-Peterson 推定量). 個体数 N を推定するために, n 匹を捕まえて, 標識をつけて再放流する. 十分な時間経過後に sampling し, 標識がついているものが r 匹だった場合, $\tilde{N} := \frac{nr}{x}$ によって N が推定できる. これは, 確率 $p(x; N, n, r)$ が最大になるような N の選び方で, 最尤法の考え方である. \square

例 4.1.25 (Fisher の正確検定 (exact test 1922)). 2 つの 2 値変数の間に有意な関連があるかどうかの検定においては, 超幾何分布が使える. これは Fisher が発見した. しかし, 標本が大きいなどの場合は計算困難になるが (式が階乗を含むため), その場合は χ^2 -検定が可能である. \square

命題 4.1.26 (二項近似). $HG(N, n, r)$ は, $N \rightarrow \infty$ のとき, $\lim_{N \rightarrow \infty} \frac{n}{N} =: p$ とすれば, $B(n, p)$ で近似される.

4.1.9 超幾何関数

問題 4.1.27 (hypergeometric differential equation).

$$(GE) \quad t(1-t)x'' + (\gamma - (\alpha + \beta + 1)z)x' - \alpha\beta x = 0, \quad \alpha, \beta, \gamma \in \mathbb{C}.$$

を **Gauss の超幾何微分方程式** という. 特異点は $\{0, 1, \infty\} \subset \hat{\mathbb{C}}$ に持つ.

- (1) $U \subset^{\text{open}} \hat{\mathbb{C}} \setminus \{0, 1, \infty\}$ を領域, $\mathcal{F}(U)$ を U 上の正則解全体とする.
 (2) $(\alpha)_n := \alpha(\alpha+1) \cdots (\alpha+n-1) = \frac{\Gamma(\alpha+n)}{\Gamma(\alpha)}$ とする. これを Pochhammer 記号という.
 (3) 次の冪級数が収束するならば, これが定める Δ 上の関数を **Gauss の超幾何関数** という:

$$F(\alpha, \beta, \gamma; x) := \sum_{n \in \mathbb{N}} \frac{(\alpha)_n (\beta)_n}{(\gamma)_n n!} x^n, \quad |x| < 1, \alpha, \beta, \gamma \in \mathbb{C}.$$

例えば, γ が負整数のとき, 定義されない.

存在するならば, これは方程式 (GE) の解になっていることが分かる. Gauss の超幾何関数は解析接続によって大域的な性質がよく分かる稀有な関数である.

命題 4.1.28 (Gauss の積分表示).

- (1) $0 < \operatorname{Re} \beta < \operatorname{Re} \gamma, 0 < \arg(x-1) < 2\pi$ において, 次の積分表示を持つ:

$$F(\alpha, \beta, \gamma; x) = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)} \int_0^1 t^{\beta-1} (1-t)^{\gamma-\beta-1} (1-tx)^{-\alpha} dt.$$

- (2) 特に, $\operatorname{Re}(\gamma - \beta - \alpha) > 0$ のとき, $x = 1$ とすれば, 右辺の積分は Beta 関数となり, このとき

$$F(\alpha, \beta, \gamma; 1) = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)} \frac{\Gamma(\beta)\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)}$$

よって,

$$C_{\alpha, \beta, \gamma} := F(\alpha, \beta, \gamma; 1) = \frac{\Gamma(\gamma)\Gamma(\gamma-\alpha-\beta)}{\Gamma(\gamma-\alpha)\Gamma(\gamma-\beta)}.$$

さらに, $\operatorname{Re} \alpha + \operatorname{Re} \beta < \operatorname{Re} \gamma$ ならば, 左辺の級数は絶対収束する.

注 4.1.29. (1) の積分表示も,

$$(1 - tx)^{-\alpha} = \sum_{n \in \mathbb{N}} \frac{(\alpha)_n}{n!} t^n x^n$$

として, x^n の係数を見ると, 各項の積分が Beta 関数になり,

$$\frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)} \frac{(\alpha)_n}{n!} \frac{\Gamma(\beta+n)\Gamma(\gamma-\beta)}{\Gamma(\gamma+n)} = \frac{(\alpha)_n(\beta)_n}{(\gamma)_n n!}$$

を得るため, 等式の成立がわかる.

命題 4.1.30 (Riemann). $F(\alpha, \beta, \gamma; x)$ は Gauss の超幾何微分方程式の解であり, 原点で正則でかつ 1 の値を取る解はこれに限る.

4.1.10 負の超幾何分布

4.1.11 対数分布

定義 4.1.31. \mathbb{N} 上の分布・対数分布 $\text{Log}(\theta) := (p_x)_{x \in \mathbb{N}}$ ($\theta \in (0, 1)$) とは, 確率関数

$$p_x = \frac{-1}{\log(1-\theta)} \frac{\theta^x}{x} 1_{\mathbb{N}^+}$$

で定まる分布をいう.

命題 4.1.32. $\text{Log}(\theta)$ の確率母関数は $g(z) = \frac{\log(1-\theta z)}{\log(1-\theta)}$ である.

4.1.12 Ord 族

定義 4.1.33 (Ord family). \mathbb{Z} 上の分布であって, 確率関数が差分方程式

$$p_x - p_{x-1} = \frac{(a-x)p_x}{(a+b_0) + (b_1-1)x + b_2x(x-1)}$$

を満たすものの全体を **Ord 族**という.

例 4.1.34. Katz 族と, 超幾何分布, 負の超幾何分布は Ord 族である. □

4.2 多次元純不連続分布の例

4.2.1 多項分布

二項分布は多項分布の周辺分布である. 確率母関数が

$$g_I(z_I; p) = (p_1 z_1 + \cdots + p_k z_k)^n$$

と表せることは多項定理の主張に外ならず, ここから特性量が得られる.

定義 4.2.1 (multinomial distribution). $n \in \mathbb{N}^+$ の k 成分への分解 (n 回の試行のそれぞれがどの結果 E_i で終わるかの割り当て) の全体

$$T_I := \left\{ x_I = (x_j)_{j \in I} \in \mathbb{N}^k \mid \sum_{j=1}^k x_j = n \right\}$$

と 1 の k 成分への分解 (それぞれの事象 E_i の生起確率) $p_1, \dots, p_k, \sum_{j=1}^k p_j = 1$ について, 次の質量関数

$$f_I(x_I; p) = \frac{n!}{x_1! \cdots x_k!} \prod_{j=1}^k p_j^{x_j} \quad (x_I = (x_j)_{j \in I} \in T_I)$$

で定まる分布を k -項分布, 一般に**多項分布**といい, $\text{Mult}_n(p_1, \dots, p_k)$ で表す.

要諦 4.2.2. これは、各事象 E_1, \dots, E_k の生起確率が $P[E_i] = p_i$ で、それぞれの起きた回数が x_1, \dots, x_k 回であるという確率変数が従う分布となる。それぞれの変数 X_i の周辺分布は $X_i \sim B(n, p_i)$ である。

定理 4.2.3. $X = (X_1, \dots, X_k) \sim \text{Mult}(n; p_1, \dots, p_k)$ とする。

(1) $g_I(z_I; p) = (p_1 z_1 + \dots + p_k z_k)^n$ が確率母関数である。特に、積率母関数は、

$$M(t) = (p_1 e^{t_1} + \dots + p_k e^{t_k})^n.$$

(2) $E[X_i] = np_i$.

(3) $\text{Var}[X_i] = np_i(1 - p_i)$.

(4) $\text{Cov}[X_{i_1}, X_{i_2}] = -np_{i_1}p_{i_2} \ (i_1 \neq i_2 \in [k])$.

(5) $g_I(z_I; p) = (p_1 z_1 + \dots + p_k z_k)^n$ が確率母関数である。

[証明].

(1) 多項定理より、

$$\begin{aligned} g(z) &= \sum_{x_I \in T_I} f_I(x_I; p) z_1^{x_1} \cdot z_k^{x_k} \\ &= (p_1 z_1 + \dots + p_k z_k)^n. \end{aligned}$$

特に、 $z_i = e^{u_i}$ として、積率母関数を得る。

(2) 両辺の $(u_1, \dots, u_k) = (0, \dots, 0)$ における u_i 偏微分係数より、 $np_i = \sum_{x_1, \dots, x_k} *x_i P^X[\{(x_1, \dots, x_k)\}] = E[X_i]$.

(3) 同様に u_i の2階微分を考えて、 $E[X_i^2] = n(n-1)p_i^2 + np_i$ を得る。よって、 $\text{Var}[X_i] = E[X_i^2] - (E[X_i])^2 = np_i(1 - p_i)$.

(4) $i_1 \neq i_2$ に関して、順に u_{i_1}, u_{i_2} での偏微分を考えることにより、 $E[X_{i_1} X_{i_2}] = n(n-1)p_{i_1}p_{i_2}$ 。よって、共分散公式??より、

$$\text{Cov}[X_{i_1}, X_{i_2}] = E[X_{i_1} X_{i_2}] - E[X_{i_1}]E[X_{i_2}] = -np_{i_1}p_{i_2}.$$

■

要諦 4.2.4. 初等的な証明が技巧的すぎる、多項定理に指数関数を代入する。

命題 4.2.5 (順序統計量). r 番目の順序統計量 $X_{n:r}$ の分布関数 F_r は多項分布に従う。

4.2.2 2変量 Poisson 分布

定義 4.2.6 (bivariate Poisson distribution). 確率変数 U_1, U_2, U_3 が独立で $U_i \sim \text{Pois}(\lambda_i)$ を満たすとする。このとき、 $X_1 = U_1 + U_3, X_2 = U_2 + U_3$ で定まる確率変数 $(X_1, X_2) : \Omega \rightarrow \mathbb{N}^2$ が定める分布を **2変量 Poisson 分布**といい、 $\text{BPois}(\lambda_1, \lambda_2, \lambda_3)$ で表す。周辺分布は $X_1 \sim \text{Pois}(\lambda_1 + \lambda_3), X_2 \sim \text{Pois}(\lambda_2 + \lambda_3)$ である。

命題 4.2.7 (分散や混合積率は確率母関数の項別微分で求める).

(1) 確率母関数は $g(z_1, z_2) = \exp(\lambda_1(z_1 - 1) + \lambda_2(z_2 - 1) + \lambda_3(z_1 z_2 - 1))$.

(2) $\text{Cov}[X_1, X_2] = \lambda_3$.

4.2.3 負の多項分布

負の二項分布 $NB(k, p)$ の確率母関数は、 $\hat{q} := 1/p, \hat{p} := q/p = (1 - p)/p$ とすると、

$$g(z) = p^k (1 - qz)^{-k} = (\hat{q} - \hat{p}z)^{-k}$$

と表わせ、この形で一般化を考える。

定義 4.2.8. $k > 0, P_i > 0$ をパラメータとする**負の多項分布** $NM(k, (P_1, \dots, P_d))$ とは, $Q := 1 + \sum_{i=1}^d P_i$ とするとき, 確率母関数

$$g(z_1, \dots, z_d) = \left(Q - \sum_{i=1}^d P_i z_i \right)^{-k}$$

が定める \mathbb{N}^d 上の分布を言う.

命題 4.2.9 (確率母関数の微分からわかること).

$$(1) \operatorname{Cov}[X_i, X_j] = \begin{cases} kP_i(1 + P_i) & i = j \\ kP_iP_j & i \neq j \end{cases}$$

4.3 正規分布

4.3.1 定義と特性値

特性関数 $\varphi(u) = \exp\left(i\mu \cdot u - \frac{u^\top \Sigma u}{2}\right)$ で定まる分布である.

定義 4.3.1 (normal distribution). $(\mathbb{R}, \mathcal{B}_1)$ 上の**正規分布** $N(\mu, \sigma^2)$ ($\mu \in \mathbb{R}, \sigma > 0$) とは, 確率密度関数

$$\phi(x; \mu, \sigma^2) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$

が定める分布をいう.

補題 4.3.2 (Gamma 関数の性質).

$$(1) \int_{\mathbb{R}} \exp(-x^2) dx = \sqrt{\pi} = \Gamma(1/2).$$

$$(2) \forall \mu, \sigma \in \mathbb{C} \quad \int_{\mathbb{R}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx = \sqrt{2\pi\sigma^2}.$$

[証明].

- (1) この積分を I とおき, I^2 を Fubini の定理と極座標変換を用いて計算すれば良い. あるいは, $u = x^2$ と変換すると Gamma 関数の定義式を得る.
- (2) 変数変換により $\mu, \sigma \in \mathbb{R}$ の場合で求めてから, $\mu, \sigma \in \mathbb{C}$ の場合について解析接続する. ■

命題 4.3.3 (分布の母関数). $N(\mu, \sigma^2)$ について,

- (1) well-definedness : $\int_{\mathbb{R}} \phi(x; \mu, \sigma^2) dx = 1$.
- (2) 特性関数 : $\varphi(u) = \exp\left(i\mu u - \frac{1}{2}\sigma^2 u^2\right)$.
- (3) 平均と分散 : $\alpha_1 = \mu, \mu_2 = \sigma^2$.
- (4) 積率母関数 : $\mathfrak{M}(t) = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$.
- (5) キュムラントは $\kappa_1 = \mu, \kappa_2 = \sigma^2, \kappa_r = 0$ ($r \geq 3$).^{†1}

[証明].

- (1) $u = \frac{x - \mu}{\sqrt{2\sigma}}$ と変数変換すると, $dz = \frac{dx}{\sqrt{2\sigma}}$ より,

$$\int_{\mathbb{R}} \phi(x; \mu, \sigma^2) dx = \frac{1}{\sqrt{\pi}} \int_{\mathbb{R}} e^{-u^2} du = 1.$$

^{†1} この3次以上のキュムラントが消えることが正規分布の特徴で, 積率による中心極限定理の証明に利用される.

(2) $X \sim N(\mu, \sigma^2)$ について, 次のように計算できる:

$$\begin{aligned}\varphi(u) &= E[e^{iuX}] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{iux} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2} \{(x - (\mu + iu\sigma^2))^2 - 2\mu iu\sigma^2 + u^2\sigma^4\}\right) dx \\ &= \exp\left(iu\mu - \frac{u^2\sigma^2}{2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2} ((x - (\mu + iu\sigma^2))^2)\right) dx = \exp\left(iu\mu - \frac{u^2\sigma^2}{2}\right).\end{aligned}$$

(3) 特性関数より 1.7.12,

$$\begin{aligned}\alpha_1 &= \frac{1}{i} \varphi'(0) = \frac{1}{i} (i\mu) = \mu, \\ \mu_2 &= -\varphi''(0) + (\varphi'(0))^2 = \sigma^2 - (i\mu)^2 + (i\mu)^2 = \sigma^2.\end{aligned}$$

(4) $\varphi(u)$ は明らかに整関数だから, 変数変換 $u = it$ による.

(5) キュムラント母関数は $\mathfrak{C}(t) = \mu t + \frac{\sigma^2 t^2}{2}$ であることがわかった. よって, $\kappa_1 = \mathfrak{C}'(0) = \mu, \kappa_2 = \mathfrak{C}''(0) = \sigma^2$ で, あとは消える.

■

命題 4.3.4 (正規分布の積率). $N(\mu, \sigma^2)$ について,

(1) 積率: n が奇数のときは $\alpha_n = 0$, n が偶数のときは

$$\alpha_n = \sum_{m=0}^n \binom{n}{m} \sigma^m \mu^{n-m} I_m, \quad I_m = (m-1)!!.$$

(2) 中心積率:

$$\mu_{2r+1} = 0, \quad \mu_{2r} = \frac{(2r)!}{2^r r!} \sigma^{2r} = (2r-1)!! (\sigma^2)^r.$$

(3) 歪度と尖度: $\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = 0, \gamma_2 = \frac{\mu_4}{\mu_2^2} = 3.$

[証明].

(1) 部分積分を繰り返す.

(2) 部分積分を繰り返して計算するのみである. $X - \mu$ を X とおき直すことで, $X \sim N(0, \sigma^2)$ について,

$$\begin{aligned}\mu_3 &= \int_{\mathbb{R}} x^3 e^{-\frac{x^2}{2\sigma^2}} dx \\ &= -\sigma^2 \int_{\mathbb{R}} x^2 \left(e^{-\frac{x^2}{2\sigma^2}} \right)' dx \\ &= -\sigma^2 \left[x^2 e^{-\frac{x^2}{2\sigma^2}} \right]_{-\infty}^{\infty} + 2\sigma^2 \int_{\mathbb{R}} x e^{-\frac{x^2}{2\sigma^2}} dx \\ &= -\sigma^2 \cdot 0 + 2\sigma^2 \cdot E[X] = 0.\end{aligned}$$

同様にして,

$$\begin{aligned}\mu_4 &= \int_{\mathbb{R}} x^4 e^{-\frac{x^2}{2\sigma^2}} dx \\ &= -\sigma^2 \int_{\mathbb{R}} x^3 \left(e^{-\frac{x^2}{2\sigma^2}} \right)' dx \\ &= -\sigma^2 \left[x^3 e^{-\frac{x^2}{2\sigma^2}} \right]_{-\infty}^{\infty} + 3\sigma^2 \int_{\mathbb{R}} x^2 e^{-\frac{x^2}{2\sigma^2}} dx = 3\sigma^4.\end{aligned}$$

あとは帰納的にわかる.

