

目次

第 1 章	線型推測論	2
第 2 章	統計的決定理論	3
2.1	枠組み	3
2.2	十分性と完備性	4
2.2.1	十分統計量	4
2.3	指数型分布族	4
2.4	統計的推定	4
第 3 章	統計的仮説検定	5
3.1	枠組みと例	5
3.1.1	枠組み	5
3.1.2	ランダム化検定	5
3.1.3	形式化	6
第 4 章	大標本理論	7
4.1	最尤推定	7
4.2	一様対数の法則	8
4.3	最小コントラスト推定	9
4.4	M -推定量	9
4.5	推定量の漸近正規性	9
4.6	ワンステップ推定量	9
4.7	M -推定量の存在	9
4.8	頑健推定	10
4.9	尤度比検定	10
4.10	多項分布の検定	10
4.11	尤度比確率場の局所漸近構造	10
4.12	情報量規準	10
4.13	密度推定	10
4.14	U -統計量	10
第 5 章	漸近展開とその応用	11
	参考文献	12

第 1 章

線型推測論

線型推測論が統計モデルの線形代数である .

第2章

統計的決定理論

統計的決定問題が設定されたとき，決定問題の下で起こる確率現象を解明するのが数理統計学の課題であって，特定の決定関数を無条件に是とするものではない．

2.1 枠組み

決定すべきは関数 = 射である．その文脈は8組で表現できる．

定義 2.1.1 (statistical decision problem, sample space, decision / action space, loss function, nonrandomized, randomized decision function (Abraham Wald)). 次の8組 $(\mathcal{X}, \mathcal{A}, \mathcal{P}, \Theta, \mathcal{D}, \mathcal{B}, W, \Delta)$ を統計的決定問題という．

- (1) $(\mathcal{X}, \mathcal{A})$ は可測空間で，標本空間という．
- (2) Θ はパラメータの空間で， $\mathcal{P} := (P_\theta)_{\theta \in \Theta}$ は $(\mathcal{X}, \mathcal{A})$ の確率分布の族．
- (3) $(\mathcal{D}, \mathcal{B})$ は可測空間で，決定空間または行動空間という．
- (4) 第二変数について可測な関数 $W: \Theta \rightarrow \text{Meas}(\mathcal{D}, \mathbb{R}_+)$ を損失関数という．
- (5) Δ は決定関数の族とする．
 - (a) 可測写像 $\delta: \mathcal{X} \rightarrow \mathcal{D}$ を非確率的決定関数という．
 - (b) 第一変数については $(\mathcal{D}, \mathcal{B})$ 上の確率測度，第二変数については \mathcal{X} 上の \mathcal{A} -可測関数となる関数 $\delta: \mathcal{B} \times \mathcal{X} \rightarrow [0, 1]$ を確率的決定関数という．

注 2.1.2. 非確率決定関数 $\delta: \mathcal{X} \rightarrow \mathcal{D}$ が与えられたとき，任意の $B \in \mathcal{B}$ に対して $\tilde{\delta}(B|x) := 1_{\{\delta(x) \in B\}}$ で $\tilde{\delta}: \mathcal{B} \times \mathcal{X} \rightarrow [0, 1]$ が定まるから，確率的決定関数の方がより一般的な設定となる．

例 2.1.3 (平均値の推定量の決定). $(\mathcal{X}, \mathcal{A}) := (\mathbb{R}^n, \mathcal{B}_n)$, $\mathcal{P}_1 := \left\{ P = P_*^n \in \mathcal{M}(\mathbb{R}^n) \mid P_* \text{ は } \mathbb{R} \text{ 上の確率分布で } \int_{\mathbb{R}} x^2 P_*(dx) < \infty \right\}$, $\Theta := \mathcal{P}_1$ とする．決定空間は平均値の全体としたいから $(\mathcal{D}, \mathcal{B}) := (\mathbb{R}, \mathcal{B}_1)$ ．推定量 $\Delta := \{T_1, T_2, T_3\}$ はそれぞれ $T_1 := X_1$, $T_2 := \sum_{j=1}^n \frac{X_j}{n}$, $T_3 := X_n$ という非確率的決定関数の族で，損失関数は $W(P, a) := (a - \mu_1(P))^2$ とする．

