応用をめざす

数理統計学

Web 補論 [*1)]

国友直人 [*2)]

2015 年 7 月
2017 年 1 月 (改訂)

[*1)] この原稿は朝倉書店『応用をめざす数理統計学』のホームページのために準備された Web 補論である.

[*2)] 明治大学政治経済学部特任教授, 東京大学名誉教授

Web 補論について

この Web 補論には『応用をめざす数理統計学』の誤植訂正，問題解答へのヒント，および書籍で説明した内容を補足する事項を掲載する．とくに書籍において省略したことを含め読者の皆様の勉学の参考に資すると著者が判断する補足的事項を Web 掲示することにした．「数理統計学とその応用」などの勉学の参考に資することがあれば幸いである．なお，Web 掲示の内容は適宜，変更する予定である．

誤植・訂正箇所

特に『応用をめざす数理統計学』(2015 年 8 月) に対する室井芳史教授 (東北大学)、加藤賢悟教授 (東京大学) からの指摘に感謝する。

Page 36, 下から 2 行-3 行：$I(T)$ を $\lim_{T\to\infty} \mathrm{I}(T)$ に変更 (数ヵ所)。
Page 47, 17 行：左辺は $\mathbf{E}\{\mathbf{E}[X_2|X_1]\}$ に変更。
Page 51-52, 定義 3.3: 「... 任意の $n(1 \leq n \leq N), 1 \leq i_2 \leq \cdots \leq i_n \leq N$ に対して $P(A_{i_1} \cap \cdots \cap A_{i_n}) = \prod_{j=1}^{n} P(A_{i_j})$ が成り立つ...」に変更。
Page 53, 12 行：x_1, x_2(小文字) に変更。
Page 55, 例 3.8:「互いに独立な確率変数」に変更。
Page 62, 13 行-14 行：$\{M_n\}$, $\mathcal{F}(\{M_1, \cdots, M_n\})$ に変更。
Page 71, 12 行：右辺 $it\frac{X_n}{np(1-p)}$ に変更。
Page 75, (5.12)：$\sqrt{n}(\bar{X}_n - \mu)^2$ に変更。
Page 76, 15 行：「... 条件として任意の $\delta > 0$ に対して...」に変更。
Page 80, 4 行, 7 行：$-\frac{t^2}{2}$, $-\frac{\sigma^2}{2}$ に変更。
Page 93, (下から)2 行：$n_i(\bar{X}_{i\cdot} - \bar{\bar{X}})$(大文字) に変更。
Page 96, 5 行：$\mathrm{I}(X_i \leq x)$ に変更。
Page 98, 1 行-2 行：$h[F(x)]^{h-1}[1-F(x)]^{n-h} - (n-h)[F(x)]^h[1-F(x)]^{n-h-1}$, $\sum_{h'=k-1}^{n-1}$ に変更。
Page 102, 10 行：$Y = [Ul]/[V/m]$ に変更。
Page 115, (下から)3 行-2 行：σ^2 で微分, $\frac{\partial l_n}{\partial \sigma^2}$ に変更。
Page 117, 1 行-3 行：...$g(\mathbf{\Sigma})$ の代わりに $g^*(\mathbf{\Sigma})$..., $\sum_{i>j} t_{ij}^2$,$(i > j)$ に変更。
Page 118, 例 7.9：「例 7.1 では」に変更。
Page 119, (下から)3 行：$\mathbf{E}\{[\cdot]^2\}$ に変換。
Page 121, 3 行, (下から)5 行-4 行：$f(\mathbf{X}), f(X_1|\theta), f(X_1|\theta)$(大文字) に変更。
Page 122, 12 行：$f(X|\mu)$(大文字) に変更。
Page 125, 9 行, 14 行：$\frac{\partial l_n(\theta|\mathbf{x})}{\partial\theta}$,$-\frac{\partial^2 l_n(\theta|\mathbf{X})}{\partial\theta^2}$ に変更。

Page 126, (下から)4 行,2 行：$-\frac{\partial^2 l_n(\theta|\mathbf{X})}{\partial\theta^2}$, $\frac{1}{2\sqrt{n}\sigma^4}$ に変更。
Page 127, 2 行：「したがって近似的には...」に変更。
Page 130, (7.60): $e^{-\frac{1}{2s_n^2}(x-\bar{x})}$ に変更。
Page 139, 2 行,4 行,6 行：$\phi(\mathbf{X})$(大文字) に変更。
Page 140, 4 行,6 行：x_i(小文字) に変更。
Page 147, 2 行：$p_{\cdot,j}$ に変更。
Page 158, 12 行,14 行：$\bar{x}_n$(小文字) に変更。
Page 159, 3 行：$X_1, \cdots, X_n$(大文字) に変更。
Page 159, 5 行：$M_n, \cdots, M_1$ を $X_n, \cdots, X_1$ に変更。
Page 159, (下から)4 行：1 個の $l(d,\theta)$ を削除。
Page 163, 注 ∗2)：「リスク $\int_\Theta l(x,\theta,d)f(x|\theta)\pi(\theta)d\theta$ を最小化する d をベイズ決定関数と呼ぶ。例えば...」に変更。
Page 165, (下から)1 行：「例えば」を削除。
Page 166, 1 行：$d^S(\mathbf{X})$ に変更。
Page 172, 7 行：$(\mathbf{y}-\mathbf{Zb})'\mathbf{Z}=\mathbf{0}$ に変更。
Page 174, 図：(ξ_i,η_i) は (η_i,ξ_i) $(i=1,2,3)$ に変更。
Page 174, 9 行, 11 行：ξ_i^* に変更。
Page 174, (下から)1 行：$+\sum_{i=1}^n\{\cdot\}^2$ に変更。
Page 196 (11.15)：$1-[1+\xi\frac{x}{\sigma(u)}]^{-1/\xi}$ に変更。
Page 199, 10 行：「期待値 $\mathbf{E}(Z)=\mathbf{w}'\boldsymbol{\mu}=$(一定), 基準化 $\mathbf{w}'\mathbf{1}=1$ という制約条件...」に変更。
Page 201, 18 行：「$U_i \sim U([0,1])$ に従うことを利用すると,」を削除。

数 理 的 補 論

2.B 補論——複素積分について

数理統計学や確率論では特性関数を利用すると様々な議論が一般的かつより簡明に説明できることが多い．しかしながら例えば大学文系・社会科学系などでは複素数や複素関数はあまり登場しない．

ここでは複素数と複素数の関数の有用性についての比較的分かりやすい読み物として，志賀浩二『複素数 30 講』(朝倉書店,1989) を挙げておく．複素関数を巡る議論は統計学の範囲内でも統計的時系列分析 [∗1)](statistical time series analysis) と呼ばれる分野では有用である．そこでは時間領域 (time domain) ではなく周波数領域 (frequency domain) での統計分析，例えば季節性 (seasonality) を巡る時系列の周期性・循環性などの分析に役立つ．

特性関数を正確に求めるには複素数値をとる関数の積分を評価する必要がある．ここで複素数関数 $f(z)$ の積分は実数軸と虚数軸を 2 次元ととらえて経路 C 上で (線積分と呼ばれている積分値) $\int_C f(z)dz$ を意味するが，通常の積分と同様に経路 C に沿って複素数 z と z の関数 $f(z)$(複素数) を有限和により

$$\int_C f(z)dz \sim \sum_{i=1}^n f(\xi_i)(z_i - z_{i-1}) \quad (\xi_i \in [z_{i-1}, z_i], i = 1, \ldots, n) \tag{2B.1}$$

して n についての極限操作により定められる．一見すると実関数の定積分とは結びつかないような複素積分の議論により，ほかの方法では評価が困難な積分が求められる．何回でも微分可能である性質の良い関数 (正

∗1) 例えば古典的な文献として T.W. Anderson “*The Statistical Analysis of Time Series*” (Wiley, 1971) を挙げておこう．

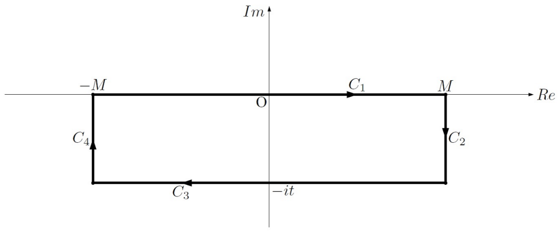

図 1　複素平面上の経路

則関数，解析関数) であれば

$$\int_C f(z)dz = 0 \tag{2B.2}$$

となることを主張するコーシーの積分定理が基本的で重要である．

標準正規分布の特性関数の評価では積分経路 $C : [-M, 0] \to [M, 0] \to [M, -it] \to [-M, -it] \to [-M, 0]$ をとる．$f(z) = e^{-z^2}$ は正則関数 (holomorphic function) なのでコーシーの積分定理を利用して閉経路 C を分割して $C_1 : (-M, 0) \to (M, 0)$ (実軸上で積分値は知られている), $C_2 : (M, 0) \to (M, -it)$, $C_3 : (M, -it) \to (-M, -it)$ (求めたい積分の経路), $C_4 : (-M, -it) \to (-M, 0)$ という 4 つの部分に積分を分割する．このとき $M \to \infty$ とすると ($f(z)$ の絶対値は 0 に収束するので M を十分に大きくすると 0 に収束する C_2, C_4 上の積分を除き) 求める積分値は通常の実数軸上の積分値に帰着される．経路の方向に注意すると (C_2 上の符号を変更する) コーシーの積分定理より結局は

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}x^2\right] dx + (-1)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}[x-it]^2\right] dx$$
$$= 0$$

(ただし上辺での x は実数) より左辺の第 1 項が 1 であるから結果が得られる．

次にコーシー分布の特性関数の場合はより複雑である．密度関数は $f(x) = 1/[\pi(1 + x^2)]$ であり，$e^{itx} = \cos(tx) + i\sin(tx)$ より $\sin(tx)$ 関数は奇関数なので積分すると 0 となるので

$$\int_{-\infty}^{\infty} e^{itx}\frac{1}{\pi}\frac{1}{1+x^2}dx = 2\int_0^{\infty} \frac{1}{\pi}\frac{\cos tx}{1+x^2}dx$$

を評価すればよい．まず $t > 0$ の場合を考える．複素関数 $f(z) = e^z/(1 + z^2)$ は解析的ではなく点 $z = i, -i$ において関数値が発散する．これらの 2 点は極と呼ばれるが，この場合には $z = i$ における留数 (residue) $\mathrm{Res}(e^{itz}/(1+z^2), i) = e^{-t}/(2i)$ と呼ばれる量 (非正則部分を部分分数分解した $1/(z-i)$ の係数 $(1/(2i))e^{it(i)}$ に対応) を利用する必要が生じる．少し込み入った議論を利用すると結局は

$$2\int_0^{\infty} \frac{\cos tx}{1+x^2}dx = 2\pi i\frac{e^{-t}}{2i} = \pi e^{-t} \tag{2B.3}$$

