2020年度3Q
応用確率統計

平均と分散の不偏推定

阪口 啓

sakaguchi@mobile.ee.titech.ac.jp

講義スケジュール後半

	日付	教科書	内容	演習
第8回	10月30日	9.1, 9.2	統計的推論（推測）	なし
第9回	11月6日	9.3, 10.1	平均と分散の不偏推定	あり
第10回	11月10日	11, 13	最尤推定とベイス推論	あり
第11回	11月13日	12	仮説検定、分類、機械学習	あり
第12回	11月17日	配布資料	回帰、予測、機械学習	あり
第13回	11月20日	配布資料	確率過程、相関、予測	あり
第14回	11月24日		理解度確認総合演習	

復習

- 母数の推定 (Parameter estimation)
  - 観測データから未知母数（統計パラメタ）を推定する方法
  - 母数の推定: $X_i = \theta + \varepsilon_i \ \ (i = 1, \cdots, n) \ \ \rightarrow \ \ \hat{\theta}$

- 仮説検定、分類 (Hypothesis test, Classification)
  - 観測データと想定するデータが一致するかどうかを確認する方法
  - 仮説検定:「 $\theta = \theta_0$ or $\theta \neq \theta_0$ 」、分類:「 $\theta = \theta_0$ or $\theta = \theta_1$ 」

- 回帰、予測 (Regression, Prediction)
  - 観測データから2つ以上の物理量の関係を表現する方法
  - 線形回帰: $Y_i = ax_i + b + \varepsilon_i \ \ (i = 1, \cdots, n) \ \ \rightarrow \ \ \hat{a}, \hat{b}$

- 確率過程、予測 (Stochastic process, Forecast)
  - 未知母数が時間の関数である観測データから未来を予測する方法
  - 予測: $Y_i(t) = \theta(t) + \varepsilon_i \ \ (i = 1, \cdots, n) \ \ \rightarrow \ \ \hat{\theta}(t+1)$

講義内容

- 平均の推定量
- 標本平均の分散
- 分散の推定量

標本平均の例

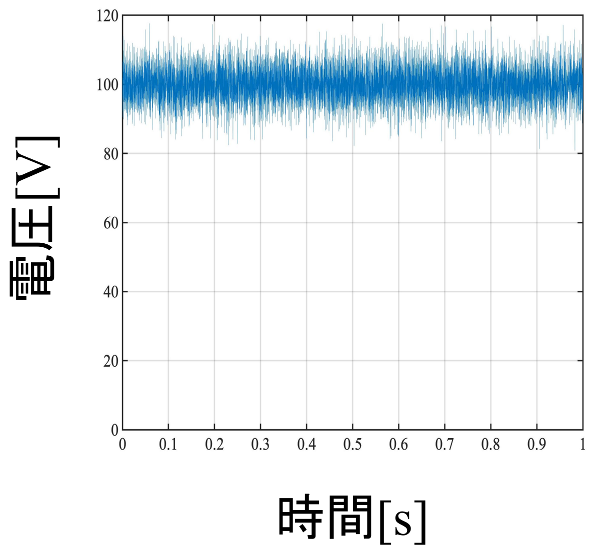
■ テストの点数

A君	B君	C君	D君	E君
75点	81点	96点	61点	88点

平均点

$$\frac{75+81+96+61+88}{5}=80.2$$

■ 電圧の測定値



電圧の平均値

$$\bar{V}=\frac{1}{n}\sum_{i=1}^{n}V_i$$

平均値の不偏推定

■ 観測データ

例: $X_i = \theta + \varepsilon_i \quad (i = 1, \cdots, n)$

確率変数　未知母数　誤差

誤差の分布: $f(\varepsilon_i)$　平均: $E(\varepsilon_i) = 0$　分散: $V(\varepsilon_i) = \sigma^2$

■ 最適な推定量

観測値に基づく推定値の平均が真の母数に一致すること

不偏: $E(\hat{\theta} \mid \theta) = E^{X_1, \cdots, X_n}\left(\hat{\theta}(X_1, \cdots, X_n) \mid \theta\right) = \theta$

