

平成22年度
統計工学 講義資料(6)

～統計的推論:推定Ⅱ～

平成22年度 統計工学
(C) Sadami SUZUKI

母分散の推定

・ 標本分散 $\hat{\sigma}_X^2$ による母分散 σ^2 の推定は...

観測値: X_1, X_2, \dots, X_n → 確率変数: X_1, X_2, \dots, X_n
による母分散 σ^2 の推定

$\hat{\sigma}_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} S_X$
 $S_X = \sum_{i=1}^n (X_i - \bar{X})^2 \dots$ 標本平均からの偏差平方和
 標本分散は母分散を**過小評価**してしまうきがある

平成22年度 統計工学
(C) Sadami SUZUKI

母分散に対する過小評価

母平均 μ が既知であれば...

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$$

実際に未知のため...

平成22年度 統計工学
(C) Sadami SUZUKI

母分散に対する過小評価

一般的には真値を用いたほうが母数パラメータに近い推定値をあたえるはず

$$\sigma^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$$

母平均 μ が既知とした場合の推定量

→

<標本分散>

$$\hat{\sigma}_X^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

母平均を標本平均で推定した場合の推定量

標本分散では母分散を**過小評価**してしまう

平成22年度 統計工学
(C) Sadami SUZUKI

標本分散による過小評価の修正

適当に修正されることの証明

$$\sum_{i=1}^n (X_i - \mu)^2 = \sum_{i=1}^n (X_i - \bar{X} + \bar{X} - \mu)^2$$

$$= \sum_{i=1}^n (X_i - \bar{X})^2 + \sum_{i=1}^n (\bar{X} - \mu)^2 + \sum_{i=1}^n 2(X_i - \bar{X})(\bar{X} - \mu)$$

ここで両辺の期待値をとると

$$E\left[\sum_{i=1}^n (X_i - \mu)^2\right] = E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] + nE[(\bar{X} - \mu)^2]$$

和の期待値=期待値の和、また右辺第2項は標本平均の分散であることに注意して

$$n\sigma^2 = E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] + \sigma^2$$

$$\therefore E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = n\sigma^2 - \sigma^2 = (n-1)\sigma^2$$

$$E[(n-1)V_x] = (n-1)E[V_x] = (n-1)\sigma^2$$

推定量の不偏性

推定量が平均的に母集団パラメータに一致する
平均的には正しく推定する

平成22年度 統計工学
(C) Sadami SUZUKI

母平均および母分散の推定量

推定量

母平均推定量
【標本平均】
$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

母分散推定量
【不偏分散】
$$\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{S_X}{n-1}$$

n-1 ... 偏差の値に関する**自由度**

n個の偏差の値は全てが自由に任意の値をとれるわけではないことに注意!

平成22年度 統計工学
(C) Sadami SUZUKI

平成22年度 統計工学

母平均の区間推定と t 分布

標準化

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

スチューデント化

$$T = \frac{\bar{X} - \mu}{\hat{\sigma} / \sqrt{n}}$$

$\hat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}$

標準正規分布に従う 自由度 $n-1$ の t 分布に従う

標準正規分布 (自由度 ∞ の t 分布)
t 分布 (自由度 20)
t 分布 (自由度 5)
t 分布 (自由度 2)

t 分布 (t-distribution) の分布形状は標本の大きさ n のみに依存し、未知の母集団パラメータには一切依存しない。
ここで $n-1$ を t 分布の自由度という。
t 分布の形状は自由度さえ与えれば一意に決まる。

(C) Sadami SUZUKI

平成22年度 統計工学

母平均の区間推定と t 分布

スチューデント化

$$T = \frac{\bar{X} - \mu}{\hat{\sigma} / \sqrt{n}}$$

$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \sim N(0,1)$

$\Pr\left\{-1.96 < \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} < 1.96\right\} = 0.95$
 $\Pr\left\{-1.96 \frac{\hat{\sigma}}{\sqrt{n}} < \bar{X} - \mu < 1.96 \frac{\hat{\sigma}}{\sqrt{n}}\right\} = 0.95$

自由度 $n-1$ の t 分布に従う

t 分布表の値を活用して...