となる [*2)]．さらに $t < 0$ の場合も似たような議論を行い，両者をまとめるとコーシー分布の特性関数は

$$\psi(t) = e^{-|t|} \tag{2B.4}$$

となる．こうして導かれた特性関数の関数形より原点で微分可能でないことが分かり，期待値が存在しないことに対応する．

こうした定積分の計算の基礎となる複素関数論は美しい理論といわれている．複素数まで考察を広げるとバラバラと思われていた三角関数や指数関数，あるいは多様な関数の性質がより深く理解でき，一見すると評価が困難な実数軸上の定積分が統一的に求めることが可能となることによるのだろう．古典的な優れた教科書として，高木貞治『解析概論』(岩波書店, 1960; 現在手に入るのは『定本 解析概論』(2012)) がある．

---

*2)　明示的に導出しているのは例えば杉浦光夫『解析入門 II』(東京大学出版会, 1985) IX-9 節などである．

3.A 数学補論——ラドン–ニコディゥム定理について

第 3 章における条件付期待値を巡る議論を数理的に理解するために最低限必要な測度論的内容に言及するが，例えば志賀浩二『ルベーグ積分 30 講』(朝倉書店, 1990) の説明などが分かりやすい．測度論的裏付けは確率・統計・ファイナンスなどの応用上で重要である．一般に測度とは $\mathbf{R}$ 上の長さ，$\mathbf{R}^2$ 上の面積，$\mathbf{R}^3$ 上の体積などルベーグ測度を一般化した議論であるが，結果として得られるルベーグ積分はリーマン積分の一般化であり，曖昧であった積分と極限操作についての見通しがよい．確率測度は集合の関数としてはとくに有限値をとる測度である．

集合 Ω 上の加法的集合関数を $\nu(\cdot)$ とする．すなわち $\nu(\cdot)$ を $A_i \cap A_j = \phi\ (i \neq j),\ A_i \subset \Omega$ に対し，

$$\nu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \nu(A_i) \tag{3A.1}$$

が成り立つ関数 $\nu(\cdot)$ とする．ここでとくに**測度** (measure) $\mu(\cdot)$ とは条件

(i) $\mu(A) \geq 0 \quad (\forall A \in \mathcal{F})$,

(ii) $\mu(\phi) = 0$,

(iii) $A_i \in \mathcal{F}, A_i \cap A_j = \phi\ (i \neq j)$ に対し $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$

を満たす加法的集合関数 $\nu(\cdot)$，3 つ組 $(\Omega, \mathcal{F}, \mu)$ を測度空間と呼ぶ．

定義 3.A.1：測度空間 $(\Omega, \mathcal{F}, \mu)$ 上の測度 $\mu_1(\cdot)$ と $\mu_2(\cdot)$ をとる．

(i) $\forall A \in \mathcal{F}$ について $\mu_1(A) = 0 \Rightarrow \mu_2(A) = 0$ のとき $\mu_2(\cdot)$ は μ_1 に対し**絶対連続** (absolutely continuous) といい，$\mu_1 \gg \mu_2$ と記す．

(ii) $\mu_1 \gg \mu_2,\ \mu_2 \gg \mu_1$ のとき測度 μ_1, μ_2 は**等価** (equivalent) といい，$\mu_1 \sim \mu_2$ と記す．

定理 3.A.1 (ラドン–ニコディゥムの定理)：測度空間 $(\Omega, \mathcal{F}, \mu)$ において Ω は μ について σ-有限 (finite) とする．有限の加法的非負集合関数 μ_2 が条件 $\mu_1 \gg \mu_2$ を満たすとき，任意の集合 $A \in \mathcal{F}$ に対し

$$\mu_2(A) = \int_A Y d\mu_1 \tag{3A.2}$$

となる測度空間 $(\Omega, \mathcal{F}, \mu)$ 上の可積分関数 Y が存在し，

$$\frac{d\mu_2}{d\mu_1} = Y \tag{3A.3}$$

と表記する．

ここでの条件「標本空間 Ω が σ-有限」は「$\Omega = \bigcup_{k\geq 1} \Omega_k,\ \mu(\Omega_k) < +\infty$ となる $\Omega_1 \subset \Omega_2 \subset \cdots$ が存在する」ことを意味する．面積や体積など一般の測度を扱うときには有用な仮定であるが確率測度の場合にはこの条件は自動的に成り立つ．

このことから条件付期待値は常に構成できることが分かる．ここで確率変数 $X_1(\omega)$ とは別の確率変数 $X_2(\omega)$ の期待値を考える．まず条件 $X_2(\omega) \geq 0,\ E[|X_2|] < +\infty$ を仮定しよう．このとき任意の集合 A に対して

$$S(A) \equiv \int_A X_2(\omega) P(d\omega) \quad (\forall A \in \mathcal{F}) \tag{3A.4}$$

により集合の関数 $S(\cdot)$ を定義する．さらに確率変数 $X_1(\omega)$ に対して σ-加法族 $\mathcal{F} \supset \mathcal{G} \equiv \{X_1^{-1}(B) | B \in \mathbf{B}(\mathbf{R})\}$ より $\mathcal{G}$ は $\mathcal{F}$ の一部分となるので

$$Q_1(X_1^{-1}(B)) \equiv P(X_1^{-1}(B)) \tag{3A.5}$$

により関数 $Q_1(X_1^{-1}(B))$ を定義できる．この $Q_1(\cdot)$ は確率測度なので

$$F_1(x_1) = Q_1(\omega | X_1(\omega) \leq x_1) \tag{3A.6}$$

により確率分布関数が定義できる．このとき $S(\cdot)$ と $Q_1(\cdot)$ の間には「$Q_1(X_1^{-1}(B)) = 0$ ならば $S(X_1^{-1}(B)) =$

0」という関係がある．すなわち確率測度 $Q_1(\cdot)$ で測ると 0 である事象は測り方を変えた測度 $S(\cdot)$ でも 0 となる意味である (絶対連続性 (absolute continuity) は実数値関数の連続性の拡張である)．したがってラドン–ニコディウムの定理 (一種の微係数である $Y(\cdot)$ は (ラドン–ニコディウム密度と呼ばれる) を用いると任意の $\mathbf{B}(\mathbf{R})$ に対して

$$
\begin{aligned}
S(X_1^{-1}(B)) &= \int_{X_1^{-1}(B)} Y(\omega)Q_1(d\omega) \\
&= \int_B y(x_1)dF_1(x_1)
\end{aligned}
$$

を満足する可測関数 $Y(\omega)$ (あるいは $y(x_1)$) が存在する．そこで $y(x_1) = \mathbf{E}[X_2|X_1 = x_1]$ と定義すれば上の関係は条件付確率の一般化であるのでこの関数 $y(\cdot)$ を条件付期待値と呼ぶことが妥当であろう．確率変数 $X_2(\omega)$ が正値関数とは限らない一般の場合には $X_2 = X_2^+(\omega) - X_2^-(\omega)$, $X_2^+ \geq 0$, $X_2^- \geq 0$ に対して，

$$
\mathbf{E}[X_2|X_1 = x_1] \equiv \mathbf{E}[X_2^+|X_1 = x_1] - \mathbf{E}[X_2^-|X_1 = x_1] \tag{3A.7}
$$

とおけば，正値関数についての議論より

$$
\mathbf{E}[X_2] = \mathbf{E}[\mathbf{E}(X_2|X_1)] \tag{3A.8}
$$

が成立する．

5.B　弱収束についての補論

ここで確率測度の分布収束・弱収束に関する重要な事項について補足しておく．

補題 5.B.1：次の条件は同等である．
(i) $F_n \xrightarrow{\mathcal{L}} F$,
(ii) $\forall f \in C_b(\mathbf{R})$ に対し $\int f dQ_n \to \int f dQ$. ただし $C_b(\mathbf{R})$ は $\mathbf{R}$ 上の有界連続関数の全体とする [*3)].

ここで補題 5.B.1 により確率分布の収束は特性関数の収束を意味することが分かる．この逆を示すには次の概念が必要となる．

定義 5.B.1：確率分布の列 P_n がタイト (tight) とは任意の ϵ に対し有限の実数 $a < b$ が存在して

$$
P_n(x \in [a, b]) > 1 - \epsilon \tag{5B.1}
$$

となる．

この条件は確率分布の列が $n \to \infty$ のとき有限区間に収まらずに遠くに逃げてしまうことを防ぐための必要十分な条件である．実は特性関数が原点で連続ならこの条件を満たす．このことは任意の $u\ (> 0)$ について

$$
\begin{aligned}
\frac{1}{u}\int_{-u}^{u}[1 - \phi_n(t)]dt &= 2\int_{-\infty}^{\infty}\Big[1 - \frac{\sin ux}{ux}\Big]dF_n(x) \\
&\geq 2\int_{|x|\geq(2/u)}\Big[1 - \frac{1}{|ux|}\Big]dF_n(x) \\
&\geq P\Big(|x| \geq \frac{2}{u}\Big)
\end{aligned}
$$

より導ける．

特性関数が収束するとき，反転公式より分布関数も収束する．上の条件を満たしていれば分布関数が退化することはないので分布関数列も分布関数に収束することになる．

補題 5.B.1 の証明：(ii) $\Rightarrow$ (i) を示す．任意の $x \in \mathbf{R}$ と $\delta > 0$ に対して関数 $h \in C_b(\mathbf{R})$ として

---

*3)　積分記号 $\int f dQ$ は $\int f(\omega)Q(d\omega)$ と同じ意味である．

$$h(y) = \begin{cases} 1 & (y \le x) \\ 1 - \delta^{-1}(y-x) & (x < y < x+\delta) \\ 0 & (y \ge x+\delta) \end{cases} \tag{5B.2}$$

とする．このとき $\mathbf{E}_{Q_n}[h(X_n)] \ge F_n(x)$ かつ $F(x+\delta) \ge \mathbf{E}_Q[h(X)]$ である (ここで $\mathbf{E}_{Q_n}[\cdot]$ および $\mathbf{E}_Q[\cdot]$ は O_n および Q についての期待値を意味する)．したがって $\lim_n \sup F_n(x) \le F(x+\delta)$ となる．ここで $\delta \downarrow 0$ とすると関数 $F(\cdot)$ は右連続より $\lim_n \sup F_n(x) \le F(x)$ となる．同様にすれば $\lim_n \inf F_n(x-) \ge F(x-\delta) \to F(x-)$ も得られる．これらを合わせると $F(x-) \le \lim_n \inf F_n(x-) \le \lim_n \sup F_n(x) \le F(x)$ となるので，連続点では $\lim F_n(x) = F(x)$ となる．**Q.E.D.**

定義 5.B.2 (分布族の tightness)：確率分布 μ_n が tight であるとは $\iff$ $\forall \epsilon > 0,\ \exists a < b$ ：

$$\forall n \ \ \mu_n([a,b]) = F_n(b) - F_n(a-) > 1 - \epsilon \tag{5B.3}$$

を満足する．

補題 5.B.2：「確率分布族 μ_n が tight」ならば

$$\exists \mu_{n_i}\ \exists \mu \in Prob(\mathbf{R}) \ \ s.t.\ \mu_{n_i} \xrightarrow{w} \mu\ . \tag{5B.4}$$