（※ E^X : X に対する期待値）

分散最小: $\hat{\theta}^* = \arg\min_{\hat{\theta}} V(\hat{\theta} \mid \theta)$

平均の推定量の例

■ 標本平均 (Sample average)

$$\hat{\theta} = \frac{X_1 + \cdots + X_n}{n}$$

■ 中央値 (Median)

大きさの順に並べ替え

$$\{\tilde{X}_1, \cdots, \tilde{X}_n\} = \mathrm{sort}\{X_1, \cdots, X_n\}$$

$$\hat{\theta} = \begin{cases} \tilde{X}_{(n+1)/2} & n : \mathrm{odd} \\ (\tilde{X}_{n/2} + \tilde{X}_{n/2+1})/2 & n : \mathrm{even} \end{cases}$$

■ 加重平均 (Weighted average)

$$\sum_{i=1}^{n} c_n = 1$$ ← データの重み

$$\hat{\theta} = c_1 X_1 + c_2 X_2 + \cdots + c_n X_n$$

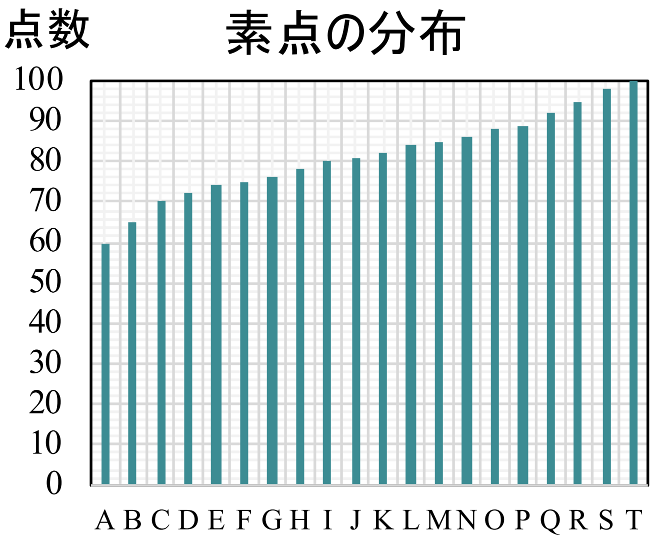
■ 幾何平均 (Geometric average)

$$\hat{\theta} = \left(\prod_{i=1}^{n} X_i \right)^{1/n}$$

X_i は正の数

標本平均の例

	点数
A君	60
B君	65
C君	70
D君	72
E君	74
F君	75
G君	76
H君	78
I君	80
J君	81
K君	82
L君	84
M君	85
N君	86
O君	88
P君	89
Q君	92
R君	95
S君	98
T君	100



標本平均

$$\bar{X} = \frac{X_1 + \cdots + X_n}{n} = 81.5$$

標本平均の不偏性

■ 標本平均

誤差の分布未知のときの最小分散推定量

$$\hat{\theta}^* = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

■ 不偏性

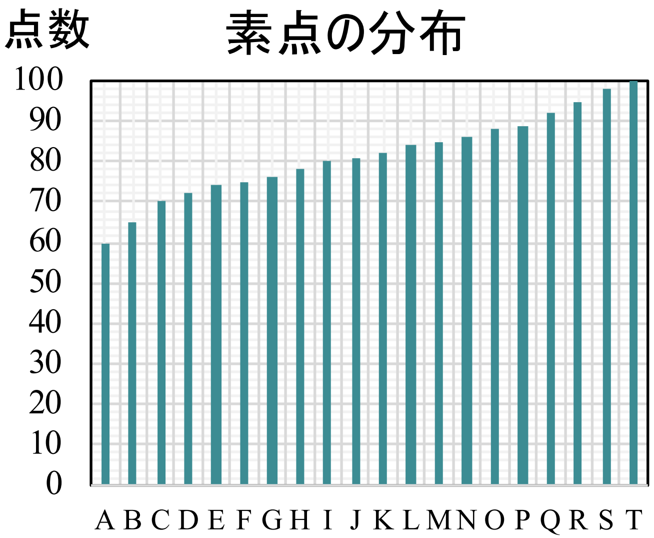
$$E\left(\hat{\theta}\right) = \frac{E\left(X_1\right) + \cdots + E\left(X_n\right)}{n} = \theta$$