(3) $\gamma_1 = 0/\sigma^3 = 0, \gamma_2 = 3\sigma^4/\sigma^4 = 3.$

■

4.3.2 多変量正規分布の定義

特性関数の収束を用いて、一般の半正定値行列についても (退化した) 多変量正規分布を定義できる。

定義 4.3.5 (multivariate normal distribution). $\mu \in \mathbb{R}^d$, $\Sigma \in \text{GL}_d(\mathbb{R})$ を正定値実対称行列として, $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ 上の測度

$$\mu_{\mu, \Sigma}(dx) = \phi(x; \mu, \Sigma) = \frac{1}{(2\pi)^{d/2}(\det \Sigma)^{1/2}} \exp\left(-\frac{1}{2}(x - \mu)^\top \cdot \Sigma^{-1}(x - \mu)\right) dx$$

を, 平均値 μ , 共分散行列 Σ に関する d 次元正規分布といい, $N_d(0, \Sigma)$ で表す。

例 4.3.6. $N(\mu, \sigma^2)^{\otimes n} = N_n(\mu \mathbf{1}_n, \sigma^2 I_n)$ のとき, $\det \Sigma = \sigma^{2n}$, $\Sigma^{-1} = \frac{1}{\sigma^2} I_n$ から,

$$-\frac{1}{2}(x - \mathbf{1}_n \mu)^\top \frac{1}{\sigma^2} I_n (x - \mathbf{1}_n \mu) = -\frac{1}{2\sigma^2} \sum_{i \in [n]} (x_i - \mu)^2.$$

より, 密度関数は

$$\phi(x; \mu \mathbf{1}_n, \sigma^2 I_n) = \frac{1}{(2\pi\sigma^2)^{n/2}} e^{-\frac{1}{2\sigma^2} \sum_{i \in [n]} (x_i - \mu)^2}.$$

係数部分が $e^{-\psi(\theta)}$ に当たり,

□

命題 4.3.7 (多変量正規分布の母関数). d 次元正規分布 $N_d(0, \Sigma)$ について,

(1) 特性関数は

$$\varphi(u) = \exp\left(i\mu \cdot u - \frac{1}{2}u^\top \Sigma u\right)$$

(2) キュムラント母関数は $\psi(u) = \log \varphi(u) = i\mu \cdot u - \frac{1}{2}u^\top \Sigma u$ である。

(3) 特に, 平均ベクトルは μ , 分散共分散行列が Σ で, 3 次以上のキュムラントはすべて消えている。

定義 4.3.8 (半正定値対称行列が定める多変量正規分布). $\Sigma \in M_d(\mathbb{R})$ を半正定値対称行列とする (特に, 非退化の可能性はある)。列 (Σ_n) を $\Sigma_n := \Sigma + n^{-1}I_n$ で定めるとこれは正定値行列の列で, $\Sigma_n \xrightarrow{n \rightarrow \infty} \Sigma$ 。対応する正規分布 $\nu_n := N_d(\mu, \Sigma_n)$ の特性関数の列は

$$\varphi_{\nu_n}(u) = \exp\left(iu^\top \mu - \frac{1}{2}u^\top \Sigma_n u\right) \xrightarrow{n \rightarrow \infty} \varphi(u) = \exp\left(iu^\top \mu - \frac{1}{2}u^\top \Sigma u\right)$$

と各点収束し, 原点で連続である。よって, Bochner の定理 1.7.24 より, これはある確率分布 ν の特性関数である。この分布を, d 変量正規分布 $N_d(\mu, \Sigma)$ と呼ぶ。

4.3.3 条件付き分布

正規確率変数の間の条件付き確率分布は再び正規である。また, 正規確率変数の affine 変換像は再び正規である。正規確率変数が直交することと独立であることは同値であるから, ユニタリ変換は成分間の独立性を変えない。

定理 4.3.9 (正規確率変数同士の条件付き分布). $X = (X_1, X_2) \sim N_n(\mu, \Sigma)$, $X_1 \in \mathbb{R}^p$, $X_2 \in \mathbb{R}^q$ で, 共分散行列は正則 $\Sigma \in \text{GL}_n(\mathbb{R})$ とする。

(1) X_1, X_2 の分布は, 対応する μ, Σ の分割について, $N_p(\mu_1, \Sigma_{11}), N_q(\mu_2, \Sigma_{22})$ に従う。

(2) $X_2|X_1$ の条件付き分布は再び q 次元正規分布で,

(a) 平均ベクトルは $E[X_2|X_1] = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1)$ 。

(b) 共分散行列は $\text{Var}[X_2|X_1] = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$ 。

命題 4.3.10 (多次元正規確率変数の成分間の独立性). $Y_1 := (X_1, \dots, X_{k_1})^\top, \dots, Y_l := (X_{k_{l-1}+1}, \dots, X_d)^\top$ ($l \geq 2$) のように, X を l 個の確率変数 Y_1, \dots, Y_l に分ける。この分割に対して, Σ のブロック $\Sigma_{a,b} := \text{Cov}[Y_a, Y_b]$ ($a, b \in [l]$) を考える。

- (1) Y_1, \dots, Y_l は独立.
 (2) $\forall a, b \in [l] \ a \neq b \Rightarrow \Sigma_{a,b} = 0$.

系 4.3.11 (ユニタリ変換は成分間の独立性を変えない). $X \sim N_d(\mu, \sigma^2 I_d), U \in M_d(\mathbb{R})$ を直交行列とする. このとき, $UX \sim N_d(U\mu, \sigma^2 I_d)$ である. 特に, UX の成分は再び独立である.

4.3.4 線型汎函数による特徴付け

線型汎函数像が正規になることは, 多変量正規分布を特徴付ける. また, 一般の確率変数について, 任意の線型変換像が分布収束することは, 結合分布が分布収束することの特徴付ける.

定理 4.3.12 (Gauss 測度の特徴付け). $X \in L(\Omega, \mathbb{R}^n)$ について, 次の2条件は同値:

- (1) $X \sim N_n(\mu, \Sigma)$.
 (2) $\forall t \in \mathbb{R}^n \ tX \sim N(t\mu, t\Sigma t)$.

[証明]. X の特性関数

$$\varphi(u) = E[e^{iu^\top X}] = \exp\left(iu^\top \mu - \frac{1}{2}u^\top \Sigma u\right)$$

について, $u = st \ (s \in \mathbb{R})$ への制限

$$\varphi|_{\mathbb{R}t}(s) = E[e^{ist^\top X}] = \exp\left(is(t^\top \mu) - \frac{1}{2}s^2(t^\top \Sigma t)\right)$$

は, $t^\top X$ の特性関数でもあり, その形は $N(t^\top \mu, t^\top \Sigma t)$ のものである. 逆も同様に辿れる. ■

要諦 4.3.13. 一般の次元 (または無限次元) において (2) の定義を用いるのは, X が退化している場合でも通用する定義であるためである.

系 4.3.14 (多変量正規分布の線型変換の分布). $X \sim N_d(\mu, \Sigma), A \in M_{dn}(\mathbb{R})$ について, $A^\top X \sim N_n(A^\top \mu, A^\top \Sigma A)$.

[証明]. 任意の $t \in \mathbb{R}^n$ について, $t^\top (A^\top X) = (At)^\top X$ は正規分布 $N_d(t^\top A^\top \mu, t^\top (A^\top \Sigma A)t)$ に従う. ■

命題 4.3.15 (Cramer-Wald). 確率変数 $X_n = (X_{n1}, \dots, X_{nk}), Y = (Y_1, \dots, Y_k)$ について, 次は同値:

- (1) $X_n \Rightarrow Y$.
 (2) $\forall t \in \mathbb{R}^k \ t^\top X_n \Rightarrow t^\top Y$.

[証明]. 連続写像定理と, 特性関数の議論による. ■

4.3.5 3 次以上のキュムラントが消える

3 次以上のキュムラントが (存在しながに) 消えることは正規確率変数を特徴付ける.

定理 4.3.16. $\mu \in P(\mathbb{R})$ について, 次は同値.

- (1) μ は正規分布である.
 (2) μ の任意次数のキュムラントが存在し, 3 次以上は消えている.

4.3.6 正規確率変数の独立和への分解

定理 4.3.17 (Cramer (36)). $X_1, X_2 \in L(\Omega)$ を独立とする. 次の2条件は同値.

- (1) 和 $X_1 + X_2$ が正規確率変数である.

(2) X_1, X_2 はいずれも正規確率変数である.

定理 4.3.18 (Skitovich-Darmois). $X := (X_1, \dots, X_n)$ を独立, これらの2つの異なる一次結合を $L_1 = a^\top X, L_2 = b^\top X$ ($a, b \in \mathbb{R}^n$) とする. (1) \Rightarrow (2) が成り立つ.

(1) $L_1 \perp L_2$.

(2) $a_j b_j \neq 0$ を満たす $j \in [n]$ について, X_j は正規.

系 4.3.19. 例えば, 独立確率変数 X_1, X_2 について, $X_1 + X_2 \perp X_1 - X_2$ ならば, X_1, X_2 は正規に限る.

定理 4.3.20 (Kac (40)). $X_1, X_2 \in L(\Omega)$ を独立とする.

$$\begin{cases} Y_1 = X_1 \cos \theta + X_2 \sin \theta, \\ Y_2 = X_1 \sin \theta - X_2 \cos \theta. \end{cases}$$

について, 次は同値.

(1) 任意の $\theta \in [0, 2\pi]$ について, $Y_1 \perp Y_2$.

(2) ある $\sigma^2 > 0$ が存在して, $X_1, X_2 \sim N(0, \sigma^2)$.

4.3.7 統計量の分布

一般の (2 次の積率をもつ) 分布について, 標本平均 \bar{X} と不偏分散 U^2 は, いずれも不偏推定量である.

定義 4.3.21. $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ を独立標本とする.

(1) 次の統計量を**標本平均**という:

$$\bar{X} := \frac{1}{n} \sum_{i \in [n]} X_i.$$

(2) 次の統計量を**標本分散**という:

$$S^2 := \frac{1}{n} \sum_{i \in [n]} (X_i - \bar{X})^2.$$

(3) 次の統計量を**不偏分散**という:

$$U^2 := \frac{1}{n-1} \sum_{i \in [n]} (X_i - \bar{X})^2 = \frac{1}{n(n-1)} \sum_{i < j \in [n]} (X_i - X_j)^2.$$

(4) 次の統計量を, **母平均を用いた標本分散**という:

$$S^2(\mu_0) := \frac{1}{n} \sum_{i \in [n]} (X_i - \mu_0)^2.$$

\bar{X}, U^2 は U -統計量である.

定理 4.3.22. $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ を独立標本とする. 標本平均 \bar{X} と標本の2次の中心積率 S^2 について, 次が成り立つ:

(1) $\bar{X} \perp U^2$ で, $\bar{X} \sim N(\mu, \sigma^2/n)$ かつ $U^2/\sigma^2 \sim \chi^2(n-1)$.

(2) (\bar{X}, S^2) は最小十分統計量である.

(3) $\bar{X}, S^2 := \sum_{i=1}^n (X_i - \bar{X})^2, W := \left(\frac{X_1 - \bar{X}}{S}, \dots, \frac{X_n - \bar{X}}{S} \right)$ は独立であり, W は $n-2$ 次元単位球面上の一様分布に従う:
 $W \sim U(S^{n-2})$. これは補助統計量の例である.

(4) 統計量 (\bar{X}, S^2) は完備である: $\forall g \in L(\mathbb{R}) \quad \forall (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+ \quad E_{N(\mu, \sigma^2)}[g(\bar{X}, S^2)] = 0 \Rightarrow \forall (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+ \quad P_{N(\mu, \sigma^2)}[g(\bar{X}, S^2) = 0] = 1$.

定理 4.3.23. X_1, \dots, X_n ($n \geq 2$) をある分布 $P \in P(\mathbb{R})$ の独立標本とする.

(1) 標本平均による特徴付け: $\bar{X} \sim N(\mu, \sigma^2/n)$ ならば, $P = N(\mu, \sigma^2)$.

- (2) 標本分散による特徴付け: P は正の有限な分散を持つ対称な分布とする. $S^2/\sigma^2 \sim \chi^2(n-1)$ ならば, $P = N(\mu, \sigma^2)$.
- (3) 非心の場合: 任意の実数 $a_1, \dots, a_n \in \mathbb{R}$ について, $\sum_{i \in [n]} (X_i + a_i)^2$ の分布が $\sum_{i \in [n]} a_i^2$ の値だけに依存するならば, $\exists \sigma^2 > 0$ $P = N(0, \sigma^2)$.
- (4) 標本平均と標本分散の独立性 (Kawata-Sakamoto): $\bar{X} \perp\!\!\!\perp S^2$ ならば, P は正規分布である.
- (5) 制約を強めた一般化 (Laha 1956): P は 2 次の積率を持つとし, $A = (a_{ij})$ は $\text{Tr}(A) \neq 0$, $\sum_{i,j=1}^n a_{ij} = 0$ を満たすとする. このとき, 次は同値.
- (a) $\bar{X} \perp\!\!\!\perp Q := X^\top A X$.
- (b) P は正規で, $\forall_{j \in [n]} \sum_{i=1}^n a_{ij} = 0$.
- (6) P を非退化とする. 一般に, $\forall_{\{a_i\}_{i \in [n]} \subset \mathbb{R} \setminus \{0\}}$ $L_1 = \sum_{i \in [n]} a_i X_i \sim P$ ならば $\sum_{i=1}^n a_i^2 \leq 1$ である. ただし, 等号に強めたものも成立する場合は P は正規である.

4.3.8 二次形式の分布

(非心率 0 の) 自由度 k の χ^2 -分布 $\chi^2(k)$ の特性関数は $\varphi(u) = (1 - 2iu)^{-k/2}$ と表せる.

定理 4.3.24 (正規確率変数の二次形式). $A \in M_n(\mathbb{R})$ が対称行列, $X \sim N(0, \Sigma)$ のとき, $Y := X^\top A X$ の積率母関数と特性関数は

$$\mathcal{M}_Y(u) = \det I - 2\theta A \Sigma^{-1/2} = \prod_{\lambda \in \text{Sp}(A\Sigma)} (1 - 2\theta\lambda)^{-1/2}, \quad \varphi_Y(u) = \exp(\mathcal{M}_Y(iu)).$$

と表せる.

要諦 4.3.25. これが χ^2 -分布の積率母関数と一致するためには, $\lambda \in 2$ が必要十分であり, 1 の数だけ χ^2 -分布の自由度が増える.

系 4.3.26 (χ^2 -分布の特徴付け). 対称行列 $A \in M_n(\mathbb{R})$, $X \sim N(0, \Sigma)$ と $Y := X^\top A X$ について, 次は同値:

- (1) $Y \sim \chi^2(k)$.
- (2) $\text{rank } A = k$ かつ $(\Sigma_{1/2})^\top A \Sigma_{1/2}$ は冪等である.

ただし, $\Sigma_{1/2}$ とは三角平方根, すなわち, Cholesky 分解 $\Sigma_{1/2} = \Sigma_{1/2} \Sigma_{1/2}^\top$ を与える正実数を対角成分にもつ下三角行列である.

命題 4.3.27 (非心 χ^2 -分布の特徴付け). 対称行列 $A \in M_n(\mathbb{R})$, $X \sim N(\mu, \Sigma)$ と $Y := X^\top A X$ について, 次は同値:

- (1) $Y \sim \chi^2(k, \mu^\top A \mu)$.
- (2) $\text{rank } A = k$ かつ $(\Sigma_{1/2})^\top A \Sigma_{1/2}$ は冪等である.

4.3.9 微分方程式

補題 4.3.28 (Stein の補題). $X \in L^1(\Omega)$ について, 次は同値:

- (1) $X \sim N(0, 1)$.
- (2) 任意の $f \in C_b^1(\mathbb{R})$ について, $E[f'(X) - f(X)X] = 0$.

[証明].

(1) \Rightarrow (2) f, f' はいずれも有界としたから, $E[f'(X)], E[f(X)] < \infty$ に注意. 部分積分により,

$$\begin{aligned} E[f'(X)] &= - \int_{\mathbb{R}} f(x) \phi'(x) dx \\ &= \int_{\mathbb{R}} f(x) x \phi(x) dx = E[f(X)X]. \end{aligned}$$

(2)⇒(1) X は可積分としたから、特性関数 $\varphi(uy) := E[e^{iuX}]$ の微分は

$$\varphi'(u) = iE[Xe^{iuX}] = E[-ue^{iuX}] = -u\varphi(u).$$

と計算できる。ただし、 $f(X) := ie^{iuX}$ とみて $f'(X) = -ue^{iuX}$ であることを用いた。すると、この微分方程式を規格化条件 $\varphi(0) = 1$ の下で解くと、 $\varphi(u) = e^{-\frac{u^2}{2}}$ 。

■

4.3.10 混合正規分布

正規な密度関数の凸結合で得る密度関数が定める分布を混合正規分布という。

定義 4.3.29 . 密度関数

$$p(x, \theta) := (1-t)\phi(x; \mu_1, \sigma_1^2) + t\phi(x; \mu_2, \sigma_2^2) \quad \theta := (t, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2), t \in (0, 1)$$

で定まる分布を混合正規分布という。

4.4 一次元絶対連続分布の例

4.4.1 一様分布

定義 4.4.1 (uniform distribution). $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の一様分布 $U((a, b))$ ($a < b \in \mathbb{R}$) とは、確率密度関数 f

$$f(x) := \frac{1}{b-a} 1_{(a,b)}$$

が定める確率分布をいう。

命題 4.4.2 .

