確率統計
Probability and Statistics
第6回講義資料
Lecture notes 6
統計学の基礎と点推定
Introduction of Statistics and Point Estimation

豊橋技術科学大学
Toyohashi University of Technology
電気・電子情報工学系
Department of Electrical and Electronic Information Engineering
准教授　竹内啓悟
Associate Professor　Keigo Takeuchi

統計学の目的(Purpose of statistics)

ある集団の性質を特徴付けるような指標を低コストで推計したい。

Estimate indicators that characterize properties of a group with low costs.

国勢調査(Census)

日本に常住する全世帯を対象とする。(Survey of all families that are living in Japan.)

利点(Merits) : 正確な推計ができる。(Accurate estimation)

欠点(Demerits) : 調査に費用がかかる。(High cost)

所得調査(Survey for income)

日本に住む全世帯の所得の傾向を特徴付けるには、どんな指標に注目すべきか？

What indicators should we focus on to characterize the trends of income of all families in Japan.

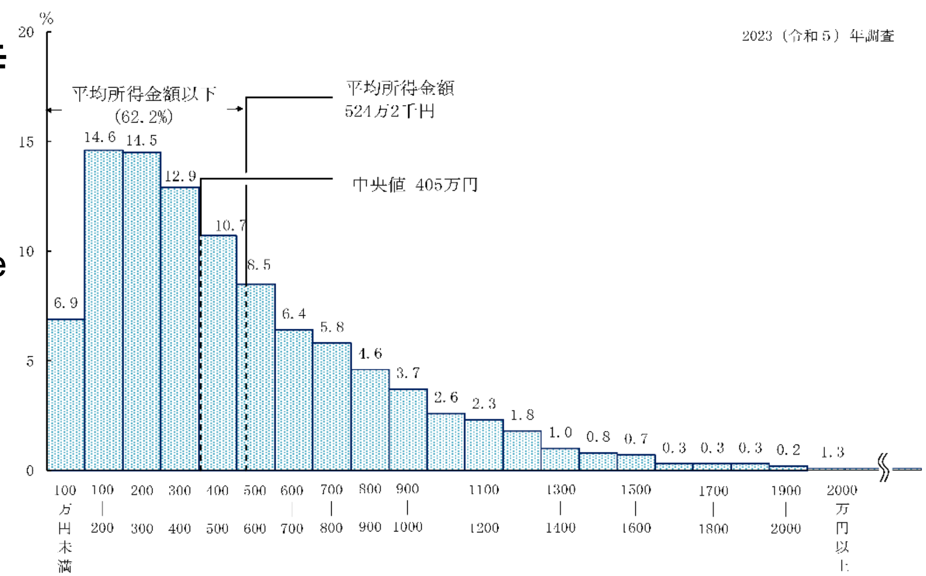
平均値、中央値、最頻値?

Mean, median, or mode?

私見 : 分布を示せば良い。

My opinion: Show the distribution.

図９　所得金額階級別世帯数の相対度数分布



厚生労働省令和5年国民生活基礎調査より抜粋

統計的推測の定式化(Formulation of statistical inference)

母集団(Population)

調査したい集団の属性値の集合

The set of attribution of a group to be investigated.

例(Example)：世帯所得(Income)

母集団分布(Population distribution)

母集団から計算される相対度数分布

Relative frequency distribution computed from a population.

以下の表記を使う。

離散分布 Discrete distribution $P(X = x)$

連続分布 Continuous distribution $p_X(x)$

仮説(Hypotheses)

母集団の要素数は十分に多い。(Population contains a huge number of members.)

真の母集団分布は、有限個の母数によって決まる分布で近似できる。

The true population distribution can be approximated by one with a finite number of parameters.

統計学の目的(Purpose of statistics)

ある集団の性質を特徴付けるような指標を低コストで推計したい。

Estimate indicators that characterize properties of a group with low costs.

⇒ 母集団分布の母数を低コストで推定したい。

Estimate the parameters in the population distribution with low costs.

標本抽出による推計(Estimation via sampling)

標本(Sample)

母集団から抽出された属性値の集合(The set of members sampled from a population.)