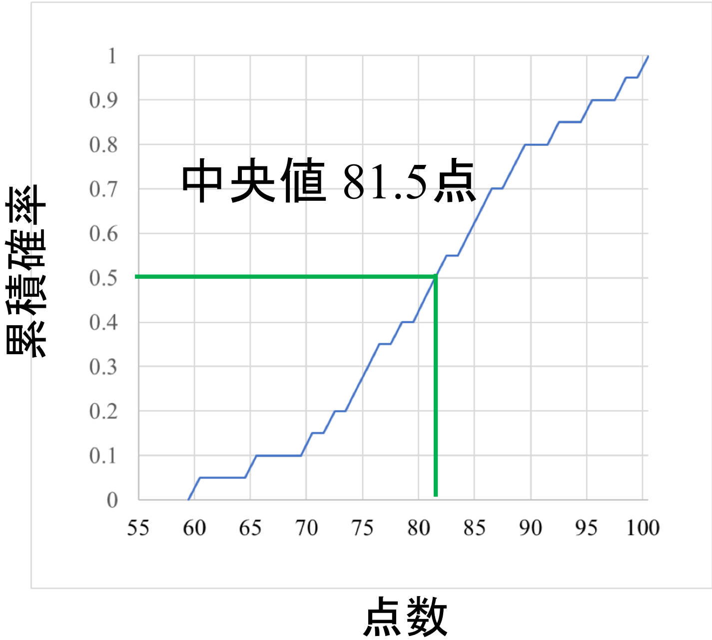
$$\Downarrow$$

標本平均（相加平均、算術平均）は不偏

中央値の例

	点数
A君	60
B君	65
C君	70
D君	72
E君	74
F君	75
G君	76
H君	78
I君	80
J君	81
K君	82
L君	84
M君	85
N君	86
O君	88
P君	89
Q君	92
R君	95
S君	98
T君	100





標本平均

$$\bar{X} = \frac{X_1 + \cdots + X_n}{n} = 81.5$$

中央値

$$\hat{\theta} = \frac{\left(\tilde{X}_{\frac{n}{2}} + \tilde{X}_{\frac{n}{2}+1}\right)}{2} = 81.5$$

中央値の不偏性

■ 推定量

大きさの順に並べ替え

$$\left\{\widetilde{X}_1,\cdots,\widetilde{X}_n\right\}=\mathrm{sort}\left\{X_1,\cdots,X_n\right\}$$

$$\hat{\theta}=\begin{cases}\tilde{X}_{(n+1)/2} & n:\mathrm{odd}\\ \left(\tilde{X}_{n/2}+\tilde{X}_{n/2+1}\right)/2 & n:\mathrm{even}\end{cases}$$

■ 不偏性

確率分布が対称なとき

$$\int_{-\infty}^{M}p(x)dx=\int_{M}^{\infty}p(x)dx=1/2$$

$E\left(\hat{\theta}\right)=M=E(X)=\theta$ ⇨ 確率分布が対称なとき中央値は不偏

幾何平均の例

■ 売上の伸び率

	2016	2017	2018	2019
売上[円]	1000万	2000万	3000万	9000万
昨対		200%	150%	300%

算術平均

$$\frac{2+1.5+3}{3}=2.17$$

1000万×2.17×3年＝6510万

幾何平均

$$\sqrt[3]{2\times1.5\times3}=2.08$$

1000万×2.08^3＝9000万

幾何平均の不偏性

■ 推定量

$$\hat{\theta} = \left(\prod_{i=1}^{n} X_i \right)^{1/n} = \exp\left(\frac{1}{n} \log \left(\prod_{i=1}^{n} X_i \right) \right) = \exp\left(\frac{1}{n} \sum_{i=1}^{n} \log X_i \right)$$