(1) 特性関数は

$$\varphi(u) = \frac{e^{ibu} - e^{iau}}{(b-a)iu}.$$

連続延長によれば、確かに $\varphi(0) = 1$ である。

$$(2) \alpha_1 = \frac{a+b}{2}.$$

$$(3) \mu_2 = \frac{(b-a)^2}{12}.$$

$$(4) \gamma_1 = 0, \gamma_2 = \frac{9}{5}.$$

[証明].

$$(1) E[e^{iux}] = \int_a^b e^{iux} \frac{1}{b-a} dx \text{ なので.}$$

(2)

$$\alpha_1 = \int_a^b x \frac{b-a}{dx} = \frac{1}{b-a} \left[\frac{x^2}{2} \right]_a^b = \frac{a+b}{2}.$$

(3)

$$\alpha_2 = \int_a^b x^2 \frac{b-a}{dx} = \frac{1}{b-a} \left[\frac{x^3}{3} \right]_a^b = \frac{a^2 + ab + b^2}{3}.$$

より,

$$\mu_2 = \alpha_2 - \alpha_1^2 = \frac{(a-b)^2}{12}.$$

(4) 同様にしてさらに高階の積率を計算することができる。

■

4.4.2 Gamma 関数の性質

定義 4.4.3. 右平面 $R := \{z \in \mathbb{C} \mid \operatorname{Im} z > 0\}$ 上の関数 $\Gamma : R \rightarrow \mathbb{C}$ を

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

と定める.

定理 4.4.4 (有理型関数への延長と反転公式).

- (1) 上の定義で, 右半平面 R 上で一様収束し, 正則関数を定める. 項別微分を考えると R 上で関数等式 $\Gamma(z+1) = z\Gamma(z)$ が成り立つ. 特に $\Gamma(n) = (n-1)!$.
- (2) \mathbb{C} 上に有理型関数として延長し, 極は $\Gamma^{-1}(\infty) = \{0, -1, -2, \dots\}$ でそれぞれ 1 位.
- (3) (反転公式) \mathbb{C} 上の有理型関数としての等式

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}$$

が成り立つ. 特に, 逆 $\frac{1}{\Gamma(z)} = \pi\Gamma(1-z)\sin \pi z$ は整関数である. また, ここから Gauss 積分 $\Gamma(1/2)^2 = \pi$ と $\Gamma(n/2) = \frac{(n-1)!}{2^{n-1}}\sqrt{\pi}$ を得る.

定理 4.4.5 (Stirling's formula).

- (1) 部分積分を繰り返すことにより, 帰納的に

$$\int_0^\infty x^n e^{-x} dx = n!.$$

- (2) Laplace の原理から,

$$\int_0^\infty e^{n \log ny - ny} \sim \sqrt{\frac{2\pi}{n}} e^{-n}.$$

- (3) (Stirling)

$$\frac{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n}{n!} \xrightarrow{n \rightarrow \infty} 1.$$

[証明]. n-Cat lab^{†2}

- (1) 部分積分による.
- (2) $f(y) = y - \log y$ に関する Laplace の大偏差原理??から,

$$\int_0^\infty e^{n \log y - ny} dy = \int_0^\infty e^{-n(y - \log y)} dy \sim \sqrt{\frac{2\pi}{n}} e^{-n}.$$

- (3) 次の計算による:

$$\begin{aligned} n! &= \int_0^\infty x^n e^{-x} dx = \int_0^\infty e^{n \log x - x} \\ &= n \int_0^\infty \exp(n \log(ny) - ny) dy \\ &= ne^{n \log n} \int_0^\infty e^{n \log y - ny} dy \sim ne^{n \log n} \sqrt{\frac{2\pi}{n}} e^{-n} = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n. \end{aligned}$$

■

定理 4.4.6 (Gamma 関数の実関数としての特徴付け (Bohr-Mollerup)). $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ が次の 3 条件を満たすならば, Gamma 関数である:

- (1) 階乗性: $f(x+1) = xf(x)$.

^{†2} https://golem.ph.utexas.edu/category/2021/10/stirlings_formula.html

- (2) 正規化: $f(1) = 1$.
 (3) 凸性: $\log f$ は凸関数である.

[証明]. [Rudin, 1976] Th'm 8.19. ■

4.4.3 Gamma 分布

パラメータの取り方はさまざまな流儀があるが, Gamma 関数の被積分関数 $g(x; 1, z) = x^{z-1}e^{-x}$ とその相似拡大 $g(x; \alpha, z) := \alpha g(\alpha x; 1, z)$ を Gamma 分布という. 簡単な変数変換により特性関数

$$\varphi(u) = \left(\frac{\alpha}{\alpha - iu} \right)^v$$

を得て, v に関する再生性が判る. これは, 指数分布と χ^2 -分布を統一的に扱う枠組みになる.

定義 4.4.7 (gamma distribution). $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 上の **Gamma 分布** $\text{Gamma}(\alpha, v)$ ($\alpha, v > 0$) とは, 確率密度関数

$$f(x) = g(x; \alpha, v) := \frac{1}{\Gamma(v)} \alpha^v x^{v-1} e^{-\alpha x} 1_{\{x>0\}}$$

が定める分布をいう. 実際, $t = \alpha x$ と変数変換すると,

$$\int_0^\infty \alpha^v x^{v-1} e^{-\alpha x} dx = \alpha^v \int_0^\infty \left(\frac{t}{\alpha} \right)^{v-1} e^{-t} \frac{dt}{\alpha} = \int_0^\infty t^{v-1} e^{-t} dt = \Gamma(v).$$

命題 4.4.8 (Gamma 分布の特性値).

- (1) 特性関数: $\varphi(u) = \frac{1}{(1 - \frac{i u}{\alpha})^v}$.
 (2) 積率母関数: $M(t) = \frac{1}{(1 - \frac{t}{\alpha})^v}$ で, 定義域は $t \in (-\infty, \alpha)$.
 (3) 平均: $\alpha_1 = \frac{v}{\alpha}$.
 (4) 分散: $\mu_2 = \frac{v}{\alpha^2}$.
 (5) 尖度と歪度: $\gamma_1 = \frac{2}{\sqrt{v}}, \gamma_2 = 3 + \frac{6}{v}$.

[証明].

- (1) 被積分関数を Gamma 関数の被積分関数に直すために, 変数変換 $z := x(\alpha - iu)$ を考えると, $dz = (\alpha - iu)dx$ で, γ を半直線 $\mathbb{R}_+ \rightarrow \mathbb{C}; t \mapsto t(\alpha - iu)$ とすると,

$$\begin{aligned} \varphi(u) &= \int_0^\infty e^{iux} \frac{1}{\Gamma(v)} \alpha^v x^{v-1} e^{-\alpha x} dx \\ &= \int_0^\infty \frac{1}{\Gamma(v)} \alpha^v x^{v-1} e^{-\alpha(1 - \frac{i u}{\alpha})x} dx \\ &= \frac{\alpha^v}{\Gamma(v)} \int_\gamma \left(\frac{z}{\alpha - iu} \right)^{v-1} e^{-z} \frac{dz}{\alpha - iu} \\ &= \left(\frac{\alpha}{\alpha - iu} \right)^v \frac{1}{\Gamma(v)} \int_\gamma z^{v-1} e^{-z} dz. \end{aligned}$$

γ と正の実軸 \mathbb{R}_+ と半径 R の円周とが囲む扇型領域の境界上の周回積分を, 特異点 $0 \in \mathbb{C}$ を迂回するように修正し, $R \rightarrow \infty, \epsilon \rightarrow 0$ の極限を考えると, 扇型の弧上での積分は e^{-z} の急減少性により 0 に収束する. よって, γ 上での積分は \mathbb{R}_+ 上での積分と値が等しく, 結論を得る.

(2)

$$\varphi'(u) = -v \left(1 - \frac{i u}{\alpha} \right)^{-v-1} \left(-\frac{i}{\alpha} \right)$$

より, $\alpha_1 = \varphi'(u)/i = \frac{v}{\alpha}$ である.

(3)

$$\varphi''(u) = -v(-v-1) \left(1 - \frac{iu}{\alpha}\right)^{-v-2} \left(-\frac{1}{\alpha^2}\right)$$

より, $\alpha_2 = -\varphi'(u) = \frac{v(v+1)}{\alpha^2}$. よって,

$$\mu_2 = \alpha_2 - \alpha_1^2 = \frac{v(v+1) - v^2}{\alpha^2} = \frac{v}{\alpha^2}.$$

(4) 同様にして, 高次の積率も計算できる.

命題 4.4.9 (Gamma 分布の再生性). Gamma 分布は, スケール変数 α を揃えた際に, $v > 0$ について再生性を持つ.

[証明]. $\text{Gamma}(\alpha, v_1) * \text{Gamma}(\alpha, v_2)$ の特性関数は

$$\left(\frac{\alpha}{\alpha - iu}\right)^{v_1} \left(\frac{\alpha}{\alpha - iu}\right)^{v_2} = \left(\frac{\alpha}{\alpha - iu}\right)^{v_1 + v_2}$$

であるため.

4.4.4 指数分布

指数分布は, 指数関数的減衰を確率分布用に規格化したもの $g(x; \gamma) = \gamma e^{-\gamma x}$ である. 無記憶性を持つ唯一の連続分布である.

定義 4.4.10 (exponential distribution). $\text{Exp}(\gamma) := \text{Gamma}(\gamma, 1)$ ($\gamma \in \mathbb{R}^+$) を母数 γ の**指数分布**という.

命題 4.4.11 (特性値). 指数分布 $X \sim \text{Exp}(\gamma)$ の,

(1) 確率密度関数は,

$$g(x; \gamma, 1) = \frac{1}{\Gamma(1)} \gamma^1 x^0 e^{-\gamma x} 1_{\{x>0\}} = \frac{\gamma}{e^{\gamma x}} 1_{\{x>0\}}.$$

(2) 特性関数は, $\varphi(u) = \frac{\gamma}{\gamma - iu}$.

(3) 積率母関数は $M(t) = \left(1 - \frac{t}{\gamma}\right)^{-1}$ で, 定義域は $\text{Dom}(M) = (-\infty, \gamma)$.

(4) $E[X] = 1/\gamma, \text{Var}[X] = 1/\gamma^2$.

(5) $\gamma_1 = 2, \gamma_2 = 9$.

[証明].

(1) 定義から明らか.

(2) 次のように計算できる:

$$\varphi(u) = \int_0^\infty e^{iux} \lambda e^{-\lambda x} dx = \mu \int_0^\infty e^{(iu-\lambda)x} dx = \mu \left[\frac{e^{(iu-\lambda)x}}{iu-\lambda} \right]_0^\infty = \frac{\lambda}{\lambda - iu}.$$

命題 4.4.12 (指数分布の生存関数). $\text{Exp}(\gamma)$ の確率密度関数を $g(x)$, 分布関数を $F(x)$ とすると, 次の関係がある:

$$F(x) = \int_0^x \gamma e^{-\gamma t} dt = -e^{-\gamma x} + 1 = 1 - \frac{g(x)}{\gamma}.$$

生存関数を $S(x) := \frac{g(x)}{\gamma} = e^{-\gamma x}$ とおくと, $\gamma = 1$ の標準指数分布のとき $F = S$ が成り立つ. よって, 指数分布の故障率は γ で一定である.

定義 4.4.13 (survival function, hazard function / failure rate function). 一般に, $T : \Omega \rightarrow \mathbb{R}_+$ を寿命を表す絶対連続確率変数, g をその確率密度関数とする.

- (1) T の分布を**寿命分布**といい, その分布関数を F とするとき, $S(x) := 1 - F(x) = P[X \geq x]$ を**生存割合関数**という.
- (2) $h(t) := \frac{g(t)}{S(t)} = -\frac{d}{dt} \log S(t)$ を**故障率関数**という. このとき,

$$\forall s, t \in \mathbb{R}_+ \quad P[T > s + t | T > s] = \exp \left(- \int_s^{s+t} h(u) du \right).$$

故障率は局所的に指数分布と見做したときのその母数に一致する.

[証明]. 故障率関数の定義式である微分方程式を $S(0) = 1$ に注意して解くと,

$$\log S(s) = - \int_0^s h(t) dt.$$

よって,

$$P[T > s + t | T > s] = \frac{S(s+t)}{S(s)} = \frac{\exp \left(- \int_0^{s+t} h(u) du \right)}{\exp \left(- \int_0^s h(u) du \right)} = \exp \left(- \int_s^{s+t} h(u) du \right).$$

■

命題 4.4.14 (指数分布の無記憶性による特徴付け).

- (1) 指数分布は無記憶性を持つ: $\forall s, t \in \mathbb{R}_+ \quad P[T > s + t | T > s] = P[T > t]$.
- (2) 無記憶性 $\forall s, t \in \mathbb{R} \quad S(s+t) = S(s)S(t)$ を持つ連続分布は指数分布に限る. すなわち, 指数分布は, 故障率が一定の連続分布として特徴付けられる.

[証明].

- (1) 任意の連続分布は, 故障率関数と生存関数についての上の関係式を満たすから, 後は

$$P[T > t] = \exp \left(- \int_s^{s+t} \gamma du \right) = e^{-\gamma t}$$

を示せば良いが, これは指数分布の性質 $P[T > t] = S(t) = e^{-\gamma t}$ である.

- (2) $S(s+t) = S(s)S(t)$ の両辺の $t = 0$ での微分係数を考えると,

$$S'(s) = g(0)S(s).$$

これを解いて, $\int_{\mathbb{R}_+} g(t) dt = 1$ の条件と併せると, $S(s) = e^{-g(0)s}$ となる. これは母数 $g(0)$ の指数分布の生存関数である.

■

4.4.5 カイ 2 乗分布

自由度 k のカイ 2 乗分布は特性関数

$$\varphi(u) = \frac{1}{(\sqrt{1-2iu})^k}$$

によって定まる. $k = 1/2$ の場合が $X \sim N(0, 1)$ の自乗 X^2 が従う分布であることに注意すれば, これは k 個の標準正規確率変数の独立和が従う分布である.

定義 4.4.15 (chi-square distribution).

- (1) スケールを $1/2$ に固定し, 位置母数を半整数とした Gamma 分布 $\chi^2(k) := \text{Gamma} \left(\frac{1}{2}, \frac{k}{2} \right)$ を (非心率 0 で) 自由度 k のカイ 2 乗分布という.

(2) 確率密度関数は,

$$\chi^2(x; k) := g(x; 1/2, k/2) = \frac{1}{\Gamma\left(\frac{k}{2}\right) 2^{k/2}} x^{\frac{k}{2}-1} e^{-\frac{x}{2}} \mathbf{1}_{\{x>0\}}.$$

(3) 特性関数は

$$\varphi(u) = \left(\frac{1}{1-2iu} \right)^{k/2} = \frac{1}{(\sqrt{1-2iu})^k}.$$

(4) 非心率 $\delta \in \mathbb{R}_+$ の非心 $\chi^2(k, \delta)$ -分布とは, 確率密度関数

$$\chi^2(x; k, \delta) = \exp\left(-\frac{\delta}{2}\right) \sum_{r=0}^{\infty} \frac{1}{r!} \left(\frac{\delta}{2}\right)^r g\left(x; \frac{1}{2}, r + \frac{k}{2}\right)$$

で定まる分布をいう. 分散は等しいが平均がバラバラで, 和が $\delta = \sum_i \mu_i$ となるような正規確率変数の自乗和はこれに従う.

要諦 4.4.16. $\delta = 0, k = 2$ のとき, 確率密度関数は

$$p_{\chi^2_2}(x) = \frac{1}{2} e^{-\frac{x}{2}} \mathbf{1}_{\{x>0\}}$$

で表され, これは指数分布 $p(x) = \lambda e^{-\lambda x} \mathbf{1}_{\{x>0\}}$ ($\lambda > 0$) の $\text{Exp}(1/2) = \text{Gamma}(1/2, 1)$ に等しい.

命題 4.4.17 (特性値). 自由度 k , 非心率 δ の χ^2 -分布 $\chi^2(k, \delta)$ について,

(1) 特性関数:

$$\varphi(u) = (1-2iu)^{-\frac{k}{2}} \exp\left(\frac{\delta i u}{1-2iu}\right).$$

非心率が 0 の χ^2 -分布では, 後ろの指数部分 $\exp\left(\frac{\delta i u}{1-2iu}\right)$ が 1 になり消える.

(2) $\delta = 0$ のとき, 平均と分散は: $\alpha_1 = n, \mu_2 = 2n$.