注意：$\mu_n \in Prob(\mathbf{R}^k)$ $(k \ge 1)$ で成立する．ただし，ここで $Prob$ は確率分布を意味する．とくに $k = 1$ のときには必要十分条件となることが知られている．

練習問題へのヒント・解の例

第　1　章

問 1.1　(i) 集合 A, B に対し集合差 $A \setminus B = A \cap B^c$ とする．$A \setminus B = A \setminus (A \cap B)$ を示せ．
(略解) 任意の $x \in A \setminus (A \cap B)$ について $x \in A \setminus B$ をまず示し，逆も示す．
(ii) 正整数 I，集合 $A_i = \{x | x \in (1/i, 1]\}$ とするとき $\bigcup_{i=1}^{\infty} A_i$ および $\bigcap_{i=1}^{\infty} A_i$ を求めよ．
(略解) 区間の端に注意すると $\bigcup_{i=1}^{\infty} A_i = (0,1]$, $\bigcap_{i=1}^{\infty} A_i = \{\o\}$.
(iii) 集合 $A_i = \{x | x \in [a/i, b]\}$ $(0 < a \le b)$, $B_i = [-a, b] \setminus A_i$ とするとき $\bigcup_{i=1}^{\infty} A_i$, $\bigcap_{i=1}^{\infty} A_i$, $\bigcup_{i=1}^{\infty} B_i$, $\bigcap_{i=1}^{\infty} B_i$ を求めよ．
(略)
問 1.2　$\mathcal{F}$ を $\mathbf{\Omega}$ 上の σ-加法族とする．加算個の事象 $A_1, A_2, \ldots \in \mathcal{F}$ に対して，A_i が単調増加 $(A_1 \subset A_2 \subset \cdots)$ なら $P(\bigcup_{i=1}^{\infty} A_i) = \lim_{n\to\infty} P(A_n)$ となることを示せ．
(略解) $A_1 \subset A_2 \subset \cdots$ より $\bigcup_{i=1}^{n} A_i = A_1 \cup (A_2 \setminus A_1) \cup \cdots \cup (A_n \setminus A_{n-1})$ となる．そこで $P(\bigcup_{i=1}^{n} A_i) = P(A_n)$ より結果を得る．
問 1.3　定義 1.4 の条件 (iii) を示せ．
(略解) 任意の $x \le y$ に対して

$$F(y) = P(\omega | X(\omega) \le y) = P(\{\omega | X(\omega) \le x\} \cup \{\omega | x < X(\omega) \le y\}) \ge F(y)$$

より結果を得る．

第　2　章

問 2.1　(i) X が密度関数 $f(x)=\alpha\beta^\alpha/x^\alpha$ $(x>\beta); f(x)=0$ $(x\le\beta)$ $(\alpha>0,\beta>0)$ のパレート分布に従うとき期待値，分散などの性質を考察せよ．

(略解) $\alpha>1$ のとき $\mathbf{E}[|X|]<\infty$ となり $\mathbf{E}[X]=\frac{\alpha}{\alpha-1}\beta$．$\alpha>2$ のとき $\mathbf{E}[|X|^2]<\infty$ となり $\mathbf{E}[X]=\frac{\alpha}{(\alpha-1)^2(\alpha-2)}\beta^2$．

(ii) 確率関数が $p(x)={}_{r+x-1}C_x p^r q^x$ $(q=1-p, 0<p<1, r\{正整数\})$ である負の 2 項分布に従うとき，期待値・分散などの性質を考察せよ．

(略解) 負の 2 項分布の特性関数は $\psi(t)=p^r(1-qe^{it})^{-r}$ より $\mathbf{E}[X]=q/p, \mathbf{V}(X)=q/p^2$ が得られる．

問 2.2　X が正規分布，Y がポワソン分布に従うとき期待値 $\mathbf{E}(X)$，分散 $\mathbf{V}(X)$，歪度，尖度などの性質を考察せよ．

(略解) 平均まわりの正規分布の特性関数は $\psi(t)=e^{-\frac{\sigma^2 t}{2}}$ より平均まわりの積率 $(k=1,2,\ldots)$ は $\mathbf{E}[(X-\mu)^{2k}]=\frac{(2k)!}{k!2^k}\sigma^{2k}$, $\mathbf{E}[(X-\mu)^{2k-1}]=0$ となる．したがって $\kappa_3=0, \kappa_4=3$ となる．ポワソン分布の特性関数は $\psi(t)=e^{\lambda(e^{it}-1)}$ より $\mathbf{E}[X]=\mathbf{X}[X]=\lambda$ となる．さらに $\kappa_3=1/\lambda^{1/2}, \kappa_4=1/\lambda+3$ である．

問 2.3　積率 $\mathbf{E}[|X|^r]$ $(r=1,2,3,4)$ が有限であるための裾確率 $P(|X(\omega)|>x)$ の条件を考察せよ．

(略解) 本文の説明と同様に積分の順序交換を行うと $\mathbf{E}[|X|^r]=r\int_0^\infty P(|X|>x)x^{r-1}dx$ となる．例えば $r=1$ に対しては $P(|X|>x)\sim 1/x^{1+\alpha}$ $(\alpha>0)$ が十分条件となる．

問 2.4　標準正規分布の分布関数 $\Phi(x)$，密度関数 $\phi(x)$ とする．任意の $X>0$ に対し $[x+x^{-1}]^{-1}\le[1-\Phi(x)]/[\phi(x)]\le x^{-1}$ を示せ．

(略解) $G_1(x)=x[1-\Phi(x)]-\phi(x)$, $G_2(x)=(x^2+1)[1-\Phi(x)-x\phi(x)]$ とすると $G_1(x)\ge 0, G_2(x)\ge 0$ を示せばよい．

問 2.5　X_i $(i=1,2)\sim N(0,1)$ が互いに独立なとき変数変換を利用し $Y=X_1^2+X_2^2$ の分布を求めよ．独立でない場合には分布はどうなるか．

(解) (i) まず $X_i\sim N(0,1)$ が $i.i.d.$ 系列の場合を考察する．$W_i=X_i^2$ と変数変換すると W_i の確率密度関数は

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

であるから，積率母関数 $\phi_{W_i}(\theta)$ は

$$\phi_{W_i}(\theta)=\mathbf{E}[e^{t\theta W_i}]=\frac{1}{2^{1/2}\Gamma(1/2)}\int_0^\infty w_i^{-\frac{1}{2}}e^{-\frac{1}{2}(1-2\theta)w_i}\,dw_i$$

である．ここで $z=(1-2\theta)w_i/2$ と変数変換すると

$$\int_0^\infty w_i^{-\frac{1}{2}}e^{-\frac{1}{2}(1-2\theta)w_i}\,dw_i=\left(\frac{1}{2}(1-2\theta)\right)^{-\frac{1}{2}}\int_0^\infty z^{-\frac{1}{2}}e^{-z}\,dz=\sqrt{2}(1-2\theta)^{-\frac{1}{2}}\Gamma(1/2)$$

なので $\phi_{W_i}(\theta)=(1-2\theta)^{-1/2}$ となる．$Y=\sum_{i=1}^n W_i$ であるから，Y の積率母関数は

$$\begin{aligned}\phi_Y(\theta)&=\prod_{i=1}^n\phi_{W_i}(\theta)\\&=(1-2\theta)^{-n/2}\end{aligned}$$

である．またガンマ分布 $Gamma(n/2,2)$ の積率母関数は $\phi(\theta)=(1-2\theta)^{-n/2}$ である．したがって自由度 n の χ^2 分布とガンマ分布 $Gamma(n/2,2)$ の積率母関数は等しいので，積率母関数の一意性から $Y\sim Gamma(n/2,2)$．

(ii) 一般の場合には少し複雑になる．

$$\psi_n(t)=\prod_{j=1}^n\mathbf{E}[e^{itX_j^2}]$$

を求めるには指数関数の中を整理 ($(x-\mu)^2-2it\sigma^2x^2=(1-2it\sigma^2)[x-\mu/(1-2it\sigma^2)]^2+\mu^2-[\mu^2/(1-2it\sigma^2)]$ を利用) して積分を実行すると (途中は省略)

$$\mathbf{E}[e^{itX_j^2}]=\frac{1}{\sqrt{2\pi}}\int_{\infty}^{\infty}e^{itx^2}e^{-\frac{1}{2\sigma^2}(x-\mu)^2}dx=\frac{1}{\sqrt{1-2it\sigma^2}}e^{\frac{\mu^2it}{1-2it\sigma^2}}$$

となるので

$$\psi_n(t)=\left[\frac{1}{1-2it\sigma^2}\right]^{n/2}e^{\frac{n\mu^2it}{1-2it\sigma^2}}$$

となる．特性関数の反転公式を利用すると密度関数は

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}\psi_n(t)dt$$

で与えられる ($f(t)$ の可積分性をチェックする必要はあるがこの場合は示せる)．特性関数の指数部分を $n\mu^2it=n\mu^2[1-(1-2it\sigma^2)]/[2\sigma^2]$ となることに注意して展開して, ガンマ分布に関する反転公式を利用すると

$$f(x)=e^{-n\mu^2/[2\sigma^2]}\sum_{k=0}^{\infty}\frac{1}{k!}\left[\frac{n\mu^2}{2\sigma^2}\right]^k g_{n/2+k,2\sigma^2}(x)$$

という表現が得られる．とくに $\mu=0,\sigma=1$ とすると $k=0$ 以外の項は消えて確かに $\chi^2(n)$ に一致する．ここで $g_{\nu,\alpha}(x)$ はガンマ分布の密度関数である．

補足事項：確率変数列 X_i $(i=1,\ldots,n)$ の期待値 $\mu\neq 0$ のときの Y の分布は非心ガンマ分布であるが，例えば検定統計量の検出力 (対立仮説の下での統計量の分布) を求めるためには重要な役割を果たす．スペースの関係もあり非心分布については本書では説明していないが，良い機会であるので議論した (なおいくつかの非心分布については竹村彰通『現代数理統計学』(創文社, 1991) に説明がある)．

問 2.6　$X\sim N(\mu,\sigma^2)$ のとき $Y=e^X$ (対数正規分布) の期待値，分散，歪度，尖度の性質を考察せよ．

(**略解**) 対数正規分布では積率母関数は存在しないが特性関数は存在する．積率は特性関数を利用しなくても次のように求めることができる．$Y=\log X$ とすると

$$\begin{aligned}\mathbf{E}[X^n]&=\mathbf{E}[e^{nY}]\\&=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}e^{ny-\frac{1}{2\sigma^2}(y-\mu)^2}dy\\&=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}e^{-\frac{1}{2\sigma^2}(y-n\sigma^2-\mu)^2}dye^{\frac{1}{2\sigma^2}(n\sigma^2+\mu)^2-\frac{1}{2\sigma^2}\mu^2}\\&=\exp\left[\frac{1}{2\sigma^2}(n\sigma^2+\mu)^2-\frac{1}{2\sigma^2}\mu^2\right]\end{aligned}$$