標本の大きさ(Sample size)

母集団から抽出された属性値の数(The number of members sampled from a population.)

標本の大きさを標本数と言ってはいけない。

Do not confuse the sample size with the number of samples.

標本数とは、母集団が複数の集団から構成される場合の集団の数のことである。(例：所得の母集団を60歳以下とその他との二集団に分けた場合、標本数は2。)

The number of samples means the number of groups when a population is composed of multiple groups. (Example: The number of groups is 2 when the income population is divided into two groups: under or over 60.)

数学的には(Mathematically)

大きさNの標本は、周辺分布が母集団分布に従うN個の確率変数列である。

A sample of size N is a sequence of random variables of which the marginal distributions are equal to the population distribution.

実現値ではないことに注意。(They are not realizations.)

標本抽出による推計(Estimation via sampling)

無作為抽出(Random sampling)

全部でM個の要素から成る母集団の中から、等確率でN個の標本を抽出する。

Sample N different members out of the population that are composed of M members with uniform probability.

Mが十分に大きいとき、無作為抽出は、母集団から独立かつ一様な確率でN個の要素を取り出す復元抽出と等価である。

When M is sufficiently large, random sampling is equivalent to sampling with replacement that picks up N members from the population independently and uniformly.

数学的には(Mathematically)

大きさNの無作為抽出された標本とは、母集団分布に従うN個の独立な確率変数列である。

A sample of size N via random sampling is a sequence of independent random variables that follows the population distribution.

統計的推測の目的(Purpose of statistical inference)

無作為抽出された標本(値)から母集団分布の母数を推定したい。

Estimate the parameters in the population distribution from (a realization of) a sample obtained via random sampling.

統計量(Statistic)

ある母数の推定量に対応する標本Xの関数$T(X)$

A function $T(X)$ of a sample X that corresponds to an estimator of a parameter.

注意(Remarks)

統計量は未知の母数に依存してはいけない。

Any statistic must not depend on unknown parameters.

統計量は確率変数である。その実現値である統計値と混同してはいけない。

A statistic is a random variable. Do not confuse it with a realization.

例(Examples)：

標本平均：母集団の平均(母平均)に対する統計量

Sample mean: A statistic of the population mean

標本分散：母集団の分散(母分散)に対する統計量

Sample variance: A statistic of the population variance

点推定の目的(Purpose of point estimation)

母集団分布の標準的推定法(Standard method for estimation of population distributions)

ある特定の母集団分布族の中から、データに基づいて最良な母集団分布を選択する。

(Select the best distribution based on data out of a certain family of population distributions.)

母集団分布族(Family of population distributions) : $P(X;\theta),\ \theta \in \Theta.$

標本Xの分布は、母数θによって分布が完全に特定されると仮定する。

The distribution of the sample X is assumed to be determined by the parameter θ.

母集団分布(Population distribution)	母数(Parameters)
二項分布(Binomial distribution)	1が出る確率 (Probability of an outcome of 1)
正規分布(Normal distribution)	平均と分散 (Mean and variance)

母数の点推定(Point estimation of the parameter)

標本Xの実現値xが得られたときに、母数θを推定せよ。

Estimate the parameter θ when a realization of the sample has been obtained.

点推定の例(Example of point estimation)

正規母集団から無作為抽出された標本$\boldsymbol{X} = \{X_1, \dots, X_N\}$

i.i.d. sample $\boldsymbol{X}$ from the normal population

$$p_{X_n}(x;\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \qquad \boldsymbol{\theta} = \{\mu, \sigma^2\}$$

標本Xから、平均μと分散σ^2をどのように推定するのが最良だろうか？

What is the best estimation of the mean and variance based on the sample?

標本平均：(Sample mean)
$$\bar{X} = \frac{1}{N}\sum_{n=1}^{N} X_n$$

標本分散：(Sample variance)
$$\hat{\sigma}^2 = \frac{1}{N}\sum_{n=1}^{N} (X_n - \bar{X})^2$$

直観的には尤もらしい推定法であるが、これよりも良い方法はないのか？

While they are intuitively reasonable methods, are not there better methods?