[証明].

(1) $\delta = 0$ のときは, Gamma 分布の特別な場合である.

(2) 直接計算

$$\begin{aligned} \alpha_1 &= \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} x x^{\frac{n}{2}-1} e^{-\frac{x}{2}} dx \\ &= \frac{2}{\Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} \frac{1}{2^{\frac{n}{2}+1}} x^{\frac{n}{2}} e^{-\frac{x}{2}} dx \\ &= \frac{2\Gamma\left(\frac{n}{2}+1\right)}{\Gamma\left(\frac{n}{2}\right)} = 2\frac{n}{2} = n. \end{aligned}$$

同様にして,

$$\begin{aligned} \alpha_2 &= \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} x^2 x^{\frac{n}{2}-1} e^{-\frac{x}{2}} dx \\ &= \frac{2^2}{\Gamma\left(\frac{n}{2}\right)} \int_0^{\infty} 2^{-\frac{n}{2}-2} x^{\frac{n}{2}+1} e^{-\frac{x}{2}} dx \\ &= \frac{2^2 \Gamma\left(\frac{n}{2}+2\right)}{\Gamma\left(\frac{n}{2}\right)} = 4\frac{n}{2} \left(\frac{n}{2}+1\right) = n(n+2). \end{aligned}$$

よって,

$$\mu_2 = \alpha_2 - \alpha_1^2 = n(n+2) - n^2 = 2n.$$

定理 4.4.18 (χ^2 -分布は再生性を持つ). $X_1 \sim N(\mu, 1), X_j \sim N(0, 1)$ ($j = 2, 3, \dots, k$) は独立であるとする. このとき,

$$Y := \sum_{j=1}^k X_j^2 \sim \chi^2(k, \mu^2).$$

[証明]. まずは $\mu = 0$ のときを考える.

(1) $n = 1$ のときは, X_1^2 の分布関数は

$$F_{X_1^2}(x) = P[|X_1| < \sqrt{x}] = \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{x}} e^{-\frac{u^2}{2}} du. \quad (x > 0).$$

これを微分することで,

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

(2) 一般の n のときは, 畳み込みにより, 特性関数は

$$\varphi_Y(t) = \frac{1}{(1 - 2it)^{n/2}}.$$

系 4.4.19 (分散は均一だが平均がバラバラな正規確率変数の二乗和は非心 χ^2 -分布に従う).

(1) $X_j \sim N(\mu_j, 1)$ ($j \in [k]$) は独立とする. このとき, $\delta := \sum_{j \in [k]} \mu_j^2$ として, $\sum_{j \in [k]} X_j^2 \sim \chi^2(k, \delta)$.

(2) $X_j \sim \chi^2(k_j, \delta_j)$ ($j \in [k]$) は独立とする. このとき,

$$\sum_{j \in [k]} X_j \sim \chi^2\left(\sum_{j \in [k]} k_j, \sum_{j \in [k]} \delta_j\right).$$

[証明]. 次の定理の系である.

定理 4.4.20. 一般に, $X \sim N_d(\mu, I_d)$, $P \in M_d(\mathbb{R})$ を階数 $r \leq d$ の直交射影行列とすると, $X^\top P X \sim \chi^2(r; \mu^\top P \mu)$.

[証明]. P はある直交行列 Q を用いて, $Q^\top P Q = \Lambda_r := \text{diag}(1, \dots, 1, 0, \dots, 0)$ と表せる. よってこのとき, $\mu^\top P \mu = \sum_{i \in [r]} (Q\mu)_i^2$ と表せている. よって, $D := QX \sim N_d(Q\mu, I_d)$ とすれば,

$$X^\top P X \stackrel{d}{=} D^\top \Lambda_r D = \sum_{i \in [r]} D_i^2 \sim \chi^2(r; \mu^\top P \mu).$$

4.4.6 Fisher-Cochran の定理

定理 4.4.21 (Fisher-Cochran). $X \sim N_d(\mu, I_d)$ とし, 対称行列 $P_1, \dots, P_n \in M_d(\mathbb{R})$ は I_d の分解であるとする. $Q_i := X^\top P_i X$, $d_i := \text{rank } P_i$ ($i \in [n]$) について, 次は同値:

- (1) $\exists \delta_i \in \mathbb{R}_+$ $Q_i \sim \chi^2(d_i, \delta_i)$ かつ Q_1, \dots, Q_n は独立.
- (2) $\sum_{i \in [n]} d_i = d$.
- (3) P_i は射影行列である: $\forall i \in [n] \ P_i^2 = P_i$.
- (4) $\forall i \neq j \in [n] \ P_i P_j = O$.
- (5) $\mathbb{R}^d = \bigoplus_{i \in [n]} \text{Im } P_i$.

また, これらが成り立つとき, $\delta_i = \mu^\top P_i \mu$ である.

[証明]. (2),(3),(4),(5) が同値であることは, ひとえに有限次元線型空間上の冪等自己準同型の性質による.

(2) \Rightarrow (1) $d_i = 0$ のとき $P_i = O$ より, $X^\top P_i X \sim \delta_0 = \chi^2(0)$. このときたしかに $\delta_i = \mu^\top P_i \mu = 0$ である. よって, 以降 $d_i > 0$ と仮定する.

Step1 P_i は冪等だから固有値は $0, 1$ のみで、さらに対称だからある直交行列 O_i が存在して

$$O_i P_i O_i^\top = D_i := \text{diag}(\underbrace{1, \dots, 1}_{d_i}, 0, \dots, 0).$$

と対角化される.

Step2 $Y_i := O_i P_i X$ とおくとこれは再び多変量正規分布に従い、共分散が

$$\begin{aligned} \text{Cov}[Y_i, Y_j] &= \text{Cov}[O_i P_i X, O_j P_j X] = O_i P_i \text{Cov}[X, X] P_j^\top O_j^\top \\ &= O_i P_i I_d P_j^\top O_j^\top = \Lambda_{d_i} \delta_{ij}. \end{aligned}$$

と表せることより、 Y_1, \dots, Y_n は独立である. よって、 $Q_i = Y_i^\top \Lambda_{d_i} Y_i$ も互いに独立.

Step3 また、 Y_i の平均ベクトルは

$$E[Y_i] = O_i P_i \mu = \Lambda_{d_i} O_i \mu$$

より、 Y_i は $N_d(\Lambda_{d_i} O_i \mu, I_d)$ に従う. Y_i の零でない成分は d_i 個に限ることに注意すれば、定理 4.4.20 より、 $Q_i = Y_i^\top \Lambda_{d_i} Y_i \sim \chi^2(d_i, \delta_i)$ で、

$$\delta_i = (O_i \mu)^\top \Lambda_{d_i} (O_i \mu) = \mu^\top P_i \mu.$$

(1) \Rightarrow (2) 2通りで $X^\top X$ の分布を記述し、それを比べる.

1 通り目

$$X^\top X = X^\top (P_1 + \dots + P_n) X = \sum_{i \in [n]} Q_i$$

より、 χ^2 -分布の再生性から、 $X^\top X \sim \chi^2\left(\sum_{i \in [n]} d_i, \sum_{i \in [n]} \delta_i\right)$.

2 通り目 $X^\top X = \sum_{i \in [d]} X_i^2$ より、 $X^\top X \sim \chi^2(d, \mu^\top \mu)$ 4.4.20.

よって、 $d = \sum_{i \in [n]} d_i$.

■

系 4.4.22. $A_1, A_2 \in M_d(\mathbb{R})$ を対称行列、 $Q_i := X^\top A_i X$ を二次形式とする.

(1) Q_1 が非心 χ^2 -分布に従うことと、 A_1 が冪等であることは同値. このとき、 $Q_1 \sim \chi^2(\mu^\top A_1 \mu, \text{rank } A_1)$.

(2) Q_1, Q_2 が共に非心 χ^2 -分布に従うとする. $Q_1 \perp Q_2$ と $Q_1 Q_2 = O$ は同値.

4.4.7 Student 分布

F 分布に従う確率変数の平方根、または正規確率変数を χ^2 -確率変数の平方根で割ったものは、 t -分布に従う.

定義 4.4.23 (t -distribution (Gosset 1908)). $Y \sim \chi^2(n), Z \sim N(\delta, 1)$ は独立とし、確率変数 $X := \frac{Z}{\sqrt{Y/n}}$ を考える.

(1) $\delta = 0$ のとき、 X の従う分布を **Student の t -分布**といい、 $t(n)$ で表す.

(2) $\delta > 0$ のとき、 X の従う分布を **非心 t -分布**といい、 $t(n, \delta)$ で表す.

命題 4.4.24 (密度関数と特性関数).

(1) 密度関数は、 $\delta = 0$ のとき、

$$p_{n,0}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{\pi n} \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} = \frac{1}{\sqrt{n} B\left(\frac{1}{2}, \frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}.$$

一般の $\delta > 0$ について、

$$p_{n,\delta}(x) = \frac{1}{\sqrt{\pi n}} e^{-\frac{\delta^2}{2}} \sum_{r \in \mathbb{N}} \frac{2^{r/2}}{r!} \frac{\Gamma\left(\frac{n+r+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \left(\frac{\delta x}{\sqrt{n}}\right)^r \left(1 + \frac{x^2}{n}\right)^{-\frac{n+r+1}{2}}.$$

(2) 特性関数は、第2種修正 Bessel 関数 K_ν について、

$$\varphi(x) = \frac{n^{\frac{n-2}{4}}}{2^{\frac{n}{2}-1}\Gamma(\frac{n}{2})} t^{\frac{n}{2}} K_{\frac{n}{2}}(\sqrt{nt}).$$

[証明].

(1) 変換

$$\begin{cases} X_1 = \frac{Z}{\sqrt{Y/n}} \\ X_2 = Z \end{cases} \Leftrightarrow \begin{pmatrix} Y \\ Z \end{pmatrix} = \begin{pmatrix} X_1^2 n \\ X_2 \end{pmatrix} = T \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$$

を考えると、 $A := \mathbb{R}^+ \times \mathbb{R}, B := \mathbb{R}^+ \times \mathbb{R}$ について、 $T: A \rightarrow B$ は可微分同相で、

$$DT(x_1, x_2) = \begin{pmatrix} -2n \frac{x_2^2}{x_1^3} & 2n \frac{x_2}{x_1^2} \\ 0 & 1 \end{pmatrix}, \quad J_T(x_1, x_2) = -2n \frac{x_2^2}{x_1^3}$$

より、 A 上で Jacobian は消えない。よって、結合密度関数は

$$\begin{aligned} p^*(x_1, x_2) &= p_Y(y)p_Z(z)|J_T(x_1, x_2)| \\ &= \left(\frac{1}{\Gamma(\frac{n}{2}) 2^{\frac{n}{2}}} y^{\frac{n}{2}-1} e^{-\frac{y}{2}} \right) \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \right) 2n \frac{x_2^2}{x_1^3} \\ &= \frac{1}{\Gamma(\frac{n}{2}) 2^{\frac{n}{2}} \sqrt{2\pi}} 2n n^{\frac{n}{2}-1} \left(\frac{x_2}{x_1} \right)^{n-2} e^{-\frac{n}{2} \frac{x_2^2}{x_1^2}} e^{-\frac{x_2^2}{2} \frac{x_2^2}{x_1^3}}. \end{aligned}$$

より、 x_1 の周辺密度関数は、積分と変数変換 $y = x_2^2$ により、 $dy = 2x_2 dx_2$ に注意して、

$$\begin{aligned} p^*(x_1) &= \frac{1}{\Gamma(\frac{n}{2}) 2^{\frac{n}{2}} \sqrt{2\pi}} n^{\frac{n}{2}} \frac{1}{x_1^{n+1}} 2 \int_{\mathbb{R}} \underbrace{x_2^n}_{=x_2^{n-1}x_2} e^{-\frac{x_2^2}{2} \left(1 + \frac{n}{x_1^2}\right)} dx_2 \\ &= \frac{1}{\Gamma(\frac{n}{2}) 2^{\frac{n}{2}} \sqrt{2\pi} \sqrt{n}} \left(\frac{n}{x_1^2} \right)^{\frac{n+1}{2}} \int_0^\infty y^{\frac{n-1}{2}} e^{-\frac{y}{2} \left(1 + \frac{n}{x_1^2}\right)} dy \\ &= \frac{1}{\Gamma(\frac{n}{2}) 2^{\frac{n}{2}} \sqrt{2\pi} \sqrt{n}} \left(\frac{n}{x_1^2} \right)^{\frac{n+1}{2}} 2^{\frac{n+1}{2}} \left(1 + \frac{n}{x_1^2}\right)^{-\frac{n+1}{2}} \Gamma\left(\frac{n+1}{2}\right) \\ &= \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right) \sqrt{\pi} \sqrt{n}} \left(1 + \frac{x_1^2}{n}\right)^{-\frac{n+1}{2}}. \end{aligned}$$

注 4.4.25.

$$f_{t(1)}(x) = \frac{1}{\sqrt{\pi}\sqrt{\pi}} (1+x^2)^{-1}$$

より、 $t(1) = \text{Cauchy}(0, 1)$ である。

命題 4.4.26 (逆 Gamma 分布の正規混合の構造). $Y \sim \chi^2(n) = \text{Gamma}\left(\frac{1}{2}, \frac{n}{2}\right), Z \sim N(0, 1)$ とする。

$$(1) \frac{Y}{n} \sim \text{Gamma}\left(\frac{n}{2}, \frac{n}{2}\right).$$

$$(2) \tilde{Y} := \left(\frac{Y}{n}\right)^{-1} \sim \text{Gamma}^{-1}\left(\frac{n}{2}, \frac{n}{2}\right) \text{ で、密度関数は}$$

$$f(x) = \frac{\left(\frac{n}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} x^{-\frac{n}{2}-1} e^{-\frac{n}{2x}} \quad (x > 0).$$

(3) $X = \sqrt{\tilde{Y}}Z$ の密度関数は

$$p(t) = \frac{\left(\frac{n}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} x^{-\frac{n}{2}-1} e^{-\frac{n}{2x}}$$

と表せる。

[証明].

- (1) 変数変換 $y = x/n$ により, 係数部分が $\alpha \mapsto \alpha/n$ と変わる.
 (2) 命題 4.4.70 で議論した.
 (3) $\sqrt{\tilde{Y}} \perp\!\!\!\perp Z$ より, $X|\tilde{Y} = y \sim N(0, y)$ から,

$$\begin{aligned} p(t) &= \int_0^\infty f_{X|Y}(t|y) f_Y(y) dy \\ &= \int_0^\infty \frac{1}{\sqrt{2\pi y}} e^{-\frac{t^2}{2y}} f(y) dy. \end{aligned}$$

系 4.4.27 (t-分布の漸近的性質).

$$t(n) \Rightarrow N(0, 1) \quad (n \rightarrow \infty).$$

[証明]. t 分布の密度関数

$$p_{n,0}(x) = \frac{1}{\sqrt{n} B\left(\frac{1}{2}, \frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}.$$

に注目する.

- (1) Beta 関数の漸近的性質 4.4.39(4) より,

$$\frac{1}{\sqrt{n} B\left(\frac{1}{2}, \frac{n}{2}\right)} \xrightarrow{n \rightarrow \infty} \sqrt{2}$$

- (2) また,

$$\left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}} = \left(1 - \frac{-x^2/2}{n/2}\right)^{-\frac{n}{2} + \frac{1}{2}} \xrightarrow{n \rightarrow \infty} e^{-\frac{x^2}{2}}.$$

より, 密度関数は $N(0, 1)$ のものに各点収束する.

命題 4.4.28 (積率の存在). $X \sim t(\nu)$ ($\nu > 0$) について,

- (1) $p \in (-1, \nu)$ ならば, $E[|X|^p] < \infty$.
 (2) $E[|X|^p] = \nu^{p/2} \frac{\Gamma\left(\frac{\nu-p}{2}\right) \Gamma\left(\frac{p+1}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{\nu}{2}\right)}.$

応用例 4.4.29. 正規母集団 $N(\mu, \sigma^2)$ からの標本 X_1, \dots, X_n を通じて, 母平均 μ の推定を考えたいとする. 標本平均 \bar{X} と不偏分散 U^2 について,

$$t(X_1, \dots, X_n | \mu) := \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

と定めると, これは $t(n-1)$ に従う. 特に, σ^2 に一切依っていないことに注意.

4.4.8 合流型超幾何関数**定義 4.4.30 (Bessel function of first kind).**

- (1) $\alpha \in \mathbb{R}$ に対して, $t^2 x'' + tx' + (t^2 - \alpha^2)x = 0$ と表される $t = 0$ に確定特異点を持つ方程式の解

$$J_{\pm\alpha}(t) = \sum_{k \in \mathbb{N}} \frac{(-1)^k}{k! \Gamma(\pm\alpha + k + 1)} \left(\frac{t}{2}\right)^{\pm\alpha + 2k}$$

を **第1種 Bessel 関数** という. $\alpha \notin \mathbb{N}$ ならば $J_{\pm\alpha}$ は線形独立で, この線形結合が一般解を与える.