となるので整理すればよい．

問 2.7　関数 $F_1(x)=\exp[-\exp(-x)]$ $(-\infty<x<\infty)$ および $F_2(x)=\exp[-x^{-1}]$ $(x>0)$ はそれぞれ確率分布関数とみなせるか．

(**略**)

問 2.8　X が指数分布 $EX(\alpha)$ ($\alpha>0$ は母数) に従うとき，任意の $s>t>0$ に対して $P(X>s|X>t)=P(X>s-t)$ となることを示せ．逆は成り立つか．

(**略**)

問 2.9　$P(X>0)=1$ のとき $\mathbf{E}[1/X]\geq 1/[\mathbf{E}[X]]$ となることを示せ．不等号が成立する例を考えよ．

(**解**) 例えば $x_i>0,P(X=x_i)=p_i$ $(i=1,2)$ とすると $p_1+p_2=1$ より

$$\begin{aligned}\mathbf{E}[X]\mathbf{E}\left[\frac{1}{X}\right]&=[p_1x_1+p_2x_2][p_1x_1^{-1}+p_2x_2^{-1}]\\&=[p_1+p_2]^2+p_1p_2\left[\sqrt{\frac{x_1}{x_2}}-\sqrt{\frac{x_2}{x_1}}\right]^2\geq 1\end{aligned}$$

となる．$p_i>0,x_1\neq x_2$ の例を作ればよい．より一般には

$$\mathbf{E}[X]\mathbf{E}\Big[\frac{1}{X}\Big] = \Big[\sum_{i=1}^{n} p_i x_i\Big]\Big[\sum_{j=1}^{n} p_j x_j^{-1}\Big]$$
$$= \Big[\sum_{i=1}^{n} p_i\Big]^2 + \frac{1}{2}\sum_{i,j=1}^{n} p_i p_j \left[\sqrt{\frac{x_i}{x_j}} - \sqrt{\frac{x_j}{x_i}}\right]^2 \geq 1$$

となる．さらに連続分布について厳密に議論しようとすると離散近似の議論が必要がある．

コーシー–シュワルツの不等式を利用すると，いま $P(X>0)=1$ であるから $\mathbf{E}(X)>0,\ \mathbf{E}(1/X)>0$ である．$\mathbf{E}\left[\frac{1}{X}\right]^{1/2}\mathbf{E}\left[X\right]^{1/2} \geq \mathbf{E}\big[\frac{1}{\sqrt{X}}\sqrt{X}\big]=1$ より $\mathbf{E}(X)>0$ に注意すれば，

$$\mathbf{E}\left[\frac{1}{X}\right] \geq \frac{1}{\mathbf{E}\left[X\right]}$$

となる．等号は $\sqrt{X}=c\frac{1}{\sqrt{X}}$ a.s. (c : (一定)) のときである．例えば X をそれぞれ $1/2$ ずつの確率で $1,3$ をとる確率変数とすれば，$\mathbf{E}\left[\frac{1}{X}\right]=2/3,\ \frac{1}{\mathbf{E}[X]}=1/2$ より厳密な不等号が成立する．

問 2.10　確率変数 X がガンマ分布 $Gamma(\alpha,\beta)$ に従うとき $Y=X^{-1}$ の確率分布を求めよ．

(略)

問 2.11　例 2.13 の説明を確かめよ．

(略)

問 2.12　独立な確率変数列 X_i $(i=1,\dots,n)$ が $N(\mu,\sigma^2)$ に従うとき $Y=\sum_{i=1}^{n} X_i^2$ の特性関数と確率分布はどのように特徴付けられるか．まず $\mu=0$ の場合を考察せよ．

(解) X_1 と X_2 が独立のとき同時密度関数は

$$f(x_1,x_2)=(2\pi)^{-1}\exp\left(-\frac{x_1^2+x_2^2}{2}\right)$$

となる. ここで $X_1=R\cos\Theta,\ X_2=R\sin\Theta$ $(R>0,\ 0\leq\Theta<2\pi)$ と変数変換するとヤコビアンは

$$\left|\frac{\partial(x_1,x_2)}{\partial(r,\theta)}\right| = r$$

であるので，(R,Θ) の同時密度関数は

$$f(r,\theta)=\frac{1}{2\pi}\cdot\frac{r}{2}\exp\left(-\frac{r^2}{2}\right)$$

となる．これより R と Θ は独立，R^2 は指数分布 $EX(2)$，ガンマ分布 $Gamma(\nu/2,2)$，$\nu=2$，すなわち自由度 2 の χ^2 分布に従うことがわかる.

X_1 と X_2 が独立でない場合には X_1 と X_2 の相関係数を ρ とすると独立に $N(0,1)$ に従う確率変数 Z_1,Z_2 を用いて

$$X_1=\frac{1}{\sqrt{2}}\left(\sqrt{1+\rho}Z_1+\sqrt{1-\rho}Z_2\right),\quad X_2=\frac{1}{\sqrt{2}}\left(\sqrt{1+\rho}Z_1-\sqrt{1-\rho}Z_2\right)$$

と表現することができる (証明は各自) ので，

$$X_1^2+X_2^2=(1+\rho)^2Z_1^2+(1-\rho)^2Z_2^2=(1+\rho)^2\chi_1^2+(1-\rho)^2\chi_1^2$$

となる. ただし χ_1^2 は自由度 1 の χ^2 分布を表す．$\rho=0$(独立) のときは $\chi_1^2+\chi_1^2$ となるので確かに χ_2^2 になる．

問 2.13　$X|Y\sim B(Y,p)$，Y は $P_o(\lambda)$ に従うとするとき X の分布を求めよ (生態学の例では Y は eggs の数，$X|Y$ は生存数が典型例).

(解) 条件付確率の式から $P(X=x,Y=y)=P(X=x|Y=y)P(Y=y)$ が成り立つ．また $X|Y$ は 2 項分布 $B(Y,p)$ に従っているので $X>Y$ となる確率は 0 であることに注意すると，X の確率関数は

$$
\begin{aligned}
P(X=x) &= \sum_{y=0}^{\infty} P(X=x, Y=y) \\
&= \sum_{y=0}^{\infty} P(X=x|Y=y)P(Y=y) \\
&= \sum_{y:y\geq x}^{\infty} \frac{y!}{x!(y-x)!} p^x (1-p)^{y-x} \frac{\lambda^y}{y!} e^{-\lambda} \\
&= e^{-\lambda} \frac{(\lambda p)^x}{x!} \sum_{y:y\geq x}^{\infty} \frac{\{\lambda(1-p)\}^{y-x}}{(y-x)!} \\
&= e^{-\lambda} \frac{(\lambda p)^x}{x!} \sum_{z=0}^{\infty} \frac{\{\lambda(1-p)\}^{z}}{z!} \\
&= e^{-\lambda} \frac{(\lambda p)^x}{x!} e^{\lambda(1-p)} = \frac{(\lambda p)^x}{x!} e^{-\lambda p}
\end{aligned}
$$

となる．これはパラメータ λp のポワソン分布の確率関数である．

第　3　章

問 3.1　定義 3.1 の条件付確率が確率測度であることを示せ．

(略)

問 3.2　(3.13) が σ-加法族であることを示せ．

(略)

問 3.3　定理 3.1 (ベイズの公式) を示せ．

(略)

問 3.4　例 3.2 において $p=2, p_1=p_2=1$ のとき $\boldsymbol{\Sigma}$ の逆行列要素を評価して同時密度関数と条件付密度関数の表現を確認せよ．

(略)

問 3.5　確率変数 X, Y について $Y \neq X$ のとき $\mathbf{E}[Y^2|X] = X^2, \mathbf{E}[Y|X] = X$ となることがあるか．

(解) 条件より

$$
\mathbf{E}[(Y-X)^2] = \mathbf{E}\left[\mathbf{E}(Y^2|X) - 2X\mathbf{E}(Y|X) + X^2\right] = 0
$$

となる．ここでもし $P(\omega||Y(\omega)-X(\omega)|>0)>0$ なら $\mathbf{E}[(Y-X)^2]>0$ となる．

問 3.6　正規分布に従う確率変数 X, Y について $\mathbf{E}[X|Y] = \mathbf{E}[Y|X]$ となることがあるか．

(解) $X_1 = X, X_2 = Y$ とすると条件は

$$
\mu_1 + \frac{\sigma_{12}}{\sigma_{22}}(X_2-\mu_2) = \mu_2 + \frac{\sigma_{12}}{\sigma_{11}}(X_1-\mu_1)
$$

であるから両辺に $(X_1-\mu_1), (X_2-\mu_2)$ を乗じて期待値をとると，

$$
\sigma_{12}\left[1-\frac{\sigma_{12}}{\sigma_{22}}\right] = 0\,, \quad \sigma_{12}\left[1-\frac{\sigma_{12}}{\sigma_{11}}\right] = 0
$$

が得られる．したがって相関係数について $\rho_{12} = 0 \text{ or } 1$ が必要になる．さらに前者なら $\mu_1 = \mu_2$ は明らか．後者なら $\mu_1 + (X_2-\mu_2) = \mu_2 + (X_1-\mu_1)$ より両辺の期待値をとると $\mu_1 = \mu_2$ となる．

問 3.7　3 つの確率変数 X_1, X_2, X_3 について任意の 2 つのペアが独立なら 3 つの確率変数が独立といえるか．

(略解) 例 3.4 を参考にせよ．

問 3.8　(3.36) で定められる確率変数列 $\{X_i\}$ の期待値，分散，(X_i, X_j) $(i \neq j)$ の共分散，相関を求めよ (ただし条件が必要なら例えば初期条件 $X_0 = 0$ と仮定せよ)．

第　4　章

問 4.1　期待値が一定である互いに独立な確率変数列 X_i $(i = 0, 1, \ldots, n)$ (ただし $\mathbf{E}[X_i] = \mu, \mathbf{V}[X_i] = \sigma_i^2 < \infty$ とする) より作られるマルチンゲールの例を挙げよ.

(略)

問 4.2　互いに独立な確率変数列 $X_i \sim N(0, \sigma^2)$ $(i = 1, 2, \ldots, n)$ とする (σ^2 は母数とする). ある関数 σ_{1n}, σ_{2n} をとり $Y_n = \sum_{i=1}^n X_i^2 - \sigma_{1n}$, $W_n = \exp[\sum_{i=1}^n X_i - \sigma_{2n}]$ がマルチンゲールとなることがあるか.