$\Pr\left\{-t_{n-1}(0.05) < \frac{\bar{X} - \mu}{\hat{\sigma} / \sqrt{n}} < t_{n-1}(0.05)\right\} = 0.95$
 $\Pr\left\{\bar{X} - t_{n-1}(0.05) \frac{\hat{\sigma}}{\sqrt{n}} < \mu < \bar{X} + t_{n-1}(0.05) \frac{\hat{\sigma}}{\sqrt{n}}\right\} = 0.95$

平均値 μ の 95% 信頼区間

$$\left(\bar{X} - t_{n-1}(0.05) \frac{\hat{\sigma}}{\sqrt{n}}, \bar{X} + t_{n-1}(0.05) \frac{\hat{\sigma}}{\sqrt{n}}\right)$$

(C) Sadami SUZUKI

平成22年度 統計工学

母分散の推定

母平均の推定と同様に...

観測値: X_1, X_2, \dots, X_n 確率変数: X_1, X_2, \dots, X_n
による母分散 σ^2 の推定

抽出 抽出

母集団 標本 (サンプル)

推定 推定

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

不偏分散なる推定値は必ずしも母分散にピッタリ一致するわけではない

不偏分散から、母分散が含まれる範囲(区間)を推定しよう

(C) Sadami SUZUKI

平成22年度 統計工学

母分散の区間推定

母分散に関する不偏推定量

母分散との比率を考えると平均的には1になる

$$E\left[\frac{\hat{\sigma}^2}{\sigma^2}\right] = E[C^2] = 1$$

$$C^2 = \frac{\hat{\sigma}^2}{\sigma^2}$$

自由度 $n-1$ の修正 χ^2 分布に従う

$$\chi^2 = (n-1)C^2 = \frac{(n-1)\hat{\sigma}^2}{\sigma^2}$$

自由度 $n-1$ の χ^2 分布に従う

χ^2 分布

(C) Sadami SUZUKI

平成22年度 統計工学

母分散の区間推定

$$C^2 = \frac{\hat{\sigma}^2}{\sigma^2}$$

$$\chi^2 = (n-1)C^2 = \frac{(n-1)\hat{\sigma}^2}{\sigma^2}$$

自由度 $n-1$ の修正 χ^2 分布に従う 自由度 $n-1$ の χ^2 分布に従う

χ^2 分布表の値を活用して...

$\Pr\left\{\chi_{n-1}^2(0.975) < \frac{(n-1)\hat{\sigma}^2}{\sigma^2} < \chi_{n-1}^2(0.025)\right\} = 0.95$
 $\Pr\left\{\frac{(n-1)\hat{\sigma}^2}{\sigma^2} < \chi_{n-1}^2(0.025) < \frac{(n-1)\hat{\sigma}^2}{\sigma^2}\right\} = 0.95$

分散 σ^2 の 95% 信頼区間

$$\left(\frac{(n-1)\hat{\sigma}^2}{\chi_{n-1}^2(0.025)}, \frac{(n-1)\hat{\sigma}^2}{\chi_{n-1}^2(0.975)}\right)$$

(C) Sadami SUZUKI

平成22年度 統計工学

χ^2 分布に関する整理

抽出 抽出

母集団 標本 (サンプル)

観測値: X_1, X_2, \dots, X_n 確率変数: X_1, X_2, \dots, X_n

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

自由度 n の χ^2 分布に従う

標準正規分布 $N(0,1)$ から n 個の独立な標本値の平方和は自由度 n の χ^2 分布に従うと言い換えることも可能。

$$\chi^2 = \sum_{i=1}^n z_i^2$$

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

自由度 $n-1$ の χ^2 分布に従う

(C) Sadami SUZUKI