すると，損失関数 W から定まる危険関数は

$$\begin{aligned} R(P, T_i) &= \int W(P, T_i(x_1, \dots, x_n)) P(dx_1, \dots, dx_n) \quad (P \in \mathcal{P}_1) \\ &= \text{Var}_P[T_i] \end{aligned}$$

となる．こうして， $R(P, T_1) = R(P, T_3) = \text{Var}_P[X_1]$, $R(P, T_2) = \frac{\text{Var}_P[X_1]}{n}$ となり， T_2 が危険関数 R を最小にするとわかる．

定義 2.1.4 (risk function, admissible). 危険関数 $R: \Theta \times \Delta \rightarrow \mathbb{R}$ を次のように定める．

- (1) 確率的な決定関数 $\delta: \mathcal{B} \times \mathcal{X} \rightarrow [0, 1]$ については， $R(\theta, \delta) := \int_{\mathcal{X}} \int_{\mathcal{D}} W(\theta, a) \delta(da|x) P_\theta(dx)$ と定める．
- (2) 非確率的な決定関数 $\delta: \mathcal{X} \rightarrow \mathcal{D}$ については， $R(\theta, \delta) := \int_{\mathcal{X}} W(\theta, \delta(x)) P_\theta(dx)$ と定める．

2つの決定関数 $\delta_1, \delta_2 \in \Delta$ について，

- (1) $\forall \theta \in \Theta R(\theta, \delta_1) \leq R(\theta, \delta_2)$ が成り立つとき, δ_1 と δ_2 は同程度に良い決定関数であるという.
- (2) 同程度に良い決定関数がさらに $\exists \theta_0 \in \Theta R(\theta_0, \delta_1) < R(\theta_0, \delta_2)$ を満たすとき, δ_1 は δ_2 より一様に良い決定関数であるという.
- (3) 「一様に良い」という関係は順序を定める. Δ に最大元が存在せず, 極大元 δ_* のみが存在するとき, 許容的であるという.

2.2 十分性と完備性

2.2.1 十分統計量

2.3 指数型分布族

2.4 統計的推定

第3章

統計的仮説検定

3.1 枠組みと例

3.1.1 枠組み

定義 3.1.1 (hypothesis testing, null hypothesis, alternative hypothesis, significance level, critical region). 確率空間 (\mathcal{X}, P) に対して, 事前に与えられた設定 $\alpha \in (0, 1)$, H_0, H_1 から定められる部分集合 $\mathcal{X}_1 \subset \mathcal{X}$ を検定という.

- (1) 仮説検定において, 仮説とは, 母数または確率分布に関する (メタ) 条件をいう.
- (2) 検定される仮説 H_0 を帰無仮説といい, それに対立する仮説 H_1 を対立仮説という.
- (3) 文脈としては, H_0 を棄却し H_1 の成立を示すことで社会的な意味を見出そうという文脈に当てはまるように H_0, H_1 を選ぶ.
- (4) $P(H_0)$ の値に意味を持たせるための指標を有意水準という. $\alpha = 0.05, 0.01$ などが用いられる.
- (5) $\mathcal{X}_1 \subset \mathcal{X}$ を α と H_0, H_1 の意義から事前に定めて棄却域といい, 観測値 x が $x \in \mathcal{X}_1$ のとき棄却し, $x \notin \mathcal{X}_1$ のとき採択する.

例 3.1.2. $\alpha = 0.05$ とする. 20 回中事象 E が 13 回起こったとする. 帰無仮説 H_0 を $P(E) = 1/2$ とし, 対立仮説 H_1 を $P(E) > 1/2$ とする. すると, 棄却域は $x_0 \in \mathbb{N}$ を用いて $\mathcal{X}_1 = \{x \in [20] \mid x \geq x_0\}$ と定めるのが良いと考えられる. いま,

$$\sum_{k=14}^{20} \binom{20}{k} \left(\frac{1}{2}\right)^{20} \approx 0.0577$$

より, $13 < x_0$ であって, 観測値 13 は \mathcal{X}_1 には入らないから, 棄却できない. 「確率 0.05 程度の事象が偶々起こった」と解釈する方が選択される.

定義 3.1.3 (type I error, type II error). H_0 が正しいのに H_0 を棄却してしまう誤りを第一種の誤りまたは偽陽性といい, H_1 が正しいのに H_0 を採択してしまう誤りを第二種の誤りまたは偽陰性という.