(解) $Y_{n+1} = [\sum_{i=1}^n X_i^2 + X_{n+1}^2] - \sigma_{1,n+1} = Y_n + X_{n+1}^2 - (\sigma_{1,n+1} - \sigma_{1n})$ と表現すると

$$\begin{aligned}\mathbf{E}(Y_{n+1}|Y_1, \ldots, Y_n) &= \mathbf{E}(Y_{n+1}|X_1, \ldots, X_n) \\ &= Y_n + \mathbf{E}(X_{n+1}^2|X_1, \ldots, X_n) - (\sigma_{1,n+1} - \sigma_{1n})\end{aligned}$$

ここで X_{n+1} と $X_1, \ldots, X_n$ は独立より $\mathbf{E}(X_{n+1}^2|X_1, \ldots, X_n) = \sigma^2$. したがって Y_n がマルチンゲールとなるためには $\sigma^2 - \sigma_{1,n+1} + \sigma_{1n} = 0$ であればよいので $\sigma_{1(n+1)} - \sigma_{1n} = \sigma^2$(等差数列) であるので $\sigma_{1n}^2 = n\sigma^2$ であればよい. 同様にして

$$W_{n+1} = \exp(X_{n+1} - (\sigma_{2(n+1)} - \sigma_{2n})) \cdot W_n$$

であるので

$$\mathbf{E}(W_{n+1}|W_1, \ldots, W_n) = \exp\left(\frac{\sigma^2}{2} - (\sigma_{2,n+1} - \sigma_{2n})\right) \cdot W_n$$

より, $\sigma_{2(n+1)} - \sigma_{2n} = \sigma^2/2$ であればよい. したがって $\sigma_{2n} = n\sigma^2/2$ ととれば W_n はマルチンゲールになる.

問 4.3　マルチンゲールについて X_n が可積分かつ条件付期待値が

$\mathbf{E}[X_{n+1}|X_n, X_{n-1}, \ldots, X_1] = (1/n)(X_1 + \cdots + X_n)$ のとき $Y_n = (1/n)(X_1 + \cdots + X_n)$ はマルチンゲールとなるか.

(解) Y_n の定義から

$$\mathbf{E}\left(X_{n+1}|Y_1, \ldots, Y_n\right) = \mathbf{E}\left(X_{n+1}|X_1, \ldots, X_n\right)$$

となることに注意する. ここで

$$Y_{n+1} = \frac{1}{n+1}(nY_n + X_{n+1})$$

と表現できることから

$$\begin{aligned}\mathbf{E}(Y_{n+1}|Y_1, \ldots, Y_n) &= \mathbf{E}(Y_{n+1}|X_1, \ldots, X_n) \\ &= \frac{n}{n+1}Y_n + \frac{1}{n+1}\mathbf{E}(X_{n+1}|X_1, \ldots, X_n) \\ &= \frac{n}{n+1}Y_n + \frac{1}{n+1}Y_n = Y_n\end{aligned}$$

であるので, Y_n はマルチンゲールである.

問 4.4　$X_n \to c$ (一定値, 一定の確率変数) となるマルチンゲールの例, $X_n \to -\infty$ (あるいは $X_n \to \infty$) となるマルチンゲールの例があるか.

(解) 簡単な例としては確率 1 で常に $X_n = c$ となるものを考えればよいが, 期待値 0 の独立確率変数列 $Z_{k,n}$ $X_{n+1} - c = Z_{1n} + \cdots + Z_{n+1,1}$ より $\mathbf{V}[Z_{1n} + \cdots + Z_{n+1,1}] \to 0$ $(n \to \infty)$ という例も作れる. X_n が発散する例としては $X_0 = c$ $(\neq 0)$ として X_{n+1} が 1/2 ずつの確率で $4X_n$ と $-2X_n$ の値をとるものを考えればよい. このとき

$$\mathbf{E}(X_{n+1}|X_1, \ldots, X_n) = \frac{1}{2}(4X_n - 2X_n) = X_n$$

が成り立つのでマルチンゲールであり,

$$|X_{n+1}| \geq 2|X_n|$$

が成り立つので $|X_n| \to \infty$ $(n \to \infty)$ となる．

問 4.5　例 4.3 における主張は確率変数列が X_i $(i = 1, \dots, n)$ が定常な 1 次自己回帰モデル (3.36) に従うときにも成立するか (ただし定常条件は $|a| < 1$ とする).

(解) X_i が定常 AR(1) に従うとき期待値 $\mathbf{E}[X_i] = 0$，2 次積率は $\mathbf{E}[X_i^2] = \sigma_2/(1 - a^2)$ となる．したがって例えば 4 次積率の存在を仮定すれば

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p} \mathbf{E}[X_1^2]$$

となる．$a \neq 0$ のとき右辺は σ^2 に一致しない．

第　5　章

問 5.1　自由度 n の t 分布は $n \to \infty$ のときに $N(0, 1)$ に収束することを示せ。

(解) $X_n \sim \chi^2(n)$ のとき X_n/n は n が大きいときに 1 に確率収束する．$Z \sim N(0, 1)$ でかつ X_n と独立とすると，(t 統計量) $Z/\sqrt{X_n/n}$ は $n \to \infty$ のとき $N(0, 1)$ に分布収束する．

問 5.2　自由度 n の χ^2 分布の自由度 $n \to +\infty$ のとき正規分布で近似できることを示せ．

(解) X_i $(i = 1, \dots, n)$ が互いに独立に $N(0, 1)$ に従うとき，$Y_n = \sum_{i=1}^{n}(X_i^2 - 1) = \sum_{i=1}^{n} Z_i$ に中心極限定理を適用する (なお $\mathbf{E}[Z_i] = 0$,$\mathbf{E}[Z_i^2] = 2$ である).

問 5.3　確率変数列 X_n が確率収束するが平均 2 乗収束するとは限らないことを示せ．

(略解) 例 4.4 を参照.

問 5.4　リヤプノフ条件の下で中心極限定理が成り立つことを説明せよ．

(略解) 任意の $\epsilon > 0, \delta > 0$ に対して

$$\sum_{k=1}^{n} \mathbf{E}[X_{nk}^2 |X_{nk}| \geq \epsilon] \leq \sum_{k=1}^{n} \mathbf{E}\left[X_{nk}^2 \left(\frac{|X_{nk}|}{\epsilon}\right)^{\delta}\right] = \left(\frac{1}{\epsilon}\right)^{\delta} \sum_{k=1}^{n} \mathbf{E}[|X_{nk}|^{2+\delta}]$$

より $n \to \infty$ のときリンドバーグ条件が成立する．

問 5.5　確率変数列 X_n の積率母関数 $M_n(\theta)$，特性関数 $\phi_n(t)$ とする．M_n と ϕ_n はある関数 M と ϕ に収束するとき X_n の分布関数は収束するといえるか．

(略解) 例えば $\phi_n(t) = e^{-nt^2/2}$ とすると $n \to \infty$ のとき 1 点分布に退化する．さらに $X_n \sim U(-n, n)$ (一様分布) に従うとき $\phi_n(t) = \sin(nt)/(nt)$ $(t \neq 0); 1$ $(t = 0)$ となる．$n \to \infty$ のとき $\phi_n(t) \to 0$ $(t \neq 0); 1$ $(t = 0)$ となるが，X_n は確率分布に収束しない．

第　6　章

問 6.1　標本平均の平均 (期待値) の公式 (6.2) を丁寧に導け．有限母集団の場合の標本 X_1 と X_2 の共分散の式を説明せよ．

(略)

問 6.2　(6.23) の分母と分子が独立となることを考察せよ．

(略解) 確率変数 $X_{ij} - (1/n_i)\sum_{j=1}^{n_i} X_{ij}$ と $(1/n_i)\sum_{j=1}^{n_i} X_{ij} - (1/n)\sum_{i=1}^{m}\sum_{j=1}^{n_i} X_{ij}$ 共分散を計算すると 0 になる．正規分布に従う確率変数の場合には無相関なら独立となる．

問 6.3　(6.25) の分解に現れる $\mathbf{P}_n, \mathbf{P}_{n1}, \mathbf{P}_{n2}$ が射影行列であることを確認せよ．

(略解) 正方行列 $\mathbf{P}$ が射影行列となる必要十分条件は $\mathbf{P}^2 = \mathbf{P}$, $\mathbf{P}^{'} = \mathbf{P}$($\mathbf{P}^{'}$ は転置) であるからこれを確かめればよい.

問 6.4　(6.18) に現れる c_m を導出に現れる a_m より導け．

(略)

問 6.5　$X_1, \dots, X_n$ が独立に $N(\mu, \sigma^2)$ に従うとき確率変数 $Y = \sum_{i=1}^{n} X_i^2$ とする．$\mu = 0$ のとき積率母関数・特性関数を求め Y がガンマ分布に従うことを示せ．さらに $\mu \neq 0$ のときの Y の従う確率分布を考察せ

よ．

(略解) 問題 2.5 の略解を参照．

問 6.6　互いに独立で同一のコーシー分布に従う場合の最大値の分布の導出を確かめよ．

(略解) 6.4 節の議論を利用すると $F^n(nz) = [1 - \frac{1}{nz}nz(1 - F(nz))^n]$ である．$y = \tan z$ とすると $\frac{dy}{dz} = 1 + \tan^2 z$ より

$$1 - F(nz) = 1 - \int_{-\infty}^{nz} \frac{1}{1+x^2}dx = \frac{1}{2} - \int_0^{nz} \frac{1}{1+x^2}dx = \frac{1}{2} - \frac{1}{\pi}\tan^{-1}[nz]\int_0^{nz} \frac{1}{1+x^2}dx$$

となる．また $n \to \infty$ のとき $(nx)\tan^{-1}(nx) \to -1$ であるから $F^n(nz) \sim [1 - \frac{1}{nz\pi}]^n$ となり結果を得る．

問 6.7　X_k $(k = 1, \ldots, n)$ がパレート分布に従うとき，X_k $(k = 1, \ldots, n)$ の最大値の確率分布を求めよ．$n \to \infty$ のとき最大値の分布の挙動を考察せよ．

(略解) パレート分布のとき

$$P(\max_{1 \le i \le n} X_i \le z) = [P(X_i \le z)]^n = \left[1 - \left(\frac{x_0}{z}\right)^{\alpha}\right]^n$$

となるので

$$P(\max_{1 \le i \le n} X_i \le n^{1/\alpha} z x_0) = \left[1 - \frac{1}{n}\left(\frac{1}{y}\right)^{\alpha}\right]^n \to e^{-(1/y)^{\alpha}}$$

となる．

問 6.8　一様分布の場合に (6.34) を求めよ．

(略解)　順序統計量 $X_{(k)}$ が従う密度関数は

$$f_k(x) = \frac{1}{B(k, n-k+1)} x^{k-1}(1-x)^{n-k}$$

であるから 1 次積率 (モーメント)・2 次積率 (モーメント) を求めればよい．例えば

$$\begin{aligned}
\mathbf{E}[X_k] &= \frac{1}{B(k, n-k+1)} \int_0^1 x^k (1-x)^{n-k} dx \\
&= \frac{B(k+1, n-k+1)}{B(k, n-k+1)} \\
&= \frac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)} \times \frac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)} = \frac{k}{n+1}
\end{aligned}$$

