$S_n = \sum_{i=1}^{n} X_n$, $m_n = \mathrm{E}(S_n)$ とする。$\sum_{i=1}^{\infty} \frac{\sigma_i^2}{i^2} < \infty$ のとき，

$$P\left(\lim_{n\to\infty} \frac{S_n - m_n}{n} = 0\right) = 1$$

が成り立つ。

8　統計的推定

8.1　推定法と標本平均および標本分散の性質

8.1.1　推定法

点推定：　母集団の分布型は既知，その分布のある特性値 θ (母数) は未知とする。
その母集団の分布は $f(x;\theta)$ で与えられている。

このとき，標本の実現値 $(x_1,\ x_2,\ \cdots,\ x_n)$ から適当な値 $\widehat{\theta}_n(x_1,\ x_2,\ \cdots,\ x_n)$ を計算する。

$\widehat{\theta}_n(x_1,\ x_2,\ \cdots,\ x_n)$ を θ の推定値とする。$\Longrightarrow$ 点推定

例：

母平均 μ の点推定値 (＝標本平均 $\overline{x}$)

$$\widehat{\mu}_n(x_1, x_2, \cdots, x_n) \equiv \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

母分散 σ^2 の点推定値 (＝標本不偏分散 s^2)

$$\widehat{\sigma}_n^2(x_1, x_2, \cdots, x_n) \equiv s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

区間推定：　母集団の分布の未知母数 θ を推定するとき，実現値 $(x_1,\ x_2,\ \cdots,\ x_n)$ より $\widehat{\theta}_L(x_1,\ x_2,\ \cdots,\ x_n)$ と $\widehat{\theta}_U(x_1,\ x_2,\ \cdots,\ x_n)$ を作り，区間 $(\widehat{\theta}_L, \widehat{\theta}_U)$

の中に θ は $1-\alpha$ の確率で入っていることを示す推定法を区間推定法という。
区間 $(\widehat{\theta}_L,\widehat{\theta}_U)$ は信頼係数 $1-\alpha$ の θ の信頼区間である。
$\widehat{\theta}_L \Longrightarrow$ 信頼下限
$\widehat{\theta}_U \Longrightarrow$ 信頼上限
信頼区間の幅はなるべく狭くなるように $\widehat{\theta}_L$, $\widehat{\theta}_U$ を選ぶ。

8.1.2　標本，統計量，推定量

母集団の分布型は既知，その分布のある特性値 θ (母数) は未知とする。
その母集団の分布は $f(x;\theta)$ で与えられている。
標本： $\boldsymbol{X_1, X_2, \cdots, X_n} \Longrightarrow$ 母集団の部分集合
実現値： $\boldsymbol{x_1, x_2, \cdots, x_n}$
母数 θ の推定量： $\widehat{\theta}(\boldsymbol{X_1},\ \boldsymbol{X_2},\ \cdots,\ \boldsymbol{X_n})$
母数 θ の推定値： $\widehat{\theta}(\boldsymbol{x_1},\ \boldsymbol{x_2},\ \cdots,\ \boldsymbol{x_n})$

例：　$\boldsymbol{\theta = (\mu, \sigma^2)}$ とする。

$\boldsymbol{\mu}$ の推定量： $\boldsymbol{\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i}$

$\boldsymbol{\mu}$ の推定値： $\boldsymbol{\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i}$

$\boldsymbol{\sigma^2}$ の推定量： $\boldsymbol{S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$

$\boldsymbol{\sigma^2}$ の推定値： $\boldsymbol{s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}$

注)

確率変数の関数 $\Longrightarrow$ 統計量

母数の推定のために使われる統計量 $\Longrightarrow$ 推定量

8.1.3 母平均，母分散の推定

X_1, X_2, $\cdots$, X_n は互いに独立で，すべて同一の分布 (すなわち，平均 μ, 分散 σ^2 ですべて同一の分布) に従うものとする。

1. 母平均 μ の推定量：
$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

2. 母分散 σ^2 の推定量：
   - 母平均 μ が既知のとき： $S^{*2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$
   - 母平均 μ が未知のとき： $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$