- (2)

$$Y_\nu(t) := \frac{J_\nu(t) \cos \nu\pi - J_{-\nu}(t)}{\sin \nu\pi}$$

を **Weber 関数** または **第2種 Bessel 関数** という. $\nu \in \mathbb{C}$ について定まり, Bessel 方程式の解であり, $J_{\pm m}$ と線形独立になる.

(3) 次の2つを**修正 Bessel 関数**, 後者を特に **Macdonald 関数**ともいう:

$$I_\alpha(x) = i^{-\alpha} J_\alpha(ix), \quad K_\alpha(x) = \frac{\pi}{2} \frac{I_{-\alpha}(x) - I_\alpha(x)}{\sin \alpha\pi}.$$

これらは, 修正された Bessel 方程式 $t^2 x'' + tx' - (t^2 + \alpha^2)x = 0$ の2つの線型独立な解である.

定義 4.4.31 (generalized hypergeometric function, confluent hypergeometric function).

- (1) 関数 ${}_rF_s(\alpha_1, \dots, \alpha_r; \beta_1, \dots, \beta_s; z) = \sum_{k \in \mathbb{N}} \frac{(\alpha_1)_k \cdots (\alpha_s)_k}{(\beta_1)_k \cdots (\beta_r)_k} \frac{z^k}{k!}$ を**一般化超幾何関数**という.
 (2) 関数 $M(\alpha, \beta; z) := {}_1F_1(\alpha, \beta; z) = \sum_{k \in \mathbb{N}} \frac{(\alpha)_k}{(\beta)_k} \frac{z^k}{k!}$ を**Kummer の合流型超幾何関数**という.

命題 4.4.32 .

- (1) Kummer の合流型超幾何関数は,
 (a) Kummer の微分方程式 $tx'' + (\beta - t)x' - \alpha x = 0$ の解の1つである.
 (b) 極限 $M(\alpha, \gamma; z) = \lim_{\beta \rightarrow \infty} F(\alpha, \beta, \gamma; z/\beta)$ である.
 (c) 積分表示 $M(\alpha, \beta; z) = \frac{\Gamma(\beta)}{\Gamma(\alpha)\Gamma(\beta - \alpha)} \int_0^1 e^{zu} u^{\alpha-1} (1-u)^{\beta-\alpha-1} du$ を持つ. 特に, $M(\alpha, \alpha + \beta; it)$ は Beta 分布の特性関数である.
 (d) α が負の整数のとき, Laguerre の陪多項式の定数倍 $L_n^{(\alpha)}(x) = \binom{n+\alpha}{n} M(-n, \alpha+1; x)$ となる.
 (2) 次も Kummer の微分方程式の線型独立な基本解で, **第二種 Kummer 関数**または **Tricomi の合流型超幾何関数**という:

$$U(\alpha, \beta; z) = \frac{\Gamma(1-\beta)}{\Gamma(\alpha+1-\beta)} M(\alpha, \beta; z) + \frac{\Gamma(\beta-1)}{\Gamma(\alpha)} z^{1-\beta} M(\alpha+1-\beta, 2-\beta; z).$$

- (3) 一般化超幾何関数は, ${}_1F_1$ のとき Kummer の合流型超幾何関数であるが,
 (a) ${}_0F_0(z) = e^z$ は指数関数になる.
 (b) ${}_1F_0(\alpha; z) = (1-z)^{-\alpha}$ は代数関数になる.
 (c) ${}_0F_1(\beta; z)$ は本質的に Bessel 関数になる. 第1種 Bessel 関数に対して,

$$J_\alpha(x) = \frac{(x/2)^\beta}{\Gamma(\beta+1)} {}_0F_1(\beta+1; -x^2/4).$$

$${}_0F_1(1/2; -z^2/4) = \cos z \text{ になる.}$$

4.4.9 Fisher-Snedecor 分布

t -確率変数の自乗, そして独立な χ^2 -確率変数の商は, F -確率変数である.

定義 4.4.33 (Fisher-Snedecor distribution / F-distribution). $Y_1 \sim \chi^2(m, \delta), Y_2 \sim \chi^2(n)$ は独立とする. このとき,

$$X := \frac{Y_1/m}{Y_2/n}$$

の分布を,

- (1) $\delta = 0$ のとき, **Fisher-Snedecor 分布**といい, $F(m, n) := \text{Beta}\left(\frac{m}{2}, \frac{n}{2}\right)$ で表す.
 (2) $\delta > 0$ のとき, **非心 F-分布**といい, $F(m, n, \delta)$ で表す.

命題 4.4.34 (F-分布の特性値).

- (1) 確率密度関数は, $\delta = 0$ のとき,

$$p_{m,n,0}(x) = \frac{1}{B\left(\frac{m}{2}, \frac{n}{2}\right)} \left(\frac{\frac{m}{n}x}{1+\frac{m}{n}x}\right)^{\frac{m}{2}} \left(\frac{1}{1+\frac{m}{n}x}\right)^{\frac{n}{2}} \frac{1}{x} 1_{\mathbb{R}^+} = \frac{1}{B\left(\frac{m}{2}, \frac{n}{2}\right)} \left(\frac{mx}{n+mx}\right)^{\frac{m}{2}} \left(1 - \frac{mx}{n+mx}\right)^{\frac{n}{2}} \frac{1}{x} 1_{\mathbb{R}^+}$$

一般の $\delta > 0$ については,

$$p_{m,n,\delta}(x) = \sum_{r=0}^{\infty} e^{-\frac{1}{2}\delta} \left(\frac{\delta}{2}\right)^r \frac{(m/n)^{m/2+r}}{r! B(m/2+r, n/2)} \frac{x^{m/2+r-1}}{(1+mx/n)^{(m+n)/2+r}} 1_{\mathbb{R}^+}.$$

(2) 中心の場合と非心の場合の確率密度関数の間には,

$$M(\alpha, \gamma; z) = \frac{\Gamma(\gamma)}{\Gamma(\alpha)} \sum_{n \in \mathbb{N}} \frac{\Gamma(\alpha+n)}{\Gamma(\gamma+n)} \frac{z^n}{n!}$$

を Kummer の合流型超幾何関数として,

$$p_{m,n,\delta}(x) = e^{-\frac{1}{2}\delta} M\left(\frac{m+n}{2}, \frac{m}{2}; \frac{\delta}{2} \frac{mx/n}{1+mx/n}\right) p_{m,n,0}(x)$$

という関係がある.

(3) 特性関数は, 第二種 Kummer 合流超幾何関数

$$U(a, c; z) := \frac{1}{\Gamma(a)} \int_0^{\infty} e^{-zs} s^{a-1} (1+s)^{c-a-1} ds$$

を用いて,

$$\varphi_{m,n}(u) = \frac{\Gamma(\frac{n+m}{2})}{\Gamma(\frac{n}{2})} U\left(\frac{m}{2}, 1 - \frac{n}{2}; -\frac{n}{m} i s\right).$$

命題 4.4.35 (F-分布に関連する分布).

- (1) $X \sim F(n, m)$ のとき, $X^{-1} \sim F(m, n)$.
- (2) $X \sim t(n)$ のとき, $X^2 \sim F(1, n)$.
- (3) $X \sim \text{Beta}(m/2, n/2)$ のとき, $\frac{nX}{m(1-X)} \sim F(m, n)$.
- (4) $X_1 \sim \Gamma(\alpha_1, \beta_1), X_2 \sim \Gamma(\alpha_2, \beta_2)$ を独立とする. このとき, $\frac{\beta_1/\alpha_1 \cdot X_1}{\beta_2/\alpha_2 \cdot X_2} \sim F(2\alpha_1, 2\alpha_2)$.

歴史 4.4.36. $X \sim F(m, n)$ のとき, $\frac{\log X}{2}$ を Fisher の z-分布と呼ぶ.

命題 4.4.37 (Hotelling T^2 -統計量). $X_1, \dots, X_n \sim N_d(\mu_0, \Sigma)$ とし, $U_n := \frac{1}{n-1} \sum_{i \in [n]} (X_i - \bar{X}_n)(X_i - \bar{X}_n)^\top$ を不偏共分散行列とする. このとき,

$$T_n^2(\mu) := n(\bar{X}_n - \mu)^\top U_n^{-1}(\bar{X}_n - \mu)$$

について,

- (1) $\delta_n := n(\mu_0 - \mu)^\top \Sigma^{-1}(\mu_0 - \mu)$ を非心率として,

$$\frac{n-d}{d(n-1)} T_n^2(\mu) \sim F_{d, n-d}(\delta_n).$$

- (2) $T_n^2(\mu_0) \Rightarrow \chi^2(d)$.

4.4.10 Beta 関数

Beta 関数には

$$B(x+1, y) = B(y, x+1) = \frac{x}{x+y} B(x, y)$$

の関係があり, Gamma 関数との関係

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

によって理解される. x について単調減少であることを思えば, 上下を間違えないだろう.

定義 4.4.38 (Beta function). $B: \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}$ を

$$B(x, y) := \int_0^1 t^{x-1} (1-t)^{y-1} dt, \quad (x, y > 0)$$

で定める。これは $\operatorname{Im} x, \operatorname{Im} y > 0$ を満たす半平面上に解析接続される。

定理 4.4.39 (Beta 関数の性質). 次の成り立つ：

(1) 対称性： $B(x, y) = B(y, x)$.

(2) Gamma 関数との関係：

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

(3) 積分表示：

$$B(x, y) = 2 \int_0^{\pi/2} (\sin \theta)^{2x-1} (\cos \theta)^{2y-1} d\theta = \int_0^\infty \frac{u^{x-1}}{(1+u)^{x+y}} du.$$

(4) Gamma 関数との漸近的關係：固定された y について、 $B(x, y) \sim \Gamma(y)x^{-y}$ ($x \rightarrow \infty$).

なお、いずれの等式も $\operatorname{Im} x, \operatorname{Im} y > 0$ にて成り立つ。

[証明].

(1) 変数変換 $s := 1 - t$ による：

$$B(y, x) = \int_0^1 t^{y-1} (1-t)^{x-1} dt = - \int_1^0 (1-s)^{y-1} s^{x-1} ds = B(x, y).$$

(2) Step1 $B(1, y) = \frac{1}{y}$ である。実際、

$$B(1, y) = \int_0^1 t^{y-1} dt = \left[\frac{t^y}{y} \right]_0^1 = \frac{1}{y}.$$

Step2 $x \mapsto \log B(x, y)$ は凸関数である。実際、任意の $p \in (1, \infty)$ とその共役指数 q について、凸結合 $\frac{1}{p} + \frac{1}{q} = 1$ を考えると、Holder の不等式より、

$$\begin{aligned} B\left(\frac{x}{p} + \frac{y}{q}, z\right) &= \int_0^1 t^{\frac{x}{p} + \frac{y}{q} - 1} (1-t)^{z-1} dt = \int_0^1 t^{\frac{x-1}{p} + \frac{y-1}{q}} (1-t)^{(z-1)(\frac{1}{p} + \frac{1}{q})} dt \\ &\leq \left(\int_0^1 \left(t^{\frac{x-1}{p}} (1-t)^{\frac{z-1}{p}} \right)^p dt \right)^{1/p} \left(\int_0^1 \left(t^{\frac{y-1}{q}} (1-t)^{\frac{z-1}{q}} \right)^q dt \right)^{1/q} = B(x, z)^{1/p} B(y, z)^{1/q}. \end{aligned}$$

両辺の対数を取ると、

$$\log B\left(\frac{x}{p} + \frac{y}{q}, z\right) \leq \frac{\log B(x, z)}{p} + \frac{\log B(y, z)}{q}.$$

を得る。

Step3 次の等式が成り立つ：

$$B(x+1, y) = \frac{x}{x+y} B(x, y).$$

実際、部分積分より、

$$\begin{aligned} B(x+1, y) &= \int_0^1 t^x (1-t)^{y-1} dt \\ &= \int_0^1 \left(\frac{t}{1-t} \right)^x (1-t)^{x+y-1} dt = -\frac{1}{x+y} \int_0^1 \left(\frac{t}{1-t} \right)^x \left((1-t)^{x+y} \right)' dt \\ &= -\frac{1}{x+y} \left[\left(\frac{t}{1-t} \right)^x (1-t)^{x+y} \right]_0^1 + \frac{1}{x+y} \int_0^1 x \left(\frac{t}{1-t} \right)^{x-1} \frac{(1-t) + t}{(1-t)^2} (1-t)^{x+y} dt \\ &= -\frac{1}{x+y} \left[t^x (1-t)^y \right]_0^1 + \frac{x}{x+y} \int_0^1 t^{x-1} (1-t)^{y-1} dt = \frac{x}{x+y} B(x, y). \end{aligned}$$

Step4 任意の $y > 0$ について,

$$f(x) := \frac{\Gamma(x+y)}{\Gamma(y)} B(x, y)$$

は $\log f$ は凸であり, $f(1) = 1$ を満たし, さらに

$$f(x+1) = \frac{\Gamma(x+y+1)}{\Gamma(y)} B(x+1, y) = x f(x)$$

が成り立つ. よって, Gamma 関数の特徴付け 4.4.6 より, $f(x) = \Gamma(x)$ である.

(3) $t =: \sin \theta$ の変数変換による.

4.4.11 Beta 分布

定義 4.4.40 (beta distribution). $((0, 1), \mathcal{B}((0, 1)))$ 上の (第1種) ベータ分布 $\text{Beta}(\alpha, \beta)$ ($\alpha, \beta \in \mathbb{R}_{>0}$) とは, 確率密度関数

$$f(x; \alpha, \beta) := \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} 1_{(0,1)}(x)$$

が定める確率分布をいう. ただし, $B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx$.

命題 4.4.41 (Beta 分布の特性値). $\text{Beta}(\alpha, \beta)$ について,

- (1) $\alpha_1 = \frac{\alpha}{\alpha + \beta}$.
- (2) $\mu_2 = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}$.
- (3) $\gamma_1 = \frac{2(\beta - \alpha)\sqrt{\alpha + \beta + 1}}{(\alpha + \beta + 2)\sqrt{\alpha\beta}}$.
- (4) $\gamma_2 = \frac{3(\alpha + \beta + 1)(\alpha^2\beta + \alpha\beta^2 + 2\alpha^2 - 2\alpha\beta + 2\beta^2)}{\alpha\beta(\alpha + \beta + 2)(\alpha + \beta + 3)}$.
- (5) 特性関数は, $M(\alpha, \beta; z) = {}_1F_1(\alpha, \beta; z) = \sum_{k \in \mathbb{N}} \frac{(\alpha)_k z^k}{(\beta)_k k!}$ を Kummer の合流型超幾何関数として,

$$\varphi(u) = M(\alpha, \alpha + \beta; iu).$$

[証明].

(1)

$$\begin{aligned} \alpha_1 &= \int_0^1 x \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} dx \\ &= \frac{B(\alpha+1, \beta)}{B(\alpha, \beta)} \underbrace{\int_0^1 \frac{1}{B(\alpha+1, \beta)} x^{\alpha} (1-x)^{\beta-1} dx}_{=1} \\ &= \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+1+\beta)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} = \frac{\alpha}{\alpha + \beta}. \end{aligned}$$

命題 4.4.42 (積率). $\text{Beta}(\alpha, \beta)$ について,

(1) 中心積率:

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

この等式は任意の実数 $r > -\alpha$ について成り立つ.

(2) $\alpha > 1$ のとき, $E[X^{-1}] = \frac{\alpha + \beta - 1}{\alpha - 1}$.

4.4.12 第2種 Beta 分布

定義 4.4.43. 次の確率密度関数で定まる分布 $\text{Beta}_2: \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow P(\mathbb{R})$ を第2種 Beta 分布という:

$$f(x; \alpha, \beta) := \frac{1}{B(\alpha, \beta)} \frac{x^{\alpha-1}}{(1+x)^{\alpha+\beta}}$$

命題 4.4.44. $X \sim \text{Beta}(\alpha, \beta)$ のとき, $\frac{X}{1-X} \sim \text{Beta}_2(\alpha, \beta)$.

4.4.13 Cauchy 分布

特性関数 $\varphi(u) = \exp(i\mu u - \sigma|u|)$ で定まる分布である. $u = 0$ で微分可能でない, 1 次の積率 (平均) さえも存在しない.

定義 4.4.45 (Cauchy distribution). $(\mathcal{X}, \mathcal{B}_1)$ 上の **Cauchy 分布** $C(\mu, \sigma)$ ($\mu \in \mathbb{R}, \sigma \in \mathbb{R}_{>0}$) とは, 確率密度関数

$$f(x) = \frac{1}{\pi\sigma} \frac{1}{1 + \left(\frac{x-\mu}{\sigma}\right)^2}$$

が定める分布をいう. μ は位置母数, σ は尺度母数と解釈される. $C(0, 1)$ の場合の確率密度は \mathbb{R} 上の Poisson 核に一致する. そして σ は総和核の添字 $P_\sigma(x) := \frac{1}{\sigma} P\left(\frac{x}{\sigma}\right) = \frac{1}{\pi\sigma} \frac{1}{1 + \left(\frac{x}{\sigma}\right)^2}$ である.

命題 4.4.46.