3.1.2 ランダム化検定

検定の構成法の例をあげる. 検定 = 棄却域の良さは, 第一種の誤りの確率を有意水準 α で制限した後に, その範囲内で第二種の誤りの確率を最小化することで得られる.

例 3.1.4. 例 3.1.2 について, $x_0 = 15$ とすると,

$$\sum_{k=15}^{20} \binom{20}{k} \left(\frac{1}{2}\right)^{20} \approx 0.0207$$

より, 棄却域のサイズは 5% を大きく切って 2% 程度となってしまう. そこで, $14 < x_0 < 15$ にあたる棄却域を構成したいが, $x_0 \in \mathbb{Z}$ の値は変えられない. そこで, 棄却域をランダム化する.

$$\binom{20}{k} \left(\frac{1}{2}\right)^{20} \varphi + \sum_{k=15}^{20} \binom{20}{k} \left(\frac{1}{2}\right)^{20} = \alpha = 0.05$$

を満たす $\varphi \in (0, 1)$ を用いて, 観測値 $x = 14$ に対しては確率 φ で H_0 を棄却し, 確率 $1 - \varphi$ で H_0 を採択することとする.

3.1.3 形式化

定義 3.1.5 (test / critical (function), size, level- α -test, power function, uniformly most powerful (UMP) test). 確率分布族 $\mathcal{P} = (P_\theta)_{\theta \in \Theta} \subset \mathcal{M}(\mathcal{X}, A)$ を考える.

- (1) 棄却域 $\mathcal{X}_1 \subset \mathcal{X}$ の定めるパラメータ空間の分割 $\Theta = \Theta_0 + \Theta_1$ を考え, 帰無仮説を $H_0 : \theta \in \Theta_0$ で表し, 対立仮説を $H_1 : \theta \in \Theta_1$ と表す.
- (2) $|\Theta_i| = 1$ の場合を単純仮説といい, そうでない場合を複合仮説という.^{†1}
- (3) 可測関数 $\varphi : \mathcal{X} \rightarrow [0, 1]$ を検定 (関数) という. $\exists A \in \mathcal{A} \varphi = 1_A$ であるとき検定 φ は非確率的である.
- (4) $\sup_{\theta \in \Theta_0} E_\theta[\varphi]$ を φ の大きさという. 大きさが α 以下の検定を水準 α 検定という. 水準 α 検定の全体を $\Phi_\alpha \subset \text{Meas}(\mathcal{X}, [0, 1])$ で表す.
- (5) $\beta_\varphi : \Theta_1 \rightarrow [0, 1]; \theta \mapsto E_\theta[\varphi]$ を検定出力関数という. $1 - \beta_\varphi(\theta)$ は第二種の誤りの確率を表す.
- (6) $\forall \varphi \in \Phi_\alpha \beta_{\varphi_0}(\theta) \geq \beta_\varphi(\theta)$ を満たす検定 $\varphi_0 \in \Phi_\alpha$ を一様最強力検定という.

要諦 3.1.6 (統計的決定理論の枠組みでの解釈). 決定空間は, 採択する仮説の添字からなる空間 $\mathcal{D} := 2 = \{0, 1\}$ である (0 が受容, 1 が棄却). 損失関数は

$$W(\theta, a) = \begin{cases} 1_{\{1\}}(a), & \theta \in \Theta_0, \\ 1_{\{0\}}(a), & \theta \in \Theta_1. \end{cases}$$

で, 検定関数 φ は確率的決定関数

$$\delta_\varphi(dz|x) = \varphi(x)\epsilon_1(dz) + (1 - \varphi(x))\epsilon_0(dz)$$

に対応する. ただし, ϵ_a はデルタ測度とした. したがって危険関数は

$$\begin{aligned} R(\theta, \delta_\varphi) &= \int_{\mathcal{X}} \int_{\mathcal{D}} W(\theta, z) \delta_\varphi(dz|x) \epsilon_0(dz) \\ &= \begin{cases} E_\theta[\varphi], & \theta \in \Theta_0, \\ 1 - E_\theta[\varphi], & \theta \in \Theta_1. \end{cases} \end{aligned}$$

となり, $\Phi_\alpha = \left\{ \varphi \in \text{Meas}(\mathcal{X}, [0, 1]) \mid \sup_{\theta \in \Theta_0} R(\theta, \delta_\varphi) \leq \alpha \right\}$ と表せる. 決定関数の全体は $\Delta \subset \Phi_\alpha$ となる.

