応用システム工学
第4回　指数型分布族と一般化線形モデル

平成22年6月25日

最尤推定1

- 確率変数ベクトル　$y = [Y_1, Y_2, \cdots, Y_n]^t$
  - 確率変数n個
- パラメータベクトル　$\theta = [\theta_1, \theta_2, \cdots, \theta_p]^t$
  - パラメータ数p個
- Y_iの同時確率密度関数　$f(y;\theta)$
  - パラメータθを固定した時の確率変数yが対象
- 尤度関数　$L(\theta; y)$
  - 同時確率密度関数の数式表現と同一
    - θとyの役割の入替
    - 確率変数yを固定した時のパラメータθが対象
  - Lは確率変数となる
    - 確率変数ベクトルyの関数

最尤推定2

- パラメータ空間Ω　　$\theta \in \Omega$
  - パラメータベクトルθがとる全ての値を含む集合
- パラメータベクトルθの最尤推定量　$\hat{\theta}$
  - 尤度関数Lを最大にするθ

$$L\left(\hat{\theta}; y\right) \geq L\left(\theta; y\right), \theta \in \Omega$$

- 対数尤度関数l
  - 尤度関数Lの対数をとったもの　$l\left(\theta; y\right) = \log L\left(\theta; y\right)$
  - 対数関数は単調関数である
    - 対数尤度関数も単調関数となる
      - 最尤推定量は対数尤度関数を最大にする

$$l\left(\hat{\theta}; y\right) \geq l\left(\theta; y\right), \theta \in \Omega$$

最尤推定3

- 最尤推定量$\hat{\theta}$の求め方
  - 対数尤度関数の各パラメータに関する微分
    - 連立方程式

$$\frac{\partial l(\theta; y)}{\partial \theta_j} = 0,\ j = 1,2,\cdots,p$$

極値の条件

    - 極大値との対応
      - 二階導関数が負定値行列

$$\left.\frac{\partial^2 l(\theta; y)}{\partial \theta_j \partial \theta_k}\right|_{\theta=\hat{\theta}} < 0$$

» 全ての局所最大(極大)値の中で，lが最大値をとるθが最尤推定量

最尤推定4

- 最尤推定量の不変性
  - パラメータθに対する任意の関数g(θ)
    - パラメータベクトルθの最尤推定量　$\hat{\theta}$
    - $g(\theta)$　の最尤推定量も　$g(\hat{\theta})$　となる

- 最尤推定量の求め方
  - 最尤推定を行うのに便利な関数を見つける
  - この関数に対して尤度関数を最大化するθを見つける
  - 最尤推定量の不変性を用いる
  - 得られたθは必要なパラメータの最尤推定量となる

残差とモデルの検証

- 反応変数Yi,正規分布　$E(Y_i)=\mu_i; Y_i \sim N(\mu_i, \sigma^2)$
  - 期待値μi,分散σ２
- 残差　$y_i - \hat{\mu}_i$
  - 推定値　$\hat{\mu}_i$
- 標準化残差　$r_i = \frac{y_i - \hat{\mu}_i}{\hat{\sigma}}$
  - 未知パラメータσの推定値　$\hat{\sigma}$
- 残差平方和　$\sum (y_i - \hat{\mu}_i)^2$
  - モデルの適切さを計る統計量を与える

残差とモデルの検証

- 反応変数は互いに独立
  - 残差は近似的に平均0，分散一定の正規分布に従うべき
- 標準化残差の正規分布との比較検討
  - 度数分布の正規分布との適合性の検討
  - 標準化残差は分布による異なる

推測と解釈

- 測定値(観測値)の構成
  - 信号と雑音
- 節約の原則に基づくモデル化
  - 効果を説明する以上の原因を仮定すべきでない
    - 説明されない変化を残さない複雑なモデルより，データを適度に説明する簡単なモデルが好ましい
  - 節約的なモデルに用いるパラメータの仮説検定

統計モデル

- 正規分布に対するデータの分析
  - 独立な確率変数Y_i
    - $Y_i \sim N(\mu_i, \sigma^2)$
  - 期待値の線形モデル
    - $E(Y_i)=\mu_i=X_i^T\beta$

説明変数(デザイン行列):X
パラメータ:β

- 線形モデルの一般化
  - 正規分布以外の確率変数
    - 指数分布族(より広いクラスの分布)→本授業の対象
  - 確率変数と説明変数の関係が非線形
    - 連結関数の適用　$g(\mu_i)=X_i^T\beta$
    - 一般化加法モデル
      - 大量の計算処理必要→本授業では扱わない

指数型分布族の確率変数

- 指数型分布族の条件
  - 指数関数で表される確率分布
    - 確率変数Y,単一パラメータθ→最尤推定に便利

$$f(y;\theta) = s(y)t(\theta)\exp\left[a(y)b(\theta)\right]$$

    - 但し，a,b,s,tは既知の関数
    - $s(y) = \exp\left[d(y)\right], t(\theta) = \exp\left[c(\theta)\right]$ とおく

$$f(y;\theta) = \exp\left[d(y)\right]\exp\left[c(\theta)\right]\exp\left[a(y)b(\theta)\right]$$
$$= \exp\left[a(y)b(\theta) + c(\theta) + d(y)\right]$$