$\overline{\boldsymbol{X}}$ の性質：

$$\begin{aligned}
\mathbf{E}(\overline{\boldsymbol{X}}) &= \mathbf{E}\left(\frac{1}{\boldsymbol{n}}\sum_{i=1}^{n}\boldsymbol{X}_i\right) \\
&= \frac{1}{\boldsymbol{n}}\mathbf{E}\left(\sum_{i=1}^{n}\boldsymbol{X}_i\right) \\
&= \frac{1}{\boldsymbol{n}}\sum_{i=1}^{n}\mathbf{E}(\boldsymbol{X}_i) \\
&= \frac{1}{\boldsymbol{n}}\sum_{i=1}^{n}\boldsymbol{\mu} \\
&= \frac{1}{\boldsymbol{n}}\boldsymbol{n}\boldsymbol{\mu} \\
&= \boldsymbol{\mu}
\end{aligned}$$

$$
\begin{aligned}
\mathrm{V}(\overline{\boldsymbol{X}}) &= \mathrm{V}\Big(\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i\Big) \\
&= \frac{1}{n^2}\mathrm{V}\Big(\sum_{i=1}^{n}\boldsymbol{X}_i\Big) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{V}(\boldsymbol{X}_i) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\boldsymbol{\sigma}^2 \\
&= \frac{1}{n^2}n\boldsymbol{\sigma}^2 \\
&= \frac{\boldsymbol{\sigma}^2}{n}
\end{aligned}
$$

必要な公式：

$\mathrm{E}(aX) = a\mathrm{E}(X)$

$\mathrm{V}(aX) = a^2\mathrm{V}(X)$

$\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y)$

$\mathrm{V}(X+Y) = \mathrm{V}(X) + \mathrm{V}(Y) \Longleftarrow X$ と Y は独立のとき

S^{*2}, S^2 の性質：

$$
\begin{aligned}
\mathrm{E}(S^{*2}) &= \mathrm{E}(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2) \\
&= \frac{1}{n}\mathrm{E}(\sum_{i=1}^{n}(X_i-\mu)^2) \\
&= \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}((X_i-\mu)^2)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{n} \sum_{i=1}^{n} \mathrm{V}(X_i) \\
&= \frac{1}{n} \sum_{i=1}^{n} \sigma^2 \\
&= \frac{1}{n} n\sigma^2 \\
&= \sigma^2
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}(S^2) &= \mathrm{E}\left(\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2\right) \\
&= \frac{1}{n-1} \mathrm{E}\left(\sum_{i=1}^{n} (X_i - \overline{X})^2\right) \\
&= \frac{1}{n-1} \mathrm{E}\left(\sum_{i=1}^{n} ((X_i - \mu) - (\overline{X} - \mu))^2\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{n-1} \mathbf{E}\Big(\sum_{i=1}^{n} \big((X_i - \mu)^2 \\
&\qquad -2(X_i - \mu)(\overline{X} - \mu) \\
&\qquad +(\overline{X} - \mu)^2\big)\Big) \\
&= \frac{1}{n-1} \mathbf{E}\Big(\sum_{i=1}^{n} (X_i - \mu)^2 \\
&\qquad -2(\overline{X} - \mu) \sum_{i=1}^{n} (X_i - \mu) \\
&\qquad +n(\overline{X} - \mu)^2\Big) \\
&= \frac{1}{n-1} \mathbf{E}\Big(\sum_{i=1}^{n} (X_i - \mu)^2 \\
&\qquad -2n(\overline{X} - \mu)^2 + n(\overline{X} - \mu)^2\Big)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{n-1}\mathrm{E}\Big(\sum_{i=1}^{n}(X_i-\mu)^2 - n(\overline{X}-\mu)^2\Big) \\
&= \frac{1}{n-1}\mathrm{E}\Big(\sum_{i=1}^{n}(X_i-\mu)^2\Big) \\
&\qquad - \frac{1}{n-1}\mathrm{E}(n(\overline{X}-\mu)^2) \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\mathrm{E}((X_i-\mu)^2) \\
&\qquad - \frac{n}{n-1}\mathrm{E}((\overline{X}-\mu)^2) \\
&= \frac{1}{n-1}\sum_{i=1}^{n}\mathrm{V}(X_i) - \frac{n}{n-1}\mathrm{V}(\overline{X})
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{n-1}\sum_{i=1}^{n}\boldsymbol{\sigma}^2 - \frac{n}{n-1}\frac{\boldsymbol{\sigma}^2}{n} \\
&= \frac{1}{n-1}n\boldsymbol{\sigma}^2 - \frac{1}{n-1}\boldsymbol{\sigma}^2 \\
&= \boldsymbol{\sigma}^2
\end{aligned}
$$