イエンセンの不等式

上に凸な関数 $f(x)$ に対して

$$\sum_{i=1}^{n} p(x_i) f(x_i) \leq f\left(\sum_{i=1}^{n} p(x_i) x_i \right)$$

■ 標本平均との関係

$$\log \left(\prod_{i=1}^{n} X_i \right)^{1/n} = \frac{1}{n} \sum_{i=1}^{n} \log X_i \leq \log \left(\frac{1}{n} \sum_{i=1}^{n} X_i \right)$$

$$\therefore \quad \left(\prod_{i=1}^{n} X_i \right)^{1/n} \leq \frac{1}{n} \sum_{i=1}^{n} X_i$$ ⇨ 幾何平均（相乗平均）は不偏ではない

講義内容

- 平均の推定量
- 標本平均の分散
- 分散の推定量

標本平均の分散

■ 標本平均

$$\hat{\theta}^* = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

■ 標本平均の分散

$$V\left(\hat{\theta}^*\right) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = E\left(\left(\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1}{n}\sum_{i=1}^{n}\theta\right)^2\right)$$

$$= \frac{1}{n^2} E\left(\left(\sum_{i=1}^{n}\left(X_i - \theta\right)\right)^2\right) = \frac{1}{n^2} E\left(\sum_{i=1}^{n}\left(X_i - \theta\right)^2\right)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} V\left(X_i\right)^2 = \frac{\sigma^2}{n}$$

→ クラーメル・ラオの下限に一致
有効（分散最小）推定量

一致性

■ 一致推定量

観測数を増やすと推定量が未知母数に一致（確率収束）

$$\lim_{n\to\infty} P\left(\left|\hat{\theta}^* - \theta\right| \geq \varepsilon\right) = 0$$

■ 標本平均

$$\hat{\theta}^* = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$E\left(\hat{\theta}^*\right) = \theta \qquad \lim_{n\to\infty} V\left(\hat{\theta}^*\right) = \lim_{n\to\infty}\frac{\sigma^2}{n} = 0$$

⟶ チェビシェフの不等式（大数の法則）

講義内容

- 平均の推定量
- 標本平均の分散
- 分散の推定量

標本分散

■ 標本分散

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \qquad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

■ 不偏性

$$n\left(\overline{X} - \theta\right)$$

$$E\left(\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2\right) = E\left(\sum_{i=1}^{n}\left(\left(X_i - \theta\right) - \left(\overline{X} - \theta\right)\right)^2\right)$$

$$= E\left(\sum_{i=1}^{n}\left(X_i - \theta\right)^2 - 2\boxed{\sum_{i=1}^{n}\left(X_i - \theta\right)}\left(\overline{X} - \theta\right) + \sum_{i=1}^{n}\left(\overline{X} - \theta\right)^2\right)$$

$$= E\left(\sum_{i=1}^{n}\left(X_i - \theta\right)^2 - n\left(\overline{X} - \theta\right)^2\right) = nV\left(X\right) - nV\left(\overline{X}\right)$$

$$= n\sigma^2 - \sigma^2 = (n-1)\sigma^2 \longrightarrow$$ 不偏ではない

不偏分散

■ 標本分散

$$E\left(\hat{\sigma}^2\right) = \frac{1}{n} E\left(\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2\right) = \frac{n-1}{n}\sigma^2$$

■ 不偏分散

$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2$$

$$E\left(\hat{\sigma}^2\right) = \frac{n-1}{n-1}\sigma^2 = \sigma^2$$

← 観測値一つ分に相当する情報を平均に使ったため

不偏分散の例

■ テストの点数

A君	B君	C君	D君	E君
75点	81点	96点	61点	88点

■ 標本平均

$$\overline{X} = \frac{75+81+96+61+88}{5} = 80.2$$

■ 不偏分散

$$\sigma^2 = \frac{1}{4}\left\{(75-80.2)^2+(81-80.2)^2+\cdots+(88-80.2)^2\right\} = 176.7$$

まとめ

■ 標本平均

$$\hat{\theta}^* = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

■ 不偏性と一致性

$$E\left(\hat{\theta}^*\right) = \theta \qquad \lim_{n\to\infty} V\left(\hat{\theta}^*\right) = \lim_{n\to\infty}\frac{\sigma^2}{n} = 0$$

■ 不偏分散

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \qquad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$