Δ の中に入る一番自然な順序は一様順序で, これによる最大元が存在するときこれを UMP と呼ぶ.

定義 3.1.7 (nuisance parameter, goodness of fit test).

- (1) Θ が 2 次元以上で, 特定のパラメータにしか興味がないとき, 分割 $\Theta_1 + \Theta_2$ は商空間となる. この時の興味の無い母数を局外母数または攪乱母数という.
- (2) 帰無仮説を「確率変数が正規分布に従う」という命題にする場合, 特に分布の適合度検定と呼ぶ. 適合度検定や「確率分布が独立である」などの数学的な仮定が現実のデータと矛盾がないかチェックするための検定のクラスを統計的モデルの診断という.

^{†1} Θ が 1 次元で, $H_0 : \theta = \theta_0, H_1 : \theta \neq \theta_0$ の形のときを両側検定, $H_0 : \theta \leq \theta_0, H_1 : \theta > \theta_0$ の形のときを片側検定という.

第 4 章

大標本理論

非線形モデル・非ガウスモデルに対しては、統計量の分布が複雑なため、標準的な統計的決定理論を適用するのは困難であるが、これに替わって漸近的方法に基づく大標本理論が有力な解析手段となる。ここで、正規近似に基づく 1 次の漸近理論を扱う。

4.1 最尤推定

Ronald Fisher 1912-22 が開発。最尤推定量は一致性を満たすが、この現象はより広いクラスについて成り立つ。

問題 4.1.1. 可測空間 $(\mathcal{X}, \mathcal{A})$ 上の分布族 $(P_\theta)_{\theta \in \Theta}$ について、独立な観測値 x_1, \dots, x_n から θ を推定する問題を考える。

定義 4.1.2 (MLE: maximum likelihood estimator). $(\mathcal{X}, \mathcal{A})$ 上の σ -有限測度 ν に関して P_θ が絶対連続であるとし、微分を $p(x, \theta) := dP_\theta/d\nu(x)$ と表す。

(1) $L_n(\theta) := \prod_{j=1}^n p(x_j, \theta)$ を尤度関数という。これは、ある $\theta \in \Theta$ について、これが定める分布 P_θ から観測値 x_1, \dots, x_n が得られる条件付き確率に当たる。

(2) 各 $(x_1, \dots, x_n) \in \mathcal{X}^n$ に対して、尤度関数を最大にする $\tilde{\theta}_n(x_1, \dots, x_n) \in \Theta$ を対応させる関数 $\tilde{\theta}_n: \mathcal{X}^n \rightarrow \Theta$ を最尤推定量という。^{†1}

注 4.1.3 (M -推定量としての最尤推定量). 単調増加関数との合成を考えて、対数尤度関数 $l_n := \log \circ L_n$ の最大化を考えるにあたって、パラメータの空間 Θ が可微分構造を持つとき、ランダムな方程式 $\partial_\theta l_n(\theta) = 0$ の解としての最尤推定量を特徴付けることが出来る。これを尤度方程式という。このとき出現した対数尤度の微分をスコア関数 $S(x, \theta) := \frac{\partial}{\partial \theta} l(\theta, x)$ という。

要諦 4.1.4 (最尤推定量の一致性). 大数尤度関数 $l_n(\theta) := \log L_n(\theta) = \sum_{j=1}^n \log p(x_j, \theta)$ は、独立同分布に従う確率変数の和である。よって、大数の法則により、真値 $\theta_0 \in \Theta$ について、

$$\forall_{\theta \in \Theta} \quad \frac{1}{n} [l_n(\theta) - l_n(\theta_0)] \rightarrow \int \log \left(\frac{p(x, \theta)}{p(x, \theta_0)} \right) P_{\theta_0}(dx) \quad P_{\theta_0}\text{-a.s.}$$

が成り立つ。 $P: \Theta \rightarrow \{P_\theta\}$ が単射であるという識別可能性条件を仮定すると、相対エントロピーの非負性より、真値 θ_0 はこの右辺を最大にする唯一の点となる。そこで、左辺を最大にする推定量が $\tilde{\theta}_n$ なのであったから、 $n \rightarrow \infty$ のとき、 $\tilde{\theta}_n \rightarrow \theta_0$ が予想される。あとは、任意の $\theta \in \Theta$ に対する一様性を確認すれば良い。