- (1) 裾が重く ($1/x^2$ のオーダー), 平均と分散は存在しない.
- (2) $\varphi(u) = \exp(i\mu u - \sigma|u|)$. 明らかに $u = 0$ で微分可能でない.
- (3) 分布関数は $F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan(x)$.

[証明].

- (1) $X \sim C(0, 1)$ のとき,

$$E[|X|] = \frac{1}{\pi} \int_{\mathbb{R}} \frac{1}{1+x^2} x dx = \infty.$$

- (2) $\mu = C(0, 1)$ のとき $\check{\mu}(u) = e^{-|u|}$ であることを2通りで示す.

天下りの方法

$$\begin{aligned} \mathcal{F}[e^{-|u|}](x) &= \int_{\mathbb{R}} e^{-|u|} e^{-iux} du \\ &= \int_0^\infty e^{-u(1+ix)} du + \int_{-\infty}^0 e^{u(1-ix)} du \\ &= \left[-\frac{e^{-u(1+ix)}}{1+ix} \right]_0^\infty + \left[\frac{e^{u(1-ix)}}{1-ix} \right]_{-\infty}^0 \\ &= \frac{1}{1+ix} + \frac{1}{1-ix} = \frac{2}{1+x^2}. \end{aligned}$$

より, 反転公式から,

$$2\pi e^{-|u|} = \mathcal{F}\left[\frac{2}{1+x^2}\right](u).$$

複素積分を直接計算

$$\varphi(u) = \frac{1}{\pi} \int_{-\infty}^\infty \frac{e^{ixu}}{1+x^2} dx = \frac{1}{2\pi i} \int_{\mathbb{R}} \left(\frac{e^{ixu}}{x-i} - \frac{e^{ixu}}{x+i} \right) dx.$$

の計算であるが, $u > 0$ のとき, 上半円 $B_R(0) \cap \mathbb{R}_+$ の境界上での積分を考えると, これは $u = i$ での留数に等しいことから,

$$\varphi(u) = \frac{1}{2\pi i} \int_{\partial B_\epsilon(i)} \frac{e^{ixu}}{x-i} dx = e^{iu} = e^{-u}.$$

歴史 4.4.47. Cauchy(1853) によると考えられていたが, Poisson が 1824 年にすでに注目していた.

4.4.14 Weibull 分布

Weibull 分布は寿命分布として用いられ, また極値分布としても現れる.

定義 4.4.48 (Weibull distribution). $(\mathcal{X}, \mathcal{G}_1)$ 上の **Weibull 分布** $W(\nu, \alpha)$ ($\nu, \alpha \in \mathbb{R}_{>0}$) とは, 確率密度関数

$$f(x) = \frac{\nu}{\alpha} \left(\frac{x}{\alpha}\right)^{\nu-1} \exp\left(-\left(\frac{x}{\alpha}\right)^\nu\right) 1_{x>0} = \int_{-\infty}^x f(y)dy$$

が定める分布である. α が尺度母数, ν が形状母数と解釈される.

命題 4.4.49.

- (1) 分布関数は $F(x) = \left\{1 - \exp\left(-\left(\frac{x}{\alpha}\right)^\nu\right)\right\} 1_{x>0}$.
- (2) $\alpha_1 = \alpha \Gamma\left(\frac{\nu+1}{\nu}\right)$.
- (3) $\mu_2 = \alpha^2 \left\{\Gamma\left(\frac{\nu+2}{\nu}\right) - \Gamma\left(\frac{\nu+1}{\nu}\right)^2\right\}$.

命題 4.4.50. 故障率関数を

$$h(t) = \left(\frac{\alpha}{\beta}\right) \left(\frac{t}{\beta}\right)^{\alpha-1} 1_{\{t>0\}}$$

としたとき, 故障時刻 T は Weibull 分布 $W(\alpha, \beta)$ にしたかう.

命題 4.4.51. $X \sim \text{Exp}\left(\frac{1}{\alpha^\nu}\right)$ ならば, $X^{\frac{1}{\nu}} \sim W(\nu, \alpha)$.

4.4.15 対数正規分布

定義 4.4.52 (log-normal distribution). $(\mathcal{X}, \mathcal{G}_1)$ 上の**対数正規分布** $L_N(\nu, \sigma)$ ($\nu \in \mathbb{R}, \sigma \in \mathbb{R}_{>0}$) とは, 確率密度関数

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{1}{x} \exp\left\{-\frac{(\log x - \mu)^2}{2\sigma^2}\right\} 1_{x>0}$$

が定める分布である.

命題 4.4.53.

$$\log X \sim N(\mu, \sigma^2) \Leftrightarrow X \sim L_N(\mu, \sigma)$$

命題 4.4.54.

- (1) $\alpha_1 = \exp\left(\mu + \frac{\sigma^2}{2}\right)$.
- (2) $\mu_2 = (e^{\sigma^2} - 1) \exp(2\mu + \sigma^2)$.
- (3) $\gamma_1 = (\omega + 2)\sqrt{\omega - 1}$. ただし, $\omega := \exp(\sigma^2)$ とした.
- (4) $\gamma_2 = \omega^4 + 2\omega^3 + 3\omega^2 - 3$.

4.4.16 Zeta 関数

命題 4.4.55 (Zeta 関数の定義).

$$\zeta(z) := \sum_{n=1}^{\infty} n^{-z} \quad (z \in \mathbb{R} + 1)$$

とする. $\mathbb{R} + 1 = \{\text{Re } z > 1\}$ 上で一様収束し, 正則関数を定める.

定理 4.4.56 (Euler 積表示). 素数を $2 = p_1 < p_2 < \cdots$ とする. このとき,

$$\zeta(z) \prod_{n=1}^{\infty} \left(1 - \frac{1}{p_n^z}\right) = 1 \text{ on } R+1.$$

特に, ζ は $R+1$ 上に零点を持たない. また, $\sum_{n \in \mathbb{N}} p_n^{-1} = \infty$ で, 素数は無限個ある.

要諦 4.4.57. このことは, $p_n \sim n^\alpha$ ($\alpha > 1$) の分布に従わないことを示唆する (さもなくば収束する). 実際には, $p_n \sim n \log n$ であることが, ζ の $\partial(R+1)$ での挙動から知られている.

補題 4.4.58. 命題 (1) の右辺の積分に関連して, 次の等式が成り立つ:

$$\frac{\zeta(z)}{\Gamma(1-z)} = \frac{1}{2\pi i} \int_C \frac{w^{-z-1}}{e^w - 1} dw = 2(2\pi)^{z-1} \sin(z\pi/2) \sum_{m=1}^{\infty} m^{z-1} = 2(2\pi)^{z-1} \sin(z\pi/2) \zeta(1-z) \text{ on } z \in \mathbb{C}$$

ただし, $C := \partial\{w \in \mathbb{C} \mid \text{dist}(w, \mathbb{R}_+) < \epsilon\}$ とした. 特に, 左半平面 $-R$ には, $\zeta(0) = 1/2, \zeta(-2m) = 0$ 以外の零点はない.

命題 4.4.59 (解析接続による有理型関数への延長).

(1) Gamma 関数と Zeta 関数について, 次の等式が成り立つ:

$$\zeta(z)\Gamma(z) = \int_0^\infty \sum_{n=1}^\infty e^{-ny} y^{z-1} dy = \int_0^\infty \frac{y^{z-1}}{e^y - 1} dy =: I(z) \text{ on } R+1.$$

(2) 命題 (1) の右辺の積分に関連して, 次の等式が成り立つ:

$$\frac{\zeta(z)}{\Gamma(1-z)} = \frac{1}{2\pi i} \int_C \frac{w^{-z-1}}{e^w - 1} dw = 2(2\pi)^{z-1} \sin(z\pi/2) \sum_{m=1}^\infty m^{z-1} = 2(2\pi)^{z-1} \sin(z\pi/2) \zeta(1-z) \text{ on } z \in \mathbb{C}$$

ただし, $C := \partial\{w \in \mathbb{C} \mid \text{dist}(w, \mathbb{R}_+) < \epsilon\}$ とした.

(3) 特に, \mathbb{C} の全域で有理型, $\mathbb{C} \setminus \{1\}$ 上で正則な関数

$$\zeta(z) = 2(2\pi)^{z-1} \sin(\pi z/2) \Gamma(1-z) \zeta(1-z)$$

に延長され, また左半平面 $-R$ には, $\zeta(0) = 1/2, \zeta(-2m) = 0$ 以外の零点はない.

要諦 4.4.60. こうして, $-R$ と $R+1$ における零点には協力的な研究道具がある. しかし, 帯領域 $0 \leq \text{Re } z \leq 1$ における零点の研究は困難を極める. Wiener による Tauber 型の定理と $\text{Re } s = 1$ 上で零点が存在しないことは同値である.

系 4.4.61. n 以下の素数の個数を $\pi(n)$ とすると,

$$\lim_{n \rightarrow \infty} \frac{\pi(n)}{n/\log n} = \lim_{n \rightarrow \infty} \frac{p_n}{n \log n} = 1.$$

4.4.17 logistic 分布

定義 4.4.62 (logistic distribution). $(\mathcal{X}, \mathcal{G}_1)$ 上のロジスティック分布 $\text{Log}(\mu, \sigma)$ ($\mu \in \mathbb{R}, \sigma \in \mathbb{R}_{>0}$) とは, 分布関数

$$F(x) = \frac{1}{1 + \exp\left(-\frac{x-\mu}{\sigma}\right)}$$

で, 確率密度関数

$$f(x) = \frac{1}{\sigma} \frac{\exp\left(-\frac{x-\mu}{\sigma}\right)}{\left\{1 + \exp\left(-\frac{x-\mu}{\sigma}\right)\right\}^2}$$

が与える分布をいう. よって, 分布は $x = \mu$ に関して対称である.

補題 4.4.63.

$$I(r) := \int_0^\infty x^r \frac{e^{-x}}{(1+e^{-x})^2} dx = (1 - 2^{-(r+1)}) \Gamma(r+1) \zeta(r) \quad (r > 1)$$

[証明].

$$\begin{aligned}
I(r) &= \int_0^\infty x^r \frac{e^{-x}}{(1+e^{-x})^2} dx = \int_0^\infty \frac{rx^{r-1}}{1+e^x} dx \\
&= \int_0^\infty rx^{r-1} e^{-x} \sum_{j=0}^\infty (-1)^j e^{-jx} dx \\
&= \Gamma(r+1) \sum_{n=1}^\infty (-1)^{n-1} \frac{1}{n^r} \quad (r > 0)
\end{aligned}$$

ここで, $\zeta(r) = \sum_{n=1}^\infty \frac{1}{n^r}$ ($r > 1$) について

$$(1 - 2^{-(r-1)})\zeta(r) = \sum_{n=1}^\infty (-1)^{n-1} n^{-r}$$

が成り立つから, 最後の変形を得る. ■

命題 4.4.64.

(1) 中心積率は, 奇数について $\mu_r = 0$, 2以上の偶数について

$$\mu_r = 2\sigma^r \Gamma(r+1) \sum_{n=1}^\infty (-1)^{n-1} \frac{1}{n^r} = 2\sigma^r (1 - 2^{-(r-1)}) \Gamma(r+1) \zeta(r).$$

$$(2) \alpha_1 = \mu, m_2 = \frac{\pi^2 \sigma^2}{3}, \mu_4 = \frac{7\pi^4 \sigma^4}{15}.$$

$$(3) \gamma_1 = 0, \gamma_2 = 4.2.$$

$$(4) \varphi(u) = e^{i\mu u} \frac{\pi \sigma u}{\sinh(\pi \sigma u)}.$$

[証明].

(1) 奇数の時は対称性より.

$$(2) \zeta(2) = \frac{\pi^2}{6}, \zeta(4) = \frac{\pi^4}{90} \text{ より.}$$

(3) $-R, R, -R + 2\pi i, R + 2\pi i$ を頂点にもつ長方形についての留数計算より. ■

4.4.18 Pareto 分布

パレート分布は所得の分布に当てはまるという.

定義 4.4.65. $(\mathcal{X}, \mathcal{G}_1)$ 上のパレート分布 $P_A(b, a)$ ($a, b \in \mathbb{R}_{>0}$) とは, 分布関数を

$$F(x) = 1 - \left(\frac{b}{x}\right)^a 1_{x \geq b}$$

確率密度関数を

$$f(x) := ab^a x^{-(a+1)} 1_{x \geq b}$$

とする分布をいう.

命題 4.4.66.

$$(1) \alpha_1 = ab(a-1)^{-1} \quad (a > 1).$$

$$(2) \mu_2 = ab^2(a-1)^{-2}(a-2)^{-1} \quad (a > 2).$$

$$(3) \gamma_1 = 2 \frac{a+1}{a-3} \sqrt{\frac{a-2}{a}} \quad (a > 3).$$

$$(4) \gamma_2 = \frac{3(a-2)(3a^2+a+2)}{a(a-3)(a-4)} \quad (a > 4).$$

4.4.19 逆正規分布

ドリフト付きの Brown 運動の到達時間の分布として現れる．キウムラント母関数 $\log M$ が，正規分布のキウムラント母関数の逆数になっていることから．

定義 4.4.67 (inverse Gaussian distribution / Wald distribution). 確率密度関数

$$f(x; \delta, \gamma) = 1_{(0, \infty)}(x) \frac{\delta e^{\gamma \delta}}{\sqrt{2\pi}} x^{-3/2} \exp\left(-\frac{1}{2} \left(\gamma^2 x + \frac{\delta^2}{x}\right)\right)$$

で定まる確率分布 $\text{IG}(\delta, \gamma) : \mathbb{R}_{>0} \times \mathbb{R}_{\geq 0} \rightarrow P(\mathbb{R})$ を **逆正規分布** または **Wald 分布** という．

補題 4.4.68 .

(1) $\gamma > 0$ のとき，確率密度関数は

$$f(x) = 1_{(0, \infty)}(x) \left(\frac{\delta^2}{2\pi}\right)^{1/2} x^{-3/2} \exp\left(-\frac{\delta^2(x - \delta\gamma^{-1})^2}{2(\delta\gamma^{-1})^2 x}\right)$$

とも表せる．

(2) $\gamma = 0$ のとき，確率密度関数は， $c := \delta^2$ と定めて，

$$f(x; c) = 1_{(0, \infty)}(x) \sqrt{\frac{c}{2\pi}} x^{-3/2} \exp\left(-\frac{c}{2x}\right)$$

と表せる．これが定める分布を **Levy 分布** といい， $\text{Levy}(c) := \text{IG}(c^{1/2}, 0)$ と表す．

(3) $\text{IG}(\delta, \gamma)$ の積率母関数は，

$$\begin{aligned} M(s) &= \int_0^\infty e^{sx} f(x) dx \\ &= \exp\left(\delta\left(\gamma - \sqrt{\gamma^2 - 2s}\right)\right) & s \leq \frac{\gamma^2}{2}, \gamma \geq 0 \\ &= \exp\left(\gamma\delta\left(1 - \sqrt{1 - \frac{2s}{\gamma^2}}\right)\right) & s \leq \frac{\gamma^2}{2}, \gamma > 0. \end{aligned}$$

(4) $\text{IG}(\delta, \gamma)$ の特性関数は $\varphi(u) = \exp\left(\gamma\delta\left(1 - \sqrt{1 - \frac{2iu}{\gamma^2}}\right)\right)$ となる．

(5) キウムラント母関数は $\log \varphi(u) = \frac{\delta}{\gamma} iu + \sum_{r=2}^\infty (2r-3)(2r-5) \cdots 1 \cdot \frac{\delta}{\gamma^{2r-1}} \cdot \frac{(iu)^r}{r!}$ で，キウムラントは

$$\kappa_1 = \alpha_1 = \frac{\delta}{\gamma}, \quad \kappa_r = (2r-3)(2r-5) \cdots 1 \cdot \frac{\delta}{\gamma^{2r-1}}$$

(6) Levy 分布 $\text{Levy}(c)$ の Laplace 変換は

$$\int_0^\infty e^{-\lambda x} f(x) dx = e^{-\sqrt{c\lambda}} \quad (\lambda \in \mathbb{R}_+)$$

である．

4.4.20 逆 Gamma 分布

定義 4.4.69 . $X \sim \text{Gamma}(\alpha, \nu)$ とし，逆数 X^{-1} が従う分布を $\text{Gamma}^{-1}(\alpha, \nu)$ とする．

命題 4.4.70 (逆 Gamma 分布の密度関数). $X \sim \text{Gamma}(n/2, n/2)$ ($n \in \mathbb{N}$) とし，確率変数 $Y := 1/X$ の確率密度関数は

$$f(x) = \frac{\left(\frac{n}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} x^{-\frac{n}{2}-1} e^{-\frac{n}{2x}} \quad (x > 0).$$

と表せる．

[証明]. $y > 0$ のとき $P[Y \leq y] = P[X \geq 1/y]$ であるから, Gamma 分布の分布関数を G とすれば, $F(x) = P[X \geq 1/x] = 1 - G\left(\frac{1}{x}\right)$ より,

$$\begin{aligned} f(x) &= F'(x) = G'\left(\frac{1}{x}\right) \frac{1}{x^2} \\ &= \frac{\left(\frac{n}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} x^{-\frac{n}{2}+1} e^{-\frac{n}{2} \frac{1}{x}} \frac{1}{x^2} \\ &= \frac{\left(\frac{n}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} x^{-\frac{n}{2}-1} e^{-\frac{n}{2x}}. \end{aligned}$$

■

4.4.21 Laplace 分布

定義 4.4.71. 密度関数

$$f(x) = \frac{1}{2} e^{-|x|}$$

が定める分布を **標準 Laplace 分布 (両側指数分布)** という.

