応用統計学 2017 第12回 情報量規準

2018年1月10日 (水)
清 智也 sei@mist.i.u-tokyo.ac.jp
http://ur0.pw/yTzt

- モデル選択，期待対数尤度，赤池情報量規準 (AIC)，変数増加法，変数減少法[1]。
- AIC に基づくモデル選択：

$$\mathrm{AIC} = -2 \times (\text{最大対数尤度}) + 2 \times (\text{パラメータ数})$$

  をモデルごとに計算し，AIC が最小となるモデルを採用する。
- 参考文献など
  - 小西・北川「情報量規準」(朝倉書店)
  - 坂元・石黒・北川「情報量統計学」(共立出版)
  - 赤池弘次からのメッセージ（稲盛財団チャンネル）
    `https://www.youtube.com/watch?v=QAugnlK74Tw`

演習問題

問題 12-1. 次の表は2015年度と2016年度における東京の日最高気温の月平均値を表す。

月	1	2	3	4	5	6	7	8	9	10	11	12
2015	10.4	10.4	15.5	19.3	26.4	26.4	30.1	30.5	26.4	22.7	17.8	13.4
2016	10.6	12.2	14.9	20.3	25.2	26.3	29.7	31.6	27.7	22.6	15.5	13.8

2015年度のデータ $\{y_t\}_{t=1}^{12}$ に正規線形モデル

$$y_t = a_0 + \sum_{j=1}^{k} \{a_j \cos(2\pi jt/12) + b_j \sin(2\pi jt/12)\} + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{N}(0, \sigma^2),$$

をあてはめ，それを用いて2016年度のデータ $\{\tilde{y}_t\}_{t=1}^{12}$ を予測するとき，二乗予測誤差がもっとも小さくなる k はどれか。ただし $0 \le k \le 5$ とする。また AIC が最小となる k を求めよ。

問題 12-2. 講義ページにあるデータ `nations.csv` は，世界各国の GDP，1人あたり GDP，人口密度，平均寿命である[2]。これら4つの変数から「その国がアジアの国かどうか」を判別するモデルを作りたい。ロジスティック回帰モデルをあてはめてみよ。さらに変数減少法によって選ばれるモデルを求めよ。

---

[1] model selection, expected log-likelihood, Akaike's information criterion, forward selection method, backward selection method.

[2] 総務省統計局のサイトにある「世界の統計」(http://www.stat.go.jp/data/sekai/0116.htm) の表 3-2, 3-3, 2-5, 2-17 をもとに（2017年1月に）作成した。

問題 12-3. 確率変数ベクトル $\boldsymbol{Y}$ と $\tilde{\boldsymbol{Y}}$ は独立に $\mathrm{N}(\boldsymbol{\mu}, \boldsymbol{I}_n)$ に従うとする。ただし $\boldsymbol{\mu} \in \mathbb{R}^n$ は未知パラメータであり，$\boldsymbol{I}_n$ は n 次単位行列である。また M を $\mathbb{R}^n$ の p 次元線形部分空間とし，$\boldsymbol{P}$ は $\mathbb{R}^n$ から M への直交射影行列とする。

(i) 任意の $\boldsymbol{\mu} \in \mathbb{R}^n$ に対して次の等式が成り立つことを示せ：

$$\mathrm{E}\left[\|\tilde{\boldsymbol{Y}} - \boldsymbol{P}\boldsymbol{Y}\|^2\right] = \|\boldsymbol{\mu} - \boldsymbol{P}\boldsymbol{\mu}\|^2 + n + p$$

この量は，「将来のデータ」$\tilde{\boldsymbol{Y}}$ を $\boldsymbol{P}\boldsymbol{Y}$ で予測したときのリスクを表している。

(ii) 任意の $\boldsymbol{\mu} \in \mathbb{R}^n$ に対して次の等式が成り立つことを示せ：

$$\mathrm{E}\left[\|\boldsymbol{Y} - \boldsymbol{P}\boldsymbol{Y}\|^2\right] = \|\boldsymbol{\mu} - \boldsymbol{P}\boldsymbol{\mu}\|^2 + n - p$$

(iii) モデル M の AIC は，定数項（すなわち M によらない項）を除けば

$$\mathrm{AIC} = \|\boldsymbol{Y} - \boldsymbol{P}\boldsymbol{Y}\|^2 + 2p$$

となることを示せ。また AIC がリスクの不偏推定量であることを示せ。

以上の結果は $\boldsymbol{\mu}$ が M に属していなくても成立することに注意しよう。この点が仮説検定の考え方とは根本的に異なっている。

お願い

講義時に配布するアンケートに回答してください。理学部用と工学部用があります。