そもそも、推定法の良し悪しをどのような基準で比較すべきか？

What criteria should we use in comparing merits or demerits of estimation?

推定量が満たすべき性質(Properties desired in estimators)

母数θの推定量を$\hat{\theta} = \hat{\theta}(X)$と書く。

Let $\hat{\theta}$ denote an estimator of the parameter θ.

不偏性(Unbiasedness)

推定量$\hat{\theta}$が$\mathbb{E}\left[\hat{\theta}\right] = \theta$を満たすとき、$\hat{\theta}$を不偏推定量と呼ぶ。

An estimator $\hat{\theta}$ is called unbiased if $\mathbb{E}\left[\hat{\theta}\right] = \theta$ is satisfied.

一致性(Consistency)

標本の大きさが無限大のとき、推定量$\hat{\theta}$が真の母数θに確率収束するならば、$\hat{\theta}$は一致推定量と呼ばれる。

An estimator $\hat{\theta}$ is called consistent if $\hat{\theta}$ converges in probability to the true parameter θ as the sample size tends to infinity.

有効性(Efficiency)

不偏推定量$\hat{\theta}$が平均二乗誤差$\mathrm{E}[(\theta - \hat{\theta})^2]$の理論限界を達成するとき、$\hat{\theta}$は有効推定量と呼ばれる。

An unbiased estimator $\hat{\theta}$ is called efficient if $\hat{\theta}$ achieves a theoretical lower bound on the mean-square error (MSE) $\mathrm{E}[(\theta - \hat{\theta})^2]$.

不偏推定量の例(Examples of unbiased estimators)

平均μと分散σ^2を母数$\boldsymbol{\theta} = \{\mu, \sigma^2\}$とする母集団分布$P(X; \boldsymbol{\theta})$から無作為抽出された標本$\boldsymbol{X} = \{X_1, \dots, X_N\}$を使って、母数を推定しよう。

Estimate the parameters $\boldsymbol{\theta} = \{\mu, \sigma^2\}$ from an independent and identically distributed (i.i.d.) sample $\boldsymbol{X}$ drawn from the population distribution $P(X; \boldsymbol{\theta})$ with mean μ and variance σ^2.

標本平均(Sample mean): $$\bar{X} = \frac{1}{N}\sum_{n=1}^{N} X_n$$

標本平均は不偏である。(The sample mean is unbiased.)

$$\because \quad \mathbb{E}[\bar{X}] = \frac{1}{N}\sum_{n=1}^{N} \mathbb{E}[X_n] = \frac{1}{N}\sum_{n=1}^{N} \mu = \mu.$$

不偏分散(Unbiased sample variance): $$\bar{S}^2 = \frac{1}{N-1}\sum_{n=1}^{N} (X_n - \bar{X})^2$$

不偏分散は不偏推定量である。(The unbiased sample variance is unbiased.)

証明: $Y_n = X_n - \mu$とおくと、$\mathbb{E}[Y_n] = 0$と$\mathbb{V}[Y_n] = \sigma^2$が成り立つので、一般性を失うことなく、$\mu = 0$を仮定できる。

Proof

Let $Y_n = X_n - \mu$. Since $\mathbb{E}[Y_n] = 0$ and $\mathbb{V}[Y_n] = \sigma^2$ hold, without loss of generality, we can assume $\mu = 0$.

不偏分散の不偏性の証明(Proof of the unbiasedness of the unbiased sample variance)

標本平均の定義を使って、(Using the definition of the sample mean yields)

$$\sum_{n=1}^{N}(X_n - \bar{X})^2 = \sum_{n=1}^{N}(X_n^2 - 2\bar{X}X_n + \bar{X}^2) = \sum_{n=1}^{N} X_n^2 - N\bar{X}^2.$$

標本$\boldsymbol{X}$に関する期待値を取ると、

Taking the expectation with respect to the sample $\boldsymbol{X}$, we have

$$(N-1)\mathbb{E}[\bar{S}^2] = N\sigma^2 - N\mathbb{E}[\bar{X}^2],$$

ここで、不偏分散の定義と仮定$\mu = 0$を使った。

where we have used the definition of the unbiased sample variance and the assumption of $\mu = 0$.