となる．

第　7　章

問 7.1　母集団として正規分布を仮定するとき (例 7.1)，$\mathbf{E}[s_n^2] = \sigma^2$ となることを示せ．s_n^2 のばらつきを分析せよ．

(略解) 自由度 k の χ^2 分布の積率母関数は

$$M(\theta) = (1 - 2\theta)^{-k/2}$$

より χ^2 分布の 1 次積率 μ_1 と 2 次積率 μ_2 はそれぞれ

$$\begin{aligned}
\mu_1 &= M'(\theta)\,|_{\theta=0} = k(1-2\theta)^{-\frac{k}{2}-1}\,|_{\theta=0} \\
\mu_2 &= M''(\theta)\,|_{\theta=0} = k(k+2)(1-2\theta)^{-\frac{k}{2}-2}\,|_{\theta=0} = k^2 + 2k.
\end{aligned}$$

したがって分散は $\mathbf{V}(X) = \mu_2 - \mu_1^2 = 2k$. $(n-1)s^2/\sigma^2$ は自由度 $n-1$ の χ^2 分布に従うので

$$\mathbf{V}\left(\frac{(n-1)s^2}{\sigma^2}\right) = 2(n-1)$$

より

$$\mathbf{V}(s^2) = \frac{2\sigma^4}{n-1}$$

を得る．

問 7.2　例 7.4 と例 7.5 における十分統計量を導け．

(略)

問 7.3　例 7.11 における不偏推定量 s_n^2 の不偏性と最尤推定量のバイアス，平均 2 乗誤差についての主張を示せ．

(略)

問 7.4　独立な確率変数列 X_i $(i=1,\ldots,n)$ が正規分布 $N(\mu,\sigma^2)$ に従うとき，μ が既知 (例えば 0) と未知の場合の σ^2 の UMVU を求めよ．

(略解) σ^2 の推定量を $\hat{\sigma}^2=\sum(X_i-\mu)^2/n$ とすると，$\mathbf{E}[\hat{\sigma}^2]=\sigma^2$ となり $\hat{\sigma}^2$ は不偏．また $X_i \sim i.i.d.N(\mu,\sigma^2)$ であるので $\sum(X_i-\mu)^2/\sigma^2 \sim \chi_n^2$，すなわち $n\hat{\sigma}^2/\sigma^2 \sim \chi_n^2$．よって $\mathbf{V}(\hat{\sigma}^2)=2\sigma^4/n$ となる．後は $\mathbf{V}(\hat{\sigma}^2)$ がクラメール–ラオの不等式の下限と等しいことを示せばよい．$X=(X_i)_{i=1}^n$ の同時分布 $f(x,\sigma^2)$ のフィッシャー情報量を求めるには，対数尤度関数が

$$l(\sigma^2,X_1)=-\frac{1}{2}\left(\log 2+\log\pi+\log\sigma^2\right)-\frac{(X_1-\mu)^2}{2\sigma^2}$$

となるから，

$$\frac{\partial^2 l}{\partial(\sigma^2)^2}=\frac{1}{2\sigma^4}-\frac{(X_1-\mu)^2}{\sigma^6}.$$

よって

$$\mathbf{E}\left[-\frac{\partial^2 l}{\partial(\sigma^2)^2}\right]=\frac{1}{2\sigma^4}$$

となる．よってクラメール–ラオの不等式の下限は $\frac{2\sigma^4}{n}$ となって不偏推定量の分散と一致する．以上より，母平均 μ が既知のとき $\sum(X_i-\mu)^2/n$ は UMVU である．

問 7.5　互いに独立な確率変数列 X_i $(i=1,\ldots,n)$ が母数 λ の指数分布に従うとする．母数 λ $(\lambda>0)$ の推定量として $\hat{\lambda}=(1/m)\sum_{i=1}^m X_i$ (m は整数，$0<m\le n$) とするとき，この推定量の統計的意味での妥当性について議論せよ．

(略解) n 個の平均ではなく n より小さい m $(0<m<n)$ 個の平均を用いると情報損失が生じるということであった．m 個の標本平均の分散は

$$\mathbf{V}\left[\frac{1}{n}\sum_{i=1}^n X_i\right]=\frac{\lambda}{n}<\mathbf{V}\left[\frac{1}{m}\sum_{i=1}^m X_i\right]=\frac{\lambda}{m}$$

となり，CR 下限を達成しない．n 個の標本平均は CR 下限を達成するので一様最小分散不偏 (UMVU) 推定量である．

問 7.6　互いに独立な確率変数列 X_i $(i=1,\ldots,n)$ が母数 λ のポワソン分布に従うとする．(i) 母数 λ の UMVU を求めよ．(ii) 母数 λ $(\lambda>0)$ の推定量として $\hat{\lambda}=(1/m)\sum_{i=1}^m X_i$ (m は整数,$0<m\le n$) とするとき，この推定量の妥当性と改善可能性について議論せよ．

(略解) 問 7.5 と同様．

問 7.7　例 7.11 において $(n-1)s_n^2/\sigma^2$ が $\chi^2(n-1)$ (ガンマ分布 $Gamma(n/2,2)$) に従うことを利用すると σ の不偏推定量が $s\sqrt{n-1}\Gamma((n-1)/2)/[\sqrt{2}\Gamma(n/2)]$ となることを示せ．

(略解) 確率変数 X がガンマ分布 $Gamma(\alpha,\beta)$ に従うとき

$$\begin{aligned}\mathbf{E}[\sqrt{X}]&=\frac{1}{\Gamma(\alpha)}\Big(\frac{1}{\beta}\Big)^{\alpha}\int_0^\infty \sqrt{x}x^{\alpha-1}e^{-x/\beta}dx\\&=\frac{1}{\Gamma(\alpha)}\Big(\frac{1}{\beta}\Big)^{\alpha-\alpha+1/2}\int_0^\infty z^{\alpha-1/2}e^{-z}dz\\&=\frac{1}{\Gamma(\alpha)}\beta^{1/2}\Gamma(\alpha+1/2)\end{aligned}$$

となる．ここで $X=(n-1)s_n^2/\sigma^2$, $\alpha=(n-1)/2, \beta=2$ とすると $\mathbf{E}[\sqrt{X}]=\sqrt{2}\Gamma(n/2)/\Gamma((n-1)/2)$ より

$$\mathbf{E}\left[\frac{\Gamma((n-1)/2)}{\sqrt{2}\Gamma(n/2)}\sqrt{n-1}s_n\right]=\sigma$$

という結果を得る．

問 7.8　(i) 平均 2 乗積分誤差 (7.62) を h について最小化した最適な h^* を求めよ．(ii) 漸近分散項 $\mathbf{V}(\hat{f}_n)$ の第 2 項が n が大きいとき相対的に無視できることを示せ．

(略解) 分散項を評価すると $((x-t)/h=y$ と変換)

$$\begin{aligned}
\mathbf{V}&=\frac{1}{nh^2}\mathbf{E}\left[K\left(\frac{x-X_i}{h}\right)-\mathbf{E}\left(K\left(\frac{x-X_i}{h}\right)\right)\right]^2\\
&=\frac{1}{nh^2}\int\left[K\left(\frac{x-t}{h}\right)-\mathbf{E}\left(K\left(\frac{x-t}{h}\right)\right)\right]^2 f(t)dt\\
&=\frac{1}{nh}\int\left[K(y)-\int K(z)f(x-zh)dz\right]^2 f(x-yh)dy\\
&=\frac{1}{nh}\left[\int K(y)^2 f(x-hy)dy\right]-\frac{1}{nh^2}\left[h^2\int K(z)f(x-zh)dzf(x-yh)dy\right]^2\\
&=\frac{1}{nh}\left[\int K(y)^2(f(x)-f^{'}(x)hy)dy\right]-O\left(\frac{1}{n}\right)\\
&\sim\frac{1}{nh}\left[\int K(y)^2 f(x)dy\right]
\end{aligned}$$

と評価できる．関数 $A(h)=c_1/(nh)+c_2h^4$ (c_1,c_2 は定数) を h について最小化すると $n\sim h^{-5}$ を得る．

第　8　章

問 8.1　$X_1,\ldots,X_n$ が独立に $N(0,\sigma^2)$ に従うとき帰無仮説 $H_0:\sigma^2\le 1$，対立仮説 $H_1:\sigma^2>1$ に対する検定方式を与えよ．

(略解) 同時密度関数 $f(\mathbf{x}|\sigma^2)=(1/\sigma^2)^{n/2}\exp[-(1/2\sigma^2)\sum_{i=1}^n x_i^2]$ より密度関数比の条件

$$\frac{f(\mathbf{x}|\sigma_1^2)}{f(\mathbf{x}|\sigma_0^2)}=\left(\frac{\sigma_0^2}{\sigma_1^2}\right)^{n/2}\exp\left[-\frac{1}{2}\left(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_0^2}\right)\sum_{i=1}^n x_i^2\right]>c$$

は $\sum_{i=1}^n x_i^2>c^{'}$ となる ($c,c^{'}$ は定数とした)．$\sigma=1$ のとき $\sum_{i=1}^n X_i^2\sim\chi^2(n)$ を利用すればよい．

問 8.2　例 8.4 における尤度比検定統計量を導き，n が大きいとき帰無仮説および対立仮説における漸近分布を求めよ．

(略)

問 8.3　X が幾何分布 $p(x)=p(1-p)^x$　$(x=0,1,\ldots)$ に従うとき母数 p の推定方法と帰無仮説 $H_0:p=p_0, H_1:p>p_0$ の検定方法を考察せよ．母数 p についての十分統計量，フィッシャー情報量を求めよ．

(略解) 尤度比は

$$\frac{p(x_1,\ldots,x_n|p_1)}{p(x_1,\ldots,x_n|p_0)}=\frac{(1-p_1)^{\sum_{i=1}^n x_i}p_1^{nr}}{(1-p_0)^{\sum_{i=1}^n x_i}p_0^{nr}}>c$$

より

$$\left[\frac{1-p_1}{1-p_0}\right]^{\sum_{i=1}^n x_i}>c^{'}$$

と書ける．$1-p_0>1-p_1>0$ より領域

$$\sum_{i=1}^n x_i<c^{''}(\text{定数})$$

が UMP である ($c,c^{'},c^{''}$ は定数とした)．なお負の 2 項分布の特性関数は

$$\psi(t) = \sum_{x=0}^{\infty} e^{itx}(r+x-1,x)(1-p)^x p^r$$
$$= \sum_{x=0}^{\infty}(r+x-1,x)[(1-p)e^{it}]^x p^r$$
$$= \frac{p^r}{[1-(1-p)e^{it}]^r} = \left[1-(e^{it}-1)\frac{q}{p}\right]^{-r}$$