補題 4.1.5. 可測空間 (Ω, \mathcal{F}) 上の σ -有限測度 ν に関して絶対連続な確率測度 P, Q について、その微分を $p(\omega) := dP/d\nu(\omega), q(\omega) := dQ/d\nu(\omega)$ とおく。このとき、

$$\int_{\Omega} \log \left(\frac{p(\omega)}{q(\omega)} \right) P(d\omega) \leq 0$$

で、等号成立条件は $P = Q$ のとき。

^{†1} ここでは推定量は可測という仮定は置かない。

定義 4.1.6 (consistency). パラメータの推定量 $\theta_n : \mathcal{X}^n \rightarrow \Theta$ について,

- (1) 任意の $\theta \in \Theta$ について, これが真値であったとき, $\widetilde{\theta}_n \rightarrow \theta$ a.s. ($n \rightarrow \infty$) が成り立つとき, この推定量は強一貫性を持つという.
- (2) 推定量 $\widetilde{\theta}_n$ が可測で, 任意の $\theta \in \Theta$ について, これが真値であったとき, $\widetilde{\theta}_n \xrightarrow{P} \theta$ ($n \rightarrow \infty$) が成り立つとき, この推定量は弱一貫性を持つという.

たったこれだけの議論で, 相対エントロピーなる確率測度間の距離概念と, 実数の乗法群と加法群の間の準同型 $\mathbb{R}_{>0} \rightarrow \mathbb{R}$ としての対数関数とのいずれもが出てくるとは思わなかった. 「尤度」なる意味論と「相対エントロピー」という意味論とが繋がるのが一番の驚き.

4.2 一様対数の法則

最尤推定量の性質を確認するためには, $\theta \in \Theta$ に依らずに一様に収束することを確認すれば良い.

記法 4.2.1.

- (1) $(\mathcal{X}, \mathcal{A}, P)$ の可算無限直積 $(\overline{\mathcal{X}}, \overline{\mathcal{A}}, \overline{P})$ 上に, $\pi_n := (\text{pr}_1, \dots, \text{pr}_n) : \overline{\mathcal{X}} \rightarrow \mathcal{X}^n$ で \mathcal{X}^n 上の可測関数を引き戻して議論する.
- (2) $U : \mathcal{X} \times \Theta \rightarrow (-\infty, \infty]$ に対して,

$$\overline{U}_n(\theta) = \overline{U}_n(x, \theta) = \frac{1}{n} \sum_{j=1}^n U(x_j, \theta)$$

と表す. ただし, $x = (x_1, \dots, x_n) \in \mathcal{X}^n$ とした.

補題 4.2.2. (Θ, d) を距離空間, $\Theta_1 \subset \Theta$ をコンパクト集合, $U : \mathcal{X} \times \Theta \rightarrow [-\infty, \infty]$ を関数とし, 以下の条件を仮定する:

- (A1) $\forall x \in \mathcal{X} \ \theta \mapsto U(x, \theta)$ は下半連続.
- (A2) $\forall \theta \in \Theta \ r$ が十分に小さければ, $x \mapsto \inf_{\theta' \in \Theta, d(\theta', \theta) < r} U(x, \theta')$ は可測.
- (A3) ある可積分関数 $M : \mathcal{X} \rightarrow \mathbb{R}$ が存在して, $\forall (x, \theta) \in \mathcal{X} \times \Theta \ U(x, \theta) \geq M(x)$.

このとき, $\overline{U}(\theta) := \int_{\mathcal{X}} U(x, \theta) P(dx) \in (-\infty, \infty]$ に対して,

$$(1) \quad \liminf_{n \rightarrow \infty} \inf_{\theta \in \Theta_1} \overline{U}_n(\theta) \geq \inf_{\theta \in \Theta_1} \overline{U}(\theta) \quad \text{a.s.}$$

(2) \overline{U} は下半連続.