標本平均は$\mu = 0$の不偏推定量なので、

Since the sample mean is an unbiased estimator of $\mu = 0$, we have

$$\mathbb{E}[\bar{X}^2] = \mathbb{E}[(\bar{X} - \mathbb{E}[\bar{X}])^2] = \mathbb{V}[\bar{X}] = \frac{\sigma^2}{N}.$$

最後の等号は、標本平均の定義と標本の独立同一性とから従う。

The last equality follows from the definition of the sample mean and the i.i.d. property of the sample.

これらの式を組み合わせると、$\mathbb{E}[\bar{S}^2] = \sigma^2$を得る。

Combining these equations, we arrive at $\mathbb{E}[\bar{S}^2] = \sigma^2$. ■

最尤推定(Maximum likelihood (ML) estimation)

母集団分布に統計値xを代入した母数θの関数$P(X=x;\theta)$を尤度と呼ぶ。

We refer to the function $P(X=x;\theta)$ of θ obtained by substituting a realization x of the sample into the population distribution as likelihood.

最尤推定量(ML estimator)：
$$\hat{\theta}=\underset{\theta\in\Theta}{\operatorname{argmax}}\,P(X=x;\theta)$$

統計値xの出現確率である尤度を最大にするような母数を選ぶ。

Select a parameter that maximizes the likelihood—the probability of the occurrence of x.

日本語と英語の違い(Difference between Japanese and English)

尤もらしさ≒道理に合っている(reasonable)様子≒適切さ(suitability)

Likelihood≒possibility(可能性)≒probability(確率)

最尤推定は最大出現確率に基づく推定という意味なので、

なぜ一番尤もらしい推定と言えるのかという議論はしない方が良い。

Since the ML estimation means estimation based on the maximum probability of outcome, we should not discuss why the ML estimation is the most reasonable estimation.

最尤推定の例(Example of ML estimation)

正規母集団から無作為抽出された標本$\boldsymbol{X}=\{X_1,\dots,X_N\}$

i.i.d. sample $\boldsymbol{X}$ from the normal population

$$p_{X_n}(x;\boldsymbol{\theta})=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\qquad \boldsymbol{\theta}=\{\mu,\sigma^2\}$$

平均μと分散σ^2の最尤推定量$\hat{\mu}_{\mathrm{ML}}$と$\hat{\sigma}^2_{\mathrm{ML}}$を計算しよう。

Compute the ML estimators of the mean and variance.

尤度を最大化することは、対数尤度を最大化することと等価である。

The maximization of the likelihood is equivalent to that of the log likelihood.

$$\{\hat{\mu}_{\mathrm{ML}},\hat{\sigma}^2_{\mathrm{ML}}\}=\underset{\mu,\sigma^2}{\mathrm{argmax}}\log\prod_{n=1}^{N}p_{X_n}(X_n;\boldsymbol{\theta})=\underset{\mu,\sigma^2}{\mathrm{argmax}}\sum_{n=1}^{N}\log p_{X_n}(X_n;\boldsymbol{\theta}).$$

対数尤度を標本平均$\bar{X}$を使って表現すると、

Representing the log likelihood with the sample mean yields

$$\sum_{n=1}^{N}\log p_{X_n}(X_n;\boldsymbol{\theta})=-\frac{N}{2}\left\{\frac{\mu^2-2\bar{X}\mu+\overline{X^2}}{\sigma^2}+\log(2\pi\sigma^2)\right\},\ \ \overline{X^2}=\frac{1}{N}\sum_{n=1}^{N}X_n^2.$$

最尤推定の例(Example of ML estimation)

平均の最尤推定量$\hat{\mu}_{\mathrm{ML}}$は、σ^2によらず標本平均と等しい。

The ML estimator $\hat{\mu}_{\mathrm{ML}}$ of the mean is independent of σ^2, and equal to the sample mean.

$$\hat{\mu}_{\mathrm{ML}} = \bar{X}$$

確かめよ。(Confirm it.)