となる (ここで $q=1-p$ とした)．したがって $X_1+\cdots+X_n$ の特性関数は

$$[\psi(t)]^n = \left[1-(e^{it}-1)\frac{q}{p}\right]^{-nr}$$

より再生性があるので $X_1+\cdots+X_n \sim NB(nr,p)$ より UMP 検定の領域を求められる．

問 8.4　互いに独立な確率変数列 X_i $(i=1,\ldots,n)$ が母数 λ の指数分布に従うとする．帰無仮説 $H_0:\lambda=\lambda_0$ (既知の値) を対立仮説 $H_1:\lambda\neq\lambda_0$ に対して検定する問題を考察する．統計的に妥当と思われる検定方法を挙げその理由を述べよ．

問 8.5　互いに独立な確率変数列 X_i $(i=1,\ldots,n)$ が母数 λ のポワソン分布に従うとする．帰無仮説 $H_0:\lambda=\lambda_0$ (既知の値) を対立仮説 $H_1:\lambda\neq\lambda_0$ に対して検定することを考える．妥当と思われる検定方法を挙げその理由を述べよ．

(略解) ポワソン分布の確率比

$$\frac{p(\mathbf{x}|\lambda_1)}{p(\mathbf{x}|\lambda_0)} = e^{-(\lambda_1-\lambda_0)}\left(\frac{\lambda_1}{\lambda_0}\right)^{\sum_{i=1}^n x_i} > c$$

は $\sum_{i=1}^n x_i > c^{'}$ となる $(c, c^{'}$ は定数とした)．$\lambda=1$ のとき $\sum_{i=1}^n X_i^2 \sim Po(n)$ を利用すればよい．ポワソン分布 $\sum_{i=1}^n X_i$ の再生性は特性関数により示せる．

問 8.6　ウェルチ検定において漸近的議論を用いて $N(0,1)$ による検定を利用することができるか否か考察せよ．

(略)

第　9　章

問 9.1　標本が正規分布 $X\sim N(0,\sigma^2)$ から互いに独立に得られるとする．σ^{-2} の事前分布にガンマ分布を仮定したときの事後分布を分析せよ．

(略解) X がガンマ分布 $Gamma(\alpha,\beta)$ に従うとき $Y=X^{-1}$ は逆ガンマ分布 $IG(\alpha,\beta)$ に従うという．この変換のヤコビアンは y^{-2} であり，密度関数は

$$g(y|\alpha,\beta) = \frac{1}{\Gamma(\alpha)}\left(\frac{1}{\beta}\right)^{\alpha}\left(\frac{1}{y}\right)^{\alpha+1} e^{-\frac{1}{\beta}\frac{1}{y}}$$

で与えられる．

σ^2 の事前分布 $p(\sigma^2)$ が $IG(\alpha_0,\beta_0)$ に従い X_i $(i=1,\ldots,n)$ が $N(0,\sigma^2)$ のとき事後分布

$$p(\sigma^2|x_1,\ldots,x_n) \sim p(\sigma^2)\prod_{i=1}^n n(x_i|0,\sigma^2)$$

を求めると $IG(\alpha_1,\beta_1)$ に従うことがわかる．ここで

$$\alpha_1 = \alpha_0 + \frac{n}{2}\,,\quad \beta_1^{-1} = \beta_0^{-1} + \frac{\sum_{i=1}^n x_i^2}{2}$$

となる．

問 9.2　定理 8.1 の L 集合 (リスク集合) の凸性を示せ．さらに例 9.5 で述べられているベイズ解の説明を確認せよ．

(略)

問 9.3 定理 9.2 の証明で用いた式変形をより丁寧に確認せよ．

(略)

問 9.4 次に引用するエコノミストが書いた文書を読み，数理統計学の観点からコメントせよ．「いったい，なぜ人間は想起しやすさを過度に重視するのだろうか．……たった一度，されど一度．統計学が語るところの大数の法則に対して小数の法則と呼んでもよい．大数の法則とは何度も実験を繰り返せば，観察される頻度はその確率に近づくことを表した法則である．統計学が何度も繰り返すことができる事柄を扱うのに対して，我々は繰り返しのきかない人生を生きている．繰り返しのきかない人生だからこそ，有限の経験に頼ってしまうのであろう．」

ヒント：経済学者の解説の解釈なので一意の正解は存在しない．数理統計学の立場からは大数の法則 (law of large numbers) の意味の解釈，少数の法則 (law of small numbers) の意味の解釈, 例えば 2 項分布のポワソン近似などの説明が考えられる．確率の解釈としては市場確率，ベイズ確率，などは必ずしも頻度論的確率・統計の解釈は必要としない．

問 9.5 n 個の標本がポワソン分布 $X \sim P_o(\lambda)$ から独立に得られ，母数 λ についての事前分布がガンマ分布 $Gamma(\alpha_0, 1)$ とする．

(i) 標本が得られたときの母数 λ に関する事後分布を求め，合理的と考えられる規準に基づき λ の推定値を導け．

(ii) (保守的な上司の) 損失関数として，推定値が真の売り上げ値より大きいときの損失が推定値が真の売り上げ値より小さいときの損失より大きい，つまり $l(\lambda, \hat{\lambda}_n) = 3(\hat{\lambda}_n - \lambda)_+ + (\hat{\lambda}_n - \lambda)_-$ とする ($|x|_+ = x \ (x \geq 0); 0 \ (x < 0)$, $|x|_- = -x \ (x < 0); 0 \ (x \geq 0)$ である)．このとき母数 λ の推定値をどのように上司に報告したらよいか．

(略解) 同時分布

$$p(x_1, \ldots, x_n|\lambda) = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}$$

および事前分布

$$\xi(\lambda) = \frac{1}{\Gamma(\alpha_0)} \left(\frac{\lambda}{\beta_0}\right)^{\alpha_0 - 1} e^{-\lambda/\beta_0} \frac{1}{\beta_0}$$

より事後分布は

$$\begin{aligned} \xi(\lambda|x_1, \ldots, x_n) &\sim \xi(\lambda) p(x_1, \ldots, x_n|\lambda) \\ &= e^{-(n+1/\beta_0)\lambda} \lambda^{\alpha_0 + \sum_i x_i - 1} \end{aligned}$$

となる．これは $\Gamma(\alpha_1, \beta_1), \alpha_1 = \alpha_0 + \sum_i x_i, 1/\beta_1 = n + 1/\beta_0$ を意味する．ガンマ分布の性質から事後分布の期待値は

$$\mathbf{E}[\lambda|x_1, \ldots, x_n] = \alpha_1 \beta_1 \ ,$$

分散は

$$\mathbf{V}[\lambda|x_1, \ldots, x_n] = \alpha_1 \beta_1^2$$

となる．

損失関数より事後分布によりリスクを評価しよう．

$$\begin{aligned} R(d) &= \mathbf{E}[(\mu - d)_+ + 3(\mu - d)_-] \\ &= \int_d^\infty (\mu - d)\xi(\mu)d\mu + \int_{-\infty}^d (-3)(\mu - d)\xi(\mu)d\mu \\ &= -d\left[\int_d^\infty \xi(\mu)d\mu - 3\int_{-\infty}^d \xi(\mu)d\mu\right] + \left[\int_d^\infty \mu\xi(\mu)d\mu - 3\int_{-\infty}^d \mu\xi(\mu)d\mu\right] \end{aligned}$$

と書ける．ここで例えば $\int_d^\infty \xi(\mu)d\mu = 1-F(d)$, $\int_{-\infty}^d \xi(\mu)d\mu = F(d)$ (F は分布関数) となっていることに注意する．リスクを最小化するために，$R(d)$ を d について微分すると

$$\frac{\partial R(d)}{\partial d} = -\left[\int_d^\infty \xi(\mu)d\mu - 3\int_{-\infty}^d \xi(\mu)d\mu\right] - d[-\xi(d)-3\xi(d)] + [-d\xi(d)-3d\xi(d)] = 0$$

より

$$\int_d^\infty \xi(\mu)d\mu = 3\int_{-\infty}^d \xi(\mu)d\mu$$

を得る．これは d が事後分布の下側 25% 点となることを意味する．直観的にはこの上司は過大評価に対しては厳しいが，過小評価には甘いので点推定値はかなり低めに報告しておけば大過はより少ない．なお，通常の議論では過大評価も過小評価も同等に損失を考えると事後分布の中央値 (メディアン) がベイズ (点推定) 解となる．なお，n ($n=4m+1$，m を整数としておくと分かりやすい) 標本上での評価関数

$$l_n(d) = \sum_{i=1}^n [(X_i-d)_+ + 3(X_i-d)_-]$$

を最小化すると $[n/4]+1$ 番目の順序統計量が得られる．こうして得られる標本分位点を用いることも自然であろう．ちなみに $\sum_{i=1}^n |(X_i-d||$ を最小化するのは中央値 (median, メディアン) である．

第　10　章

問 10.1　式 (10.1) の最小化により係数 b_0, b_1 が一意に定まらないことがあるか．

(**略解**) 2 乗損失関数

$$L(\beta_0,\beta_1) = \sum_{i=1}^n (y_i-\beta_0-\beta_1 z_i)^2$$

を最小化解は $z_i = c$ (定数) のとき不定となる．

問 10.2　単回帰 ($k=2$) のとき推定量を $b_0 = \sum_{j=1}^n a_j y_j, b_1 = \sum_{j=1}^n c_j y_j$ (a_j, c_j は実数列) としてガウス–マルコフの定理 (最小 2 乗推定量は最小線形不偏推定量) を示せ．

(**略解**) 2 次元ベクトル $\mathbf{b} = (b_0, b_1)^{'}$ とすると問 10.3 に帰着される．

問 10.3　式 (10.13) を導け．さらに $\mathbf{c}_j\mathbf{y}$ ($j=1,\ldots,p$) の期待値，分散・共分散を導き不偏性の条件を導き，定理 10.1 を示せ．

(**略解**) 最小 2 乗法による正規方程式より

$$\mathbf{Z}^{'}\mathbf{Z}\mathbf{b} = \mathbf{Z}^{'}\mathbf{y} = \mathbf{Z}^{'}(\mathbf{Z}\boldsymbol{\beta}+\mathbf{u}) = \mathbf{Z}^{'}\mathbf{Z}\boldsymbol{\beta} + \mathbf{Z}^{'}\mathbf{u}$$

となる．仮定より期待値 $\mathbf{E}[\mathbf{b}] = \mathbf{0}$, $\mathbf{Z}^{'}\mathbf{Z}(\mathbf{b}-\boldsymbol{\beta}) = \mathbf{Z}^{'}\mathbf{u}$ より $\mathbf{V}[\mathbf{Z}^{'}\mathbf{Z}(\mathbf{b}-\boldsymbol{\beta})] = \mathbf{Z}^{'}\mathbf{E}[\mathbf{u}\mathbf{u}^{'}]\mathbf{Z} = \sigma^2\mathbf{Z}^{'}\mathbf{Z}$. したがって $\mathbf{b}$ の分散・共分散行列は $\mathbf{V}[\mathbf{b}] = \sigma^2(\mathbf{Z}^{'}\mathbf{Z})^{-1}$ である．