定理 4.2.3 (一様大数の法則). (Θ, d) を距離空間, $\Theta_1 \subset \Theta$ をコンパクト集合, $U : \mathcal{X} \times \Theta \rightarrow [-\infty, \infty]$ を関数とし, 以下の条件を仮定する:

- (A1) $\forall x \in \mathcal{X} \ U(x, -) : \Theta \rightarrow \mathbb{R}$ は連続.
- (A2) $\forall \theta \in \Theta \ U(-, \theta) : \mathcal{X} \rightarrow \mathbb{R}$ は可測.
- (A3) ある可積分関数 $M \in \mathcal{L}^1(\mathcal{X})$ が存在して, $\forall (x, \theta) \in \mathcal{X} \times \Theta \ |U(x, \theta)| \leq M(x)$.

このとき, $\overline{U}(\theta) := \int_{\mathcal{X}} U(x, \theta) P(dx) \in \mathbb{R}$ は連続で,

$$\lim_{n \rightarrow \infty} \sup_{\theta \in \Theta_1} |\overline{U}_n(\theta) - \overline{U}(\theta)| = 0 \quad \text{a.s.}$$

4.3 最小コントラスト推定

最尤推定の枠組みを一般化し、このクラスについてより一般的な形で一致性を示す。

定義 4.3.1 (contrast function, minimum contrast estimator). (Θ, d) を距離空間, $(\mathcal{X}, \mathcal{A}, P)$ を確率空間とする。

- (1) コントラスト関数とは、関数 $\Psi: \mathcal{X} \times \Theta \rightarrow (-\infty, \infty]$ をいう。
- (2) これを最小にする写像 $\tilde{\theta}: \mathcal{X} \rightarrow \Theta$ を最小コントラスト推定量という。

例 4.3.2. 尤度関数は $L_n: \mathcal{X}^n \times \Theta \rightarrow [0, 1]$ と定め、これを最大化したが、一致性の議論の中で、真の分布との対数尤度の差 $|l_n(\theta) - l_n(\tilde{\theta})|$ の最小化を考えた。

4.4 M-推定量

Huber が 1964 年に最尤推定を一般化し、"maximum likelihood-type" の頭文字から M-推定量のクラスを定めた。^a

^a <https://en.wikipedia.org/wiki/M-estimator>

定義 4.4.1. $\Theta \subset \mathbb{R}^p$ とする。関数 $\psi: \mathcal{X} \times \Theta \rightarrow \mathbb{R}^p$ に対して、

$$\bar{\psi}_n(\theta) = \bar{\psi}_n(x, \theta) = \frac{1}{n} \sum_{j=1}^n \psi(x_j, \theta)$$

とおき、 $\bar{\psi}_n(\hat{\theta}_n) = 0$ を満足する未知パラメータの推定量 $\hat{\theta}_n$ を M-推定量という。 ψ を推定関数という。

例 4.4.2. 最小コントラスト推定量は、 $\partial_\theta \Psi_n(\theta) = 0$ なる推定関数に対応する M-推定量である。

4.5 推定量の漸近正規性

一致性は大数の法則に当たるならば、その真値への収束の速度が同様に気になる。これが漸近正規性である。これを、M-推定量について議論する。

4.6 ワンステップ推定量

M-推定量は推定関数の零点として定まるが、実際に方程式を解くことは容易でない場合も多い。一方で、有効でないが、一致性をもつ推定量は容易に見つかることがしばしばある。このとき、それを用いて、最尤推定量と同じ漸近分散を持つ推定量を構成できる。

4.7 M-推定量の存在

Cramer による。

4.8 頑健推定

4.9 尤度比検定

4.10 多項分布の検定

4.11 尤度比確率場の局所漸近構造

4.12 情報量規準

4.13 密度推定

4.14 U -統計量

第 5 章

漸近展開とその応用

中心極限定理は分布の正規近似であるが，Edgeworth による展開はその近似をより精密にしたものである．標本数がそれほど大きくない実際の状況では，分布のより精密な近似を基礎として統計量を構成することが必要になる．

今日，漸近展開法も確率過程にまでその領域を広げ，新しい確率統計学が発展しつつあるが，このような新領域を理解する上でも，独立確率変数列における現象を理解することが重要である．

参考文献

- [1] 吉田朋広『数理統計学』(朝倉書店, 2006)
- [2] 竹村彰道『現代数理統計学』(学術図書, 2020)
- [3] 久保川達也『現代数理統計学の基礎』(共立出版, 2017)
- [4] 西山陽一『マルチンゲール理論による統計解析』(近代科学社, 2011)
- [5] Rabi Bhattacharya - Course in Mathematical Statistics and Large Sample Theory