分散の最尤推定量$\hat{\sigma}^2_{\mathrm{ML}}$は、標本分散と等しい。

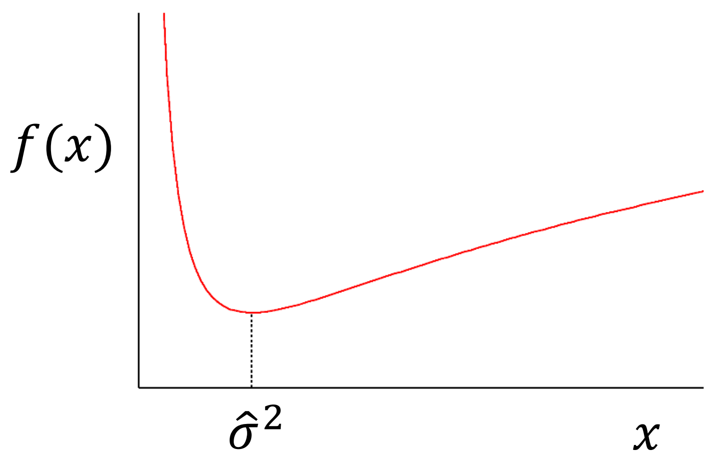
The ML estimator $\hat{\sigma}^2_{\mathrm{ML}}$ of the variance is equal to the sample variance.

$$\hat{\sigma}^2_{\mathrm{ML}} = \operatorname*{argmin}_{\sigma^2>0} \left\{ \frac{\mu^2 - 2\bar{X}\mu + \overline{X^2}}{\sigma^2} + \log(2\pi\sigma^2) \right\} \Bigg|_{\mu=\bar{X}} = \operatorname*{argmin}_{\sigma^2>0} \left\{ \frac{\hat{\sigma}^2}{\sigma^2} + \log(\sigma^2) \right\},$$

ここで、$\hat{\sigma}^2$は標本分散を表す。(where $\hat{\sigma}^2$ denotes the sample variance.)

関数$f(x) = \log x + \hat{\sigma}^2/x$は、
$x = \hat{\sigma}^2$のときに最小値を取る。

The function $f(x)$ is minimized at $x = \hat{\sigma}^2$.



$$\therefore \quad \hat{\sigma}^2_{\mathrm{ML}} = \hat{\sigma}^2.$$

最尤推定の性質(Properties of ML estimation)

適切な正則条件を満たす母集団確率密度関数$p_X(x;\theta)$から無作為抽出された標本$\boldsymbol{X}=\{X_1,\dots,X_N\}$に基づく母数$\theta$の最尤推定量に関して、

Consider the ML estimator of the parameter based on the i.i.d. sample drawn from a population pdf $p_X(x;\theta)$ satisfying regularity conditions.

不偏性(Unbiasedness)

最尤推定量は必ずしも不偏ではない。(ML estimators are not necessarily unbiased.)

一致性(Consistency)

最尤推定量は一致推定量である。(ML estimators are consistent.)

それ故、標本の大きさが無限大のときに、最尤推定量は不偏推定量である。

Thus, the ML estimators are unbiased as the sample size tends to infinity.

有効性(Efficiency)

標本の大きさが無限大のときに、最尤推定量は有効推定量である。

ML estimators are efficient as the sample size tends to infinity.

標本の大きさが十分に大きいとき、最尤推定は最良な推定法であると言える。

We can conclude that ML estimation is the best method of estimation for a sufficiently large sample size.

課題(Exercise)

$\boldsymbol{X} = \{X_1, \ldots, X_N\}$を正規母集団から無作為抽出された標本とする。母平均μと母分散σ^2に対する最尤推定量$\hat{\mu}_{\mathrm{ML}}$と$\hat{\sigma}^2_{\mathrm{ML}}$を導出せよ。

Let $\boldsymbol{X} = \{X_1, \ldots, X_N\}$ denote an i.i.d. sample from the normal population. Derive the ML estimators $\hat{\mu}_{\mathrm{ML}}$ and $\hat{\sigma}^2_{\mathrm{ML}}$for the population mean μ and variance σ^2.