任意の線形推定量 $\mathbf{b}_c = \mathbf{C}\mathbf{y}$ が不偏性を持つには任意の $\boldsymbol{\beta}$ に対して

$$\mathbf{E}[\mathbf{b}_c] = \mathbf{C}\mathbf{X}\boldsymbol{\beta} = \boldsymbol{\beta}$$

より条件 $\mathbf{C}\mathbf{X} = \mathbf{I}_k$ が必要．この条件の下で分散・共分散行列 $\mathbf{V}[\mathbf{b}_c] = \sigma^2\mathbf{C}\mathbf{C}^{'}$ の最小化の解は最小 2 乗法となる．なぜならば $\mathbf{Q} = (\mathbf{Z}^{'}\mathbf{Z})^{-1}\mathbf{X}^{'}$ とすると

$$(\mathbf{Q}-\mathbf{C})(\mathbf{Q}-\mathbf{C})^{'} = \mathbf{Q}\mathbf{Q}^{'} + \mathbf{C}\mathbf{C}^{'} - \mathbf{Q}\mathbf{C}^{'} - \mathbf{C}\mathbf{Q}^{'} = \mathbf{C}\mathbf{C}^{'} - \mathbf{Q}\mathbf{Q}^{'}$$

より非負定符号行列となる．

問 10.4　$k=2$ のとき仮定 (A4) の下で統計量 $T = (b_1-\beta_1)/\sqrt{a^{22}\hat{\sigma}^2}$ が自由度 $n-2$ の t 分布に従うことを示せ (a^{22} は行列 $(\mathbf{Z}^{'}\mathbf{Z})^{-1}$ の $(2,2)$ 要素を意味する)．

(**略解**) 残差ベクトル $\hat{\mathbf{u}} = \mathbf{y} - \mathbf{X}\mathbf{b} = [\mathbf{I} - \mathbf{X}(\mathbf{X}^{'}\mathbf{X})^{-1}\mathbf{X}]\mathbf{u}$ よりベクトル $\mathbf{b}$ と $\hat{\mathbf{u}}$ は直交している．したがって互いに無相関となる．さらに射影行列の性質を利用すると 2 次形式 $\hat{\mathbf{u}}^{'}\hat{\mathbf{u}}/\sigma^2$ は自由度 $n-k$ の χ^2 分布に

従うことがわかる．したがって t 統計量は $t(n-k)$ に従う．

問 10.5　解 (10.16) および (10.18) を導け．

(略解) 関数 $(1/n)L_d$ を b について微分すると 2 次方程式 $s_{12}b^2+(s_{22}-s_{11})b-s_{12}=0$ が得られる．

他方，$L_d = (1,-b)\mathbf{S}_y(1,-b)'/(1+b^2)$ は b についての 2 次形式なのでラグランジュ形式を $L = (1,-b)\mathbf{S}_y(1,-b)' - \lambda(1+b^2)$ として b について最適化すると $(0,-1)\mathbf{S}_y(1,-b)'+2\lambda b=0$ が得られる．これより 2 次方程式

$$\lambda^2-(s_{11}+s_{22})\lambda+s_{11}s_{22}-s_{12}^2=0$$

の 2 根の中で小さい方の根は $\lambda_1=[s_{11}+s_{22}-\sqrt{D}]/2,\ D=(s_{11}+s_{22})^2-4(s_{11}s_{22}-s_{12}^2)$ で与えられる．この最小根に対応する b は (10.15) で与えられる．

問 10.6　(10.18) の最小化問題の解を求め，$\mathbf{b}$ がスカラーの場合に直交回帰の解に一致することを示せ．

(略解) 線形代数における 2 次形式の比の最小化問題と同一であることを利用すると結果が得られる．

問 10.7　(10.28) および (10.29) の最小化問題の解を求めよ．

(略解) ベクトル $\mathbf{c}$ についての 2 次形式および 2 次形式の比であることに注目する．2 次形式の比はラグランジュ乗数 λ を利用して最適化すればよい．

第　11　章

問 11.1　(11.19) を示せ．

(略)

問 11.2　(11.38) を示せ．

(略)

問 11.3　本文の説明を利用して (11.39) を示せ．

(略)

問 11.4　(11.40) を示せ．

(略解)

$$\tau=2P((X_1-X_2)(Y_1-Y_2)>0)-1$$

とする．

$$\begin{aligned}P(X_1>X_2,Y_1>Y_2)&=\int\int P(X_2\le v,Y_2\le w)dC_1(F_1(v),F_2(w))\\&=\int\int C_2(F_1(v),F_2(w))dC_1(F_1(v),F_2(w))\end{aligned}$$

である．他方

$$\begin{aligned}P(X_1<X_2,Y_1<Y_2)&=\int\int P(X_2\le v,Y_2\le w)dC_1(F_1(v),F_2(w))\\&=\int\int[1-F_1(v)-F_2(w)+C_2(F_1(v),F_2(w))]dC_1(F_1(v),F_2(w))\\&=\int\int_{I^2}[1-u_1-u_2+C_2(u_1,u_2)]dC_1(u_1,u_2)\\&=1-\frac{1}{2}-\frac{1}{2}+\int\int_{I^2}C_2(u_1,u_2)dC_1(u_1,u_2)\end{aligned}$$

となる．これより

$$\tau=4\int\int_{I^2}C_2(u_1,u_2)dC_1(u_1,u_2)-1$$

が得られる．

問 11.5　2 次自己回帰過程 (AR(2)，式 (11.1) において $p=2$) が (弱) 定常的となる係数の条件を求めよ．ここで (弱) 定常性は $\mathbf{E}(Y_t)=\mathbf{E}(Y_{t-1})=\mathbf{E}(Y_{t-2})=\cdots$, $\mathbf{E}(Y_t^2)=\mathbf{E}(Y_{t-1}^2)=\mathbf{E}(Y_{t-2}^2)=\cdots$,

$\mathbf{E}(Y_tY_{t-1}) = \mathbf{E}(Y_{t-1}Y_{t-2}) = \mathbf{E}(Y_{t-2}Y_{t-3}) = \cdots$ などを意味する．1 次移動平均過程 ((11.3) において $q=1$) のときは定常性の条件は何か．

(略解) (i) AR(2) モデルの定常条件は特性方程式

$$g(\lambda) = \lambda^2 - \beta_1\lambda - \beta_2 = 0$$

の根の絶対値が 1 より小さい範囲を求めればよい．

(ii) 有限次数の MA モデルは常に定常過程である．

問 11.6　AR(1), AR(2) に対し 1 期先予測・2 期先予測量を構成せよ．

問 11.7　ピアソン相関係数 (式 (11.28)) に関連して $(X_1, Y_1) = (e^Z, e^{\sigma Z})$, $(X_2, Y_2) = (e^Z, e^{-\sigma Z})$, $Z \sim N(0,1)$ のとき (X_1, Y_1) および (X_2, Y_2) の相関係数がそれぞれ $(e^\sigma - 1)/\sqrt{(e-1)(e^{\sigma^2}-1)}$, $(e^{-\sigma} - 1)/\sqrt{(e-1)(e^{\sigma^2}-1)}$ となることを示し，(11.28) について本文で述べている主張を確かめよ．

(略解) $(X_1, X_2) \sim (e^Z, e^{\sigma Z})$, $Z \sim N(0,1)$ とする．このとき

$$\begin{aligned}
\mathbf{E}[X_2] &= \int e^{\sigma z}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz \\
&= \int e^{\sigma^2/2}\frac{1}{\sqrt{2\pi}}e^{-(z-\sigma)^2/2}dz \\
&= e^{\sigma^2/2}\ ,
\end{aligned}$$

$$\begin{aligned}
\mathbf{E}[X_2^2] &= \int e^{2\sigma z}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz \\
&= \int e^{2\sigma^2}\frac{1}{\sqrt{2\pi}}e^{-(z-2\sigma)^2/2}dz\ ,
\end{aligned}$$

$$\begin{aligned}
\mathbf{E}[X_1X_2] &= \int e^{(1+\sigma)z}\frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz \\
&= \int e^{(1+\sigma)^2/2}\frac{1}{\sqrt{2\pi}}e^{-(z-(1+\sigma))^2/2}dz
\end{aligned}$$

となる．したがって相関係数は

$$\rho = \frac{e^{(1+\sigma)^2/2} - e^{\sigma^2/2}e^{1/2}}{\sqrt{(e^{2\sigma^2} - e^{\sigma^2})(e^2 - e)}}$$

より

$$\rho = \frac{e^\sigma - 1}{\sqrt{(e-1)(e^{\sigma^2}-1)}}$$

が得られる．

問 11.8　楕円分布族と正規分布の 1〜4 次の積率を評価・比較せよ．

(略)

問 11.9　$p=3$ の場合に定理 11.3 を確かめよ．

(略)

問 11.10　(11.47) に続く 2 次元正規分布の裾確率評価を確かめよ．

(略解)　定数 $\alpha = \frac{2}{\sqrt{2(1+\rho)}}$ とおく．2 次元正規分布の性質より

$$\begin{aligned}
P(1-s, 1-s) &\le P(X_1 + X_2 > 2\Phi^{-1}(1-s)) \\
&= P\left(\frac{X_1+X_2}{\sqrt{2(1+\rho}} > \alpha\Phi^{-1}(1-s)\right) \\
&= 1 - \Phi(\alpha\Phi^{-1}(1-s)) \\
&\sim \frac{1}{\sqrt{2\pi}}\left[\frac{1}{\alpha\Phi^{-1}(1-s)}\right]e^{-\frac{1}{2}[\alpha\Phi^{-1}(1-s)]^2} \\
&\sim \frac{1}{\sqrt{2\pi}}\left[\frac{-\frac{\alpha^2}{2}z^2}{\alpha z}\right]
\end{aligned}$$

である．ここで正規分布の裾確率の評価を用いて z が大きいとき

$$s = 1 - \Phi(z) \sim \left(\frac{1}{\sqrt{2\pi}}\right)\left(\frac{1}{z}\right)e^{-z^2/2}$$

を利用した．したがって

$$\frac{P(1-s, 1-s)}{s} = \frac{1}{\alpha}[e^{-\frac{z^2(\alpha^2-1)}{2}}] \to 0 \quad (z \to \infty, s \to 0)$$

となる (ここで $|\alpha| = |\sqrt{2/(1+\rho)}| > 1$ を利用した).

問 11.11　国際的な銀行業・保険業に対して当局は「銀行に対して損失分布 (その価値が将来に変動する資産を保有する場合には，発生する可能性がある資産価値の損失分布) に対して対数値の変化額に正規分布を当てはめて 1% 分位点を管理する (VaR (value-at-risk) 基準と呼ばれる)」ことを推奨したことがある．この方法の妥当性について数理統計学的見地よりコメントし，妥当でないと考えるならば対案を示せ．

ヒント：例えば統計的極値論の応用が考えられる．