必要な公式：

$$\mathbf{E}(\boldsymbol{X}_i - \boldsymbol{\mu})^2 = \mathbf{V}(\boldsymbol{X}_i) = \boldsymbol{\sigma}^2$$

$$\mathbf{E}(\overline{\boldsymbol{X}} - \boldsymbol{\mu})^2 = \mathbf{V}(\overline{\boldsymbol{X}}) = \frac{\boldsymbol{\sigma}^2}{\boldsymbol{n}}$$

$$\sum_{i=1}^{n}(\boldsymbol{X}_i - \boldsymbol{\mu}) = \boldsymbol{n}(\overline{\boldsymbol{X}} - \boldsymbol{\mu})$$

したがって，$\boldsymbol{S}^{**2} = \frac{1}{\boldsymbol{n}}\sum_{i=1}^{n}(\boldsymbol{X}_i - \overline{\boldsymbol{X}})^2$ とすると，

$$\begin{aligned}
\mathbf{E}(\boldsymbol{S}^{**2}) &= \mathbf{E}\Big(\frac{1}{\boldsymbol{n}}\sum_{i=1}^{n}(\boldsymbol{X}_i - \overline{\boldsymbol{X}})^2\Big) \\
&= \mathbf{E}\Big(\frac{\boldsymbol{n}-1}{\boldsymbol{n}}\frac{1}{\boldsymbol{n}-1}\sum_{i=1}^{n}(\boldsymbol{X}_i - \overline{\boldsymbol{X}})^2\Big) \\
&= \mathbf{E}\Big(\frac{\boldsymbol{n}-1}{\boldsymbol{n}}\boldsymbol{S}^2\Big) \\
&= \frac{\boldsymbol{n}-1}{\boldsymbol{n}}\mathbf{E}(\boldsymbol{S}^2)
\end{aligned}$$

$$= \frac{n-1}{n}\sigma^2$$

8.2　点推定法： 最適性

母数： θ

推定値： $\widehat{\theta}_n(x_1, x_2, \cdots, x_n)$

推定量： $\widehat{\theta}_n(X_1, X_2, \cdots, X_n)$

$\widehat{\theta}_n(X_1, X_2, \cdots, X_n)$ を $\widehat{\theta}_n$ と書く。

$\widehat{\theta}_n$ の望ましい性質：不偏性，有効性，十分性，一致性

$\Longrightarrow$ 最適性

$\Longrightarrow$ 最適推定量

不偏性：　$\mathrm{E}(\widehat{\theta}_n) = \theta$

$\Longrightarrow \theta$ を中心にして $\widehat{\theta}_n$ は分布している。
$\widehat{\theta}_n$ は θ の不偏推定量であるという。
$\mathrm{E}(\widehat{\theta}_n) - \theta$ をバイアス (bias) と呼ぶ。

$X_1, X_2, \cdots, X_n$ は互いに独立で，すべて同一の分布 (すなわち，平均 μ, 分散 σ^2 ですべて同一の分布) に従うものとする。

1. 母平均 μ の推定量：
$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$
2. 母分散 σ^2 の推定量：
$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

$\mathrm{E}(\overline{X}) = \mu$, $\mathrm{E}(S^2) = \sigma^2$ なので，$\overline{X}$, S^2 は μ, σ^2 の不偏推定量である。

有効性：　**2 つの不偏推定量 $\widehat{\theta}_n$，$\widetilde{\theta}_n$ を考える。**
すなわち，$\mathrm{E}(\widehat{\theta}_n) = \theta$, $\mathrm{E}(\widetilde{\theta}_n) = \theta$
$\mathrm{V}(\widehat{\theta}_n) < \mathrm{V}(\widetilde{\theta}_n)$ のとき，$\widehat{\theta}_n$ が $\widetilde{\theta}_n$ より有効であるという。
$\Longrightarrow$ バラツキの小さい推定量の方が望まれる。

クラメール・ラオの不等式 (Cramer-Rao Inequality): 任意の不偏推定量 $\widehat{\theta}_n$ について，

$$\mathrm{V}(\widehat{\theta}_n) \geq \frac{\sigma^2}{n}$$

が成り立つ。ただし，

$$\begin{aligned}\sigma^2 &= \sigma^2(\theta) \\ &= \frac{1}{\mathrm{E}\left[\left(\dfrac{\partial \log f(X;\theta)}{\partial \theta}\right)^2\right]}\end{aligned}$$