命題 4.4.72.

- (1) 特性関数は $\varphi(u) = (1 + u^2)^{-1}$.
- (2) X, Y, X', Y' が独立に標準正規分布 $N(0, 1)$ に従うとする. このとき, $V := XY + X'Y'$ は密度 f を持つ Laplace 分布に従う.

4.4.22 Marcenko-Pastur 分布

MP($\lambda, 1$) 分布は Wishart 行列の固有値の標本スペクトル分布の漸近分布として導出された.

定義 4.4.73. パラメータ $(\lambda, \sigma^2) \in \mathbb{R}_{>0}^2$ が定める定数 $a := \sigma^2(1 - \sqrt{\lambda})^2, b := \sigma^2(1 + \sqrt{\lambda})^2$ について, 確率密度関数

$$f(x; \lambda, \sigma^2) := \frac{1}{2\pi\sigma^2} \frac{\sqrt{(b-x)(x-a)}}{\lambda x} 1_{[a,b]}(x) + 1_{[1,\infty)}(\lambda) \left(1 - \frac{1}{\lambda}\right) \delta_0(x)$$

が定める分布 $\text{MP}(\lambda, \sigma^2) : \mathbb{R}_{>0}^2 \rightarrow P(\mathbb{R})$ を **Marcenko-Pastur 分布** という.

補題 4.4.74. $X \sim \text{MP}(\lambda, \sigma^2)$ ならば, $\sigma^{-2}X \sim \text{MP}(\lambda) := \text{MP}(\lambda, 1)$.

命題 4.4.75. $X_{ij} \sim \text{i.i.d.} N(0, 1)$ ($i \in [d], j \in [n]$) を成分とする $d \times n$ ランダム行列 X に対し, $n^{-1}XX^\top$ の固有値を $0 \leq \lambda_1 \leq \dots \leq \lambda_d$ とする.

- (1) 標本スペクトル分布は $\mu_{d,n}(dx) := \frac{1}{d} \sum_{i=1}^d \delta_{\lambda_i}(dx)$ で与えられる.
- (2) $d = d_n$ が $\frac{d_n}{n} \rightarrow \lambda \in (0, \infty)$ を満たすとき, $\mu_{d_n,n} \xrightarrow{d} \text{MP}(\lambda)$ を満たす.

4.4.23 Pearson 系

Pearson が現象の起こり方の背後に確率分布を想定した初めの人であり, これの列挙として Pearson 系を提案した. これが今日の確率密度関数の原形となった. この枠組を基に, Fisher が方法論を体系化した.

定義 4.4.76. \mathbb{R} 上の確率密度関数 p であって, 微分方程式

$$\frac{d}{dx} \log p(x) = \frac{(x - \lambda) - a}{b_2(x - \lambda)^2 + b_1(x - \lambda) + b_0}$$

を満たす

$$p(x) = C \exp \left(\int \frac{(x - \lambda) - a}{b_0 + b_1(x - \lambda) + b_2(x - \lambda)^2} dx \right)$$

の形をした分布を **Pearson 系** という。以降, $\lambda = 0$ とする。

(1) I 型分布とは, $p(x) = C \left(1 - \frac{x}{a_1}\right)^{m_1} \left(1 - \frac{x}{a_2}\right)^{m_2}$ ($x \in (a_1, a_2)$) の形のものをいう。ただし, $a_1 < 0 < a_2, m_1, m_2 > -1$ 。

例 4.4.77.

(1) Beta 分布は I 型 Pearson 分布である。

(2)

□

4.5 多次元絶対連続分布の例

1 次元の絶対連続分布の概念を容易に d 次元に拡張できる。

4.5.1 Wishart 分布

χ^2 -分布の多次元化である。

定義 4.5.1 (Wishart distribution). $X_1, \dots, X_n \sim N_d(\mu_j, \Sigma)$ を独立同分布とし, d 次対称行列値確率変数

$$Y := \sum_{i \in [n]} X_i X_i^\top$$

を考える。

(1) Y の分布を, 非心行列 $\Delta := \sum_{i \in [n]} \mu_i \mu_i^\top$ の **非心 Wishart 分布** といい, $W_d(n, \Sigma; \Delta)$ で表す。

(2) $\Delta = 0$ の場合は単に **Wishart 分布** といい, $W_d(n, \Sigma)$ で表す。

命題 4.5.2. 密度関数は

$$f(M) = C_{n,d}^{-1} |\Sigma|^{-\frac{d}{2}} |M|^{\frac{n-d-1}{2}} e^{-\frac{1}{2} \text{Tr}(\Sigma^{-1} M)}, \quad C_{n,d} := 2^{\frac{nd}{2}} \pi^{\frac{d(d-1)}{4}} \prod_{k=1}^d \Gamma\left(\frac{n-k+1}{2}\right), M \in M_d(\mathbb{R})^+.$$

4.5.2 Dirichlet 分布

ベータ分布の高次元化であり, 多変量ベータ分布ともいう。多項分布の共役事前分布になるため, 3 以上のカテゴリへのカウント

定義 4.5.3. $\Delta_p := \left\{ (x_1, \dots, x_p) \in \mathbb{R}_{>0}^p \mid \sum_{i=1}^p x_i < 1 \right\}$ ($p \in \mathbb{N}$) について,

$$f(x_1, \dots, x_p) = \frac{\Gamma(v)}{\prod_{i=1}^{p+1} \Gamma(v_i)} \left(\prod_{i=1}^d x_i^{v_i-1} \right) \left(1 - \sum_{i=1}^p x_i \right)^{v_{p+1}-1} 1_{\Delta_p}(x_1, \dots, x_p) \quad v_i > 0, v := \sum_{i=1}^{p+1} v_i$$

が定める分布を **Dirichlet 分布** $\text{Dirichlet}(v_1, \dots, v_{p+1}) : \mathbb{R}_{>0}^{p+1} \rightarrow P(\mathbb{R}^p)$ という。

命題 4.5.4 (周辺分布は Beta 分布である). $(X_1, \dots, X_p) \sim \text{Dirichlet}(v_1, \dots, v_{p+1})$ とする。

$$\forall_{j \in [p]} X_j \sim \text{Beta}\left(v_j, \sum_{i \neq j} v_i\right)$$

4.6 Poisson 過程に付随する分布

4.6.1 Poisson 確率変数の和と差

定義 4.6.1 . 独立な $Y \sim \text{Pois}(\alpha), Z \sim \text{Pois}(\beta)$ に対して,

- (1) $X = Y + 2Z$ の分布を, パラメータ $(\alpha, \beta) \in \mathbb{R}_{>0}^2$ の **Hermite 分布** とよび, $\text{Hermite}(\alpha, \beta)$ で表す.
- (2) $X = Y - Z$ の分布を, パラメータ $(\alpha, \beta) \in \mathbb{R}_{>0}^2$ の **Skellam 分布** とよび, $\text{Skellam}(\alpha, \beta)$ で表す.

補題 4.6.2 .

- (1) Hermite 分布の特性関数は, $\varphi(u) = \exp(\alpha(e^{iu} - 1) + \beta(e^{2iu} - 1))$.
- (2) Hermite 分布の確率母関数は $g(z) = \exp(\alpha(z - 1) + \beta(z^2 - 1))$.
- (3) Skellam 関数の特性関数は $\varphi(u) = \exp(\alpha(e^{iu} - 1) + \beta(e^{-iu} - 1))$.

4.6.2 Poisson 過程に付随する分布

命題 4.6.3 (Poisson 分布の区間的性質). (N_t) を強度 λ の定常 Poisson 過程とし, $T(x) := \inf\{t \in \mathbb{R}_+ \mid N_t \geq x + 1\}$ とする. このとき, $T(x) \sim G(\lambda, x + 1)$.

命題 4.6.4 (Poisson 過程の区間的性質). (N_t) を強度 λ の定常 Poisson 過程とする.

- (1) $T(x) := \inf\{t \in \mathbb{R}_+ \mid N_t \geq x + 1\}$ とする. このとき, $T(x) \sim \text{Gamma}(\lambda, x + 1)$.
- (2) $T := \min\{t \in \mathbb{R}_+ \mid N_t \geq 1\}$ とすると, $T \sim \text{Exp}(\lambda)$.

4.7 分布の変換による関係

命題 4.7.1 . $X \sim \text{Exp}(1)$ ならば, $e^{-X} \sim U([0, 1])$ である.

[証明].

$$P[e^{-X} \leq a] = P[X \geq -\log a] = S(\log a) = e^{\log a} = a.$$

■

第 5 章

古典数理統計学

5.1 指数型分布族

小標本理論では、統計量理論を用いるために、完備十分統計量の存在が肝要になる。これが組織的に議論出来る枠組みが指数型分布族である。

要諦 5.1.1 (連続エントロピーを最大にする分布族としての特徴付け). 一般化線型モデルでは、残差に平均と分散以外の仮定を置かないから、観測 Y は $\varphi: \mathbb{R}^p \rightarrow \mathbb{R}^n$ を通じた条件 $E[\varphi(X)] = 0$ のみ与えられている。このとき、エントロピーを最大にするような絶対連続分布は $Ae^{-\varphi(x)}$ という形の密度を持つものであり、 A の部分にパラメータが入る余地がある。なおこのとき、連続エントロピーは $h(X) = -\log A + E[\varphi(X)]$ 。

5.1.1 定義と例

指数型分布族は、有限個の x の関数 $T_i \in \text{Meas}(\mathcal{X}, \mathbb{R})$ の 1 次式の指数関数の $\text{Meas}(\mathcal{X}, \mathbb{R})$ -倍で表せる Radon-Nikodym 微分を持つクラスである。このクラスに対しては、

定義 5.1.2 (exponential family). 分布族 $\mathcal{P} = \{P_\theta\}_{\theta \in \Theta} \subset P(\mathcal{X})$ が σ -有限測度 μ に支配されており、その Radon-Nikodym 微分が

$$\frac{dP_\theta}{d\mu}(x) = g(x) \exp \left(\sum_{i=1}^m a_i(\theta) T_i(x) - \psi(\theta) \right) \quad (T_i, g \in L(\mathcal{X}), a_i, \psi: \Theta \rightarrow \mathbb{R})$$

と表せるとき、 \mathcal{P} を m -径数指数型分布族という。

要諦 5.1.3. g は測度 μ に吸収させることが出来るので、本質的には「 T_i ($i \in [m]$) を通してのみ x に依存する密度関数の族」をいう。

例 5.1.4.

(1) 正規分布の量み込みが与える族 $\{N(\theta_1, \theta_2)^{\otimes n}\}_{\theta \in \mathbb{R} \times \mathbb{R}^+}$ は

$$\begin{aligned} \frac{dP_\theta}{dx}(x) &= \prod_{j=1}^n \exp \left(\frac{\theta_1}{\theta_2} x_j - \frac{1}{2\theta_2} x_j^2 - \frac{1}{2} \left(\frac{\theta_1^2}{\theta_2^2} + \log(2\pi\theta_2) \right) \right) \\ &= \exp \left(\frac{\theta_1}{\theta_2} \sum_{j \in [n]} x_j - \frac{1}{2\theta_2} \sum_{j \in [n]} x_j^2 - \frac{n}{2} \left(\frac{\theta_1^2}{\theta_2^2} + \log(2\pi\theta_2) \right) \right). \end{aligned}$$

であるから、1 次と 2 次の標本積率 $T_1(x) := \sum_{j \in [n]} x_j, T_2(x) := \sum_{j \in [n]} x_j^2$ について指数型である。特に、 $T := (T_1, T_2)$ は完備十分統計量である。

(2) 分散 $\theta_2 \in \mathbb{R}^+$ を既知として、平均のみを母数と捉えても、 $g(x) := e^{-\frac{1}{2\theta_2} \sum_{j \in [n]} x_j^2}$ とすることで、 $T(x) := \sum_{j \in [n]} x_j$ を完備十分統計量として指数型である。

(3) 二項分布族 $\{B(n, \theta)\}_{\theta \in (0,1)}$ は、計数測度 μ に関して

$$\frac{dB(n, \theta)}{d\mu}(x) = \binom{n}{x} \exp \left(x \log \frac{\theta}{1-\theta} + n \log(1-\theta) \right)$$

と表せるから指数型である.

(4) Poisson 分布族, Gamma 分布族, Beta 分布族, 負の二項分布族, 逆正規分布族も指数型である.

(5) 独立列 $\epsilon_t \sim N(0, \sigma^2)$ ($t \in [n], \sigma^2 > 0$) と $\phi \in \mathbb{R}$ が定める確率差分方程式

$$X_t = \phi X_{t-1} + \epsilon_t, \quad X_0 = x$$

が定める確率過程 $X = (X_t)_{t \in \mathbb{N}+1} \in \mathbb{R}^{n+1}$ が押し出す分布族 $\{P_\theta\}_{\theta \in \mathbb{R} \times \mathbb{R}^+}$ は指数型である.

(6) 標準 Brown 運動 $(w_t)_{t \in [0, T]}$ が定める確率微分方程式

$$dX_t = \theta X_t dt + dw_t, \quad X_0 = x \quad (\theta \in \mathbb{R})$$

の一意な解 $X = (X_t)_{t \in [0, T]}$ が $C([0, T])$ 上に押し出す分布 $(P_\theta)_{\theta \in \mathbb{R}}$ は指数型である:

$$\frac{dP_\theta}{dP_0}(X) = \exp \left(\theta \int_0^T X_t dX_t - \frac{\theta^2}{2} \int_0^T X_t^2 dt \right).$$

□

5.1.2 指数型分布族の性質

命題 5.1.5.

(1) \mathcal{P} が指数型分布ならば、任意の 2 つの元は互いに絶対連続である.

(2) 指数型分布族が定める統計的実験の列 $(\mathcal{G}^{(i)}, \mathcal{A}^{(i)}, \mathcal{P}^{(i)})$ ($i \in [k]$) について、直積 $\mathcal{P} := \left\{ \prod_{i \in [k]} P_\theta^{(i)} \right\}_{\theta \in \Theta} \subset P \left(\bigotimes_{i \in [k]} \mathcal{G}^{(i)} \right)$ も指数型である.

(3) 可測関数 $T := (T_1, \dots, T_m) : \mathcal{X} \rightarrow \mathbb{R}^m$ は \mathcal{P} に関して十分である.

命題 5.1.6. 指数型分布族 \mathcal{P} の $T := (T_1, \dots, T_m) : \mathcal{X} \rightarrow \mathbb{R}^m$ による押し出し $(P_\theta^T)_{\theta \in \Theta}$ は再び指数型である: ある σ -有限測度 λ について $\mathcal{P}^T \ll \lambda$ で,

$$\frac{dP_\theta^T}{d\lambda}(t) = \exp \left(\sum_{i \in [m]} a_i(\theta) t_i - \psi(\theta) \right) \quad (t \in \mathbb{R}^m).$$

5.1.3 密度関数が生成する標準指数型分布族

参照測度 μ と可測関数 $g, T \in L(\mathcal{X})$ から指数型分布族が構成出来る.

定義 5.1.7 (natural parameter, natural exponential family). 可測空間 $(\mathcal{X}, \mathcal{A})$ 上の σ -有限測度 μ と 2 つの可測関数 $g \in L(\mathcal{X})_+, T \in L(\mathcal{X}; \mathbb{R}^m)$ を用いて、指数型分布族を定める.

(1) 母数の空間 $A \subset \mathbb{R}^m$ を

$$A := \left\{ a \in \mathbb{R}^m \mid \int_{\mathcal{X}} e^{a^\top T(x)} g(x) \mu(dx) \in \mathbb{R}^+ \right\}$$

と定める. これを自然母数空間という. $A = \Theta_0(g)$ と書く.

(2) $A \neq \emptyset$ とする. g のキュムラント母関数

$$\Psi(a) := \log \left(\int_{\mathcal{X}} e^{a^\top T(x)} g(x) \mu(dx) \right)$$

について、分布 $P_a \in P(\mathcal{X})$ を

$$\frac{dP_a}{d\mu}(x) := g(x) \exp(a^\top T(x) - \Psi(a))$$

と定めると、 $\{P_a\}_{a \in A}$ は指数型分布族である. これを $g d\mu$ が生成する m 次の自然指数型分布族という.

命題 5.1.8. 自然母数空間 $A \subset \mathbb{R}^m$ は凸集合である.

補題 5.1.9 (キュムラント母関数の導関数). $T: \mathcal{G} \rightarrow \mathbb{R}^m, \varphi: \mathcal{G} \rightarrow \mathbb{R}$ を測度空間 $(\mathcal{G}, \mathcal{A}, \mu)$ 上の可測関数とし, 任意の $a \in V^{\text{open}} \subset \mathbb{R}^m$ 上で φ のキュムラント母関数

$$f(a) := \int_{\mathcal{G}} e^{a^\top T(x)} \varphi(x) \mu(dx)$$

が存在するとする. このとき,

- (1) f は $D := \{z \in \mathbb{C}^m \mid \operatorname{Re} z \in V^m\}$ 上に解析接続される.
- (2) $\forall n \in \mathbb{N}^m \quad \forall z \in D \quad \frac{\partial^n f}{\partial z^n}(z) = \int_{\mathcal{G}} e^{z^\top T(x)} \varphi(x) T(x)^n \mu(dx).$

命題 5.1.10. $g d\mu$ に関する m 次の自然指数型分布族 $\{P_a\}_{a \in A} \subset P(\mathcal{G})$ について,

- (1) g のキュムラント母関数 $\Psi: \mathbb{R}^m \supset A \rightarrow \mathbb{R}$ は C^∞ -級である.
- (2) P_a に関する $T: \mathcal{G} \rightarrow \mathbb{R}^m$ の特性関数は

$$\varphi(u) = \exp(\Psi(a + iu) - \Psi(u)) \quad (u \in \mathbb{R}^m)$$

で表される.

- (3) P_a に関する $T: \mathcal{G} \rightarrow \mathbb{R}^m$ のキュムラントは $\kappa_n = \partial_a^n \Psi(a)$ ($n \in \mathbb{N}^m$) で表される.
- (4) この分布族の Fisher 情報行列は, $\left(\frac{\partial^2 \Psi(a)}{\partial a_i \partial a_j} \right)$ である.

5.1.4 指数型分布族の完備十分統計量

$(\operatorname{Im} a)^\circ \subset \mathbb{R}^m$ が空でないならば T は十分である上に完備でもある.

定理 5.1.11 (指数型分布族が完備であるための十分条件). 指数型分布族 \mathcal{P} について, $a: \Theta \rightarrow \mathbb{R}^m$ の像は疎でないならば, $T: \mathcal{G} \rightarrow \mathbb{R}^m$ は完備統計量である.

補題 5.1.12. $\mathbb{R}^m = \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$ に対応して, $T = (T_1, T_2), a = (a_1, a_2)$ と分解して考える. これらが定める自然な指数型分布族 $\{P_a\}_{a \in A} \subset P(\mathcal{G})$ について, 次を満たすような \mathbb{R}^{m_1} 上の σ -有限測度 $\nu_{t_2}(dt_1)$ ($t_2 \in \mathbb{R}^{m_2}$) が存在する:

- (1) 任意の $a \in A$ に対して次は $P_a^{T_2}$ -零集合:

$$N_{a_1} = \left\{ t_2 \in \mathbb{R}^{m_2} \mid \int_{\mathbb{R}^{m_1}} e^{a_1^\top t_1} \nu_{t_2}(dt_1) \in \{0, \infty\} \right\}.$$

- (2) 次の確率分布 $P_a(-|t_2) \in P(\mathbb{R}^{m_1})$ は, $T_2 = t_2$ の下での P_a に関する T_1 の正則条件付き確率である:

$$P_a(dt_1|t_2) = \begin{cases} \frac{e^{a_1^\top t_1}}{\int_{\mathbb{R}^{m_1}} e^{a_1^\top \tau_1} \nu_{t_2}(d\tau_1)} \nu_{t_2}(dt_1) & t_2 \in N_{a_1}^c \\ \nu_0 & \text{otherwise} \end{cases}$$

ただし, $\nu_0 \in P(\mathbb{R}^{m_1})$ は任意.

5.1.5 共役事前分布

Howard Raiffa と Robert Schlaifer によるベイジアン決定理論で作られた概念である. 共役というのは, 事後分布が, 事前分布の代数的な閉式で表せることを指している (したがって, 数値積分が必要なく, 解析的に計算可能).

定理 5.1.13 (Bayes の定理 (密度版)). 確率変数 (θ, X) は, 参照測度 $\mu(dx)\nu(d\theta)$ に関する同時密度関数 $f(x, \theta)\mu(dx)\nu(d\theta)$ をもつとし, $f_X(x), \pi(\theta), f(x|\theta)$ をそれぞれ, X, θ の周辺密度と, θ を与えた下での x の条件付き密度とする. このとき, x を与えた下での θ の条件付き密度は,

$$\forall x \in \mathcal{X} \quad f_X(x) > 0 \Rightarrow f(\theta|x) = \frac{\pi(\theta)f(x|\theta)}{\int_{\Theta} \pi(\theta)f(x|\theta)\nu(d\theta)} \quad \nu\text{-a.e.}$$

と表せる. このことを, $f(\theta|x) \propto \pi(\theta)f(x|\theta)$ と表す.

定義 5.1.14 (conjugate prior). 事後分布 $g(-|x)$ が事前分布 $\pi(-)$ と同じ分布型であるとき, π を θ の共役事前分布といい, $f(x|\theta)$ をその尤度関数という.

5.2 指数分散モデル

5.2.1 構成

記法 5.2.1. 一次元確率分布 $\mu(dx)$ が生成する 1 次元自然指数型分布族

$$\exp(x\theta - \psi(\theta))\mu(dx), \quad \theta \in \Theta_0$$

を考える. ただし, ψ は μ の (第 2) キュムラント母関数とした.

$$\Theta_0 := \left\{ \theta \in \mathbb{R} \mid \int e^{\theta x} \nu(dx) < \infty \right\}$$

は $\Theta_0^\circ \neq \emptyset$ を満たすとする. μ の積率母関数 \mathfrak{M} が存在するとし, 集合

$$\Lambda := \{ \lambda > 0 \mid \mathfrak{M}(\theta)^\lambda \text{ はある確率分布 } \mu_\lambda \text{ の積率母関数} \}$$

を考える.

補題 5.2.2 (exponential dispersion model).

- (1) 任意の $\lambda \in \Lambda$ について, $\lambda\psi(\theta)$ はある分布 μ_λ のキュムラント母関数であり, $\exp(x\theta - \lambda\psi(\theta))\mu_\lambda(dx)$ ($\theta \in \Theta_0$) は分布を定める.
- (2) (1) の分布に従う確率変数 X について, $Y := \lambda^{-1}X$ の分布は $\tilde{\mu}_\lambda := (\lambda^{-1})_*\mu_\lambda$ について次を満たす:

$$\forall_{B \in \mathcal{B}^1} \quad P^Y(B) = \int_B \exp(\lambda(y\theta - \psi(\theta)))\tilde{\mu}_\lambda(dy), \quad \theta \in \Theta_0, \lambda \in \Lambda.$$

(2) が定める分布 $(P^Y)_{\lambda \in \Lambda}$ を μ または ψ が定める **指数分散モデル**といい, $\text{EDM}(\theta, \lambda)$ で表す.

注 5.2.3. 定義から $\mathbb{N} \subset \Lambda$ である.

補題 5.2.4. $Y \sim \text{EDM}(\theta, \lambda)$ とする.

- (1) Y のキュムラント母関数は $\kappa_Y(u) = \lambda \left(\psi \left(\theta + \frac{u}{\lambda} \right) - \psi(\theta) \right)$ である. 特に, $E[Y] = \psi'(\theta)$, $\text{Var}[Y] = \lambda^{-1}\psi''(\theta)$.
- (2) 平均値関数を $\tau(\theta) := \psi'(\theta) : \Theta^\circ \rightarrow \mathbb{R}$ で, その値域を $D := \tau(\Theta^\circ)$ で表す. このとき, τ は単射である.
- (3) $V(\mu) := \tau'(\tau^{-1}(\mu)) : D \rightarrow \mathbb{R}_+$ について, $\text{Var}[Y] = \lambda^{-1}V(\mu)$ と表せる.

この $V(\mu)$ を **単位分散関数**という.

注 5.2.5. $\lambda = 1$ でない限り, $V(\mu)$ そのものは分散ではない. これより, $\text{EDM}(\theta, \lambda)$ を $\text{EDM}(\mu, \lambda)$ で表す.

補題 5.2.6. 定数 $w_i \geq 0, w := \sum_{i=1}^n w_i > 0$ と, 独立確率変数列 $Y_i \sim \text{EDM}(\mu, \lambda w_i)$ について,

$$\frac{1}{w} \sum_{i=1}^n w_i Y_i \sim \text{EDM}(\mu, \lambda w)$$

5.2.2 Tweedie 分布族

定義 5.2.7. $\text{EDM}(\mu, \lambda)$ のうち, 単位分散関数が $V(\mu) = \mu^p$ ($p \in \mathbb{R} \setminus (0, 1)$) と表せる分布族を, **Tweedie 分布族**といい, 記号 $\text{Tw}_p(\mu, \lambda)$ で表す.

例 5.2.8. $p = 0$ ならば正規分布族, $p = 1$ ならば Poisson 分布, $p = 2$ ならば Gamma 分布, $p = 3$ ならば逆正規分布となる. \square

命題 5.2.9. $\text{EDM}(\mu, \lambda)$ について, $1 \in D, V(1) = 1$ とし, さらにある関数 $s: (0, \infty) \times \Lambda^{-1} \rightarrow \Lambda^{-1}$ で以下を満たすものが存在すると仮定する: $Y \sim \text{EDM}(\mu, \lambda) \Rightarrow [\forall_{c>0} cY \sim \text{EDM}(c\mu, 1/s(c, \lambda))]$. このとき, 以下が成り立つ.

- (1) $\exists_{p \in \mathbb{R}} Y \sim \text{Tw}_p(\mu, \lambda)$.
- (2) $s(c, \lambda) = c^{2-p}/\lambda$. すなわち, $c\text{Tw}_p(\mu, \lambda) = \text{Tw}_p(c\mu, \lambda/c^{2-p})$.
- (3) $p = 0$ ならば $D = \mathbb{R}$. また, $p \neq 0$ ならば $D = (0, \infty)$. 特に, $\mathcal{L}(Y)$ は無限分解可能である.

5.3 漸近関係

第 6 章

参考文献

参考文献

- [Hazewinkel, 1995] Hazewinkel, M. (1995). Encyclopaedia of Mathematics: Volume 3 Heaps and Semi-Heaps — Moments, Method of (in Probability Theory), volume 3 of Encyclopaedia of Mathematics. Springer New York.
- [Kawata and Sakamoto, 1949] Kawata, T. and Sakamoto, H. (1949). On the characterisation of the normal population by the independence of the sample mean and the sample variance. Journal of Mathematical Society of Japan, 1(2):111–115.
- [Last and Penrose, 2017] Last, G. and Penrose, M. (2017). Lectures on the Poisson Process, volume 7 of Institute of Mathematical Statistics Textbooks. Cambridge University Press.
- [Rudin, 1976] Rudin, W. (1976). Principles of Mathematical Analysis. McGraw Hill, 3 edition.
- [伊藤清, 1991] 伊藤清 (1991). 確率論. 岩波基礎数学選書. 岩波書店, 2 edition.
- [森岡毅 and 今西聖貴, 2016] 森岡毅 and 今西聖貴 (2016). 確率思考の戦略論—U S Jでも実証された数学マーケティングの力. 角川書店.
- [森口繁一, 1995] 森口繁一 (1995). 確率表現関数, volume 14 of UP 応用数学選書. 東京大学出版会.
- [盛田健彦, 2004] 盛田健彦 (2004). 実解析と測度論の基礎, volume 基礎編 4 of 数学レクチャーノート. 培風館.

参考文献

- [1] Billingsley76 Patrick Billingsley. (1976). *Probability and Measures*.
- [2] Billingsley68 Patrick Billingsley. (1968). *Convergence of Probability Measures*.
- [3] Billingsley71 Patrick Billingsley. (1971). *Weak Convergence of Measures*.
- [4] Bogachev Bogachev, V. I. (2018) *Weak Convergence of Measures*.
- [5] Rudin Rudin, W. *Functional Analysis*.
- [6] Dunford-Schwartz Dunford and Schwartz. *Linear Operators*.
- [7] HarmonicAnalysisOnSemigroups Berg, and Christensen, and Ressel. *Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions*.
- [8] 伊藤清伊藤清. 『確率論』
- [9] 柴田義貞柴田義貞 (1981) 『正規分布』 (UP 応用数学選書, 東京大学出版会).
- [10] ハンドブック『確率論ハンドブック』
- [11] 西尾西尾真喜子. 『確率論』
- [12] 高信高信敏『確率論』 (共立出版, 数学の魅力4).
- [13] Landkov Landkov, N. S. 『確率論入門』 (森北出版). Markov 過程に詳しく, ここに多くの読み残しあり.
- [14] Alexandorff Alexandorff, A. D. (1943). Additive set-functions in abstract spaces.
- [15] 清水良一清水良一. (1976). 中心極限定理.
- [16] 盛田盛田健彦 (2004) 『実解析と測度論の基礎』 (数学レクチャーノート, 培風館).
- [17] Yoshida Yoshida, Kousaku. *Functional Analysis*.
- [18] Butzer Butzer, P. L., and Oberdorster, W. (1975). Linear Functionals Defined on Various Spaces of Continuous Functions on \mathbb{R} . *Journal of Approximation Theory*. 13: 451-469.
- [19] Folland Folland, Gerald. *Real Analysis: Modern Techniques and Their Applications*.
- [20] Voevodsky Vladimir Voevodsky "Notes on categorical probability"
- [21] Kolmogorov31 Kolmogorov, A. N. (1933). Analytical methods in probability theory. 「私は Kolmogorov のこの論文 (「解析的方法」) の序文にあるアイデアからヒントを得て, マルコフ過程の軌道を表す確率微分方程式を導入したが, これが私のその後の研究の方向を決めることになった。」
- [22] Kolmogorov33 Kolmogorov, A. N. (1933). Grundbeegriffe der Wahrscheinlichkeitsrechnung. (確率論の基礎概念).
- [23] Levy37 Lévy, P. (1937). Théorie de l'addition des variables aléatoires. (独立確率変数の和の理論).
- [24] Doob37 Doob, J. L. (1937). Stochastic Processes Depending on a Continuous Parameter. *Transactions of the American Mathematical Society*. 42. 「正則化」の概念の初出.
- [25] ItoPhDthesis Ito, K. (1942). Differentiation of Stochastic Processes (Infinitely Divisible Laws of Probability). *Japanese Journal of Mathematics*. 18: 261-301. Levy-Ito の定理が示されている.
- [26] Ito42 Ito, K. (1942). Differential Equations Determining Markov Processes. 全国紙上数学談話会誌 (1077): 1352 – 1400.
- [27] Ito51 Ito, K. (1951). On Stochastic Differential Equations. *Memoir of American Mathematical Society*. 4: 1-51. 戦後間もなかったため, Doob の計らいでメモワールシリーズの1つとして米国で発行された. Levy による確率過程の見方と, Kolmogorov による Markov 過程への接近方法とを統一することにより, 確率微分方程式とそれに関連する確率解析の理論を創出した. 「Levy 過程を Markov 過程の接線として捉える」
- [28] Ito53 Ito, K. (1953). Stationary Random Distributions. *Memoirs of the College of Science, Kyoto Imperial University, Series. A*. 28: 209-223.

- [29] BeurlingDeny Beurling, A., and Deny, J. (1959). Dirichlet spaces. *Proceedings of the National Academy of Sciences of the United States of America*. 45 (2): 208 – 215. Dirichlet 形式が初めて定義された。
- [30] KunitaWatanabe Kunita, H., and Watanabe, S. (1967). On Square Integrable Martingales. *Nagoya Mathematics Journal*. 30: 209-245.
- [31] Meyer Meyer, P. A. (1967). Intégrales Stochastiques. *Séminaire de Probabilités I*. Lecture Notes in Math., 39: 72-162. 劣 martingale の Doob-Meyer 分解を用いて, 確率積分が一般の半マルチンゲールについて定義された。こうして確率解析の復権が起こった。
- [32] Ito70 Ito, K. (1970). Poisson Point Processes Attached to Markov Processes. *Berkeley Symposium on Mathematical Statistics and Probability*. 3: 225-239.
- [33] Ito-Selected Kiyosi Ito Selected Papers, edited by Stroock, D. W., and Varadhan S. R. S. (1986). Springer-Verlag.
- [34] Strook03 Stroock, D. (2003). *Markov Processes from K. Ito's Perspective*. Princeton University Press.
- [35] Adams Adams, W. J. (1974). *The Life and Times of the Central Limit Theorem*.
- [36] Kolmogorov Gnedenko, B. V., and Kolmogorov A. N. (1954). *Limit Distribution for Sums of Independent Random Variables*
- [37] Encyclopaedia3 Hazewinkel, M. (1995). *Encyclopaedia of Mathematics Volume 3*. Springer U.S.
- [38] Feller William Feller (1950). *An Introduction to Probability Theory and its Applications, Volume I*.
- [39] Andersen Erik Sparre Andersen. (1953). On the fluctuations of sums of random variables. *Mathematica Scandinavica*. 1:263-285, 2:195-223.
- [40] Frechet Maurice Fréchet. (1940). Les probabilités associées à un système d'événements compatibles et dépendants. *Actualités scientifiques et industrielles*.