指数型分布族の確率変数

$$f(y;\theta)=\exp\left[a(y)b(\theta)+c(\theta)+d(y)\right]$$

- a(y)=yのとき「正準形」の分布と言い，
  このときb(θ)を分布の自然パラメータと呼ぶ。
- 局外パラメータ
  - θ 以外のパラメータ
  - 既知として扱う必要あり
    - 単一パラメータを仮定しているため
- 指数型分布族の例
  - 正規分布，ポアソン分布，ワイブル分布，二項分布

指数型分布族の対数尤度関数
最尤推定の準備

- 指数型分布族の確率密度関数

$$f(y;\theta)=\exp\left[a(y)b(\theta)+c(\theta)+d(y)\right]$$

- 指数型分布族の尤度関数

$$L(\theta;y)=\exp\left[a(y)b(\theta)+c(\theta)+d(y)\right]$$

- 指数型分布族の対数尤度関数

$$l(\theta;y)=\log L(\theta,y)=a(y)b(\theta)+c(\theta)+d(y)$$

指数型分布族としての正規分布

- 正規分布　　$Y \sim N\left(\mu, \sigma^2\right)$
  - 対称な分布を持つ連続データに対するモデル
  - 確率密度関数

$$f\left(y;\mu\right) = \frac{1}{\left(2\pi\sigma^2\right)^{\frac{1}{2}}} \exp\left[-\frac{1}{2\sigma^2}\left(y-\mu\right)^2 \right]$$

$$= \exp\left[\log \frac{1}{\left(2\pi\sigma^2\right)^{\frac{1}{2}}} - \frac{1}{2\sigma^2}\left(y-\mu\right)^2 \right]$$

$$= \exp\left[-\frac{1}{2}\log\left(2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\left(y^2 - 2y\mu + \mu^2\right) \right]$$

次頁につづく

指数型分布族としての正規分布

- 確率密度関数

$$f(y;\mu)=\exp\left[\frac{y\mu}{\sigma^2}-\frac{\mu^2}{2\sigma^2}-\frac{y^2}{2\sigma^2}-\frac{1}{2}\log\left(2\pi\sigma^2\right)\right]$$

$$=\exp\left[a(y)b(\mu)+c(\mu)+d(y)\right] \Rightarrow a(y)b(\mu)=\frac{y\mu}{\sigma^2}$$

  - パラメータ μ
  - 正準形 $a(y)=y$
  - 自然パラメータ $b(\mu)=\frac{\mu}{\sigma^2}$
    - 局外パラメータ σ^2
  - その他の項 $c(\mu)=-\frac{\mu^2}{2\sigma^2}-\frac{1}{2}\log\left(2\pi\sigma^2\right), d(y)=-\frac{y^2}{2\sigma^2}$

正準形(Canonical form)
線形システムは、座標変換によって元の系と同じ挙動を持つ系に変換できる。

指数型分布族としてのポアソン分布

- ポアソン分布　$Y \sim Poisson(\theta)$
  - 計数データのモデルとして用いる
  - 確率変数:Y→離散値，y=1,2,3,・・・等
  - 確率密度関数

$$f(y;\theta) = \frac{\theta^y e^{-\theta}}{y!}$$
$$= \exp(y\log\theta - \theta - \log y!)$$
$$= \exp[a(y)b(\theta) + c(\theta) + d(y)]$$

  - 正準形:a(y)=yである
  - 自然パラメータb(θ)=logθ
  - c(θ)=-θ , d(y)=-logy!

指数型分布族としてのワイブル分布

- ワイブル分布
  - 故障時間(生存期間)の分布を表すモデル
    - 確率変数:y(時間)
    - パラメータ: λ (分布の形状), θ (尺度)
  - 確率密度関数

$$f(y;\theta,\lambda)=\frac{\lambda y^{\lambda-1}}{\theta^{\lambda}}\exp\left[-\left(\frac{y}{\theta}\right)^{\lambda}\right]$$

$$=\exp\left[\log\lambda+(\lambda-1)\log y-\lambda\log\theta-\left(\frac{y}{\theta}\right)^{\lambda}\right]$$

指数型分布族としてのワイブル分布

$$f(y;\theta,\lambda)=\exp\left[\log\lambda+(\lambda-1)\log y-\lambda\log\theta-\left(\frac{y}{\theta}\right)^{\lambda}\right]$$

$$=\exp\left[a(y)b(\theta)+c(\theta)+d(y)\right]$$

- パラメータ: θ
- $a(y)=y^{\lambda}$　$\lambda=1$のとき正準形となる
- 自然パラメータ　$b(\theta)=-\theta^{-\lambda}$
  - 局外パラメータ: λ
- その他の項　$c(\theta)=-\lambda\log\theta$

$$d(y)=\log\lambda+(\lambda-1)\log y$$

指数型分布族としての二項分布

- 二項分布 $Y \sim Binominal(n, \pi)$
  - 二値(0,1)をとるn回の独立試行に対する分布
  - 確率変数:Y→1の回数。離散値，y=1,2,3,・・・,n
  - 二項係数→n個からy個を選ぶ組み合わせの数

$$\binom{n}{y} = {}_n C_y = C(n, y) = \frac{n!}{y!(n-y)!}$$

- 確率密度関数

$$f(y;\pi) = \binom{n}{y} \pi^y (1-\pi)^{(n-y)}$$

指数型分布族としての二項分布

- 確率密度関数

$$
\begin{aligned}
f(y;\pi) &= \exp\left[\log\binom{n}{y} + y\log\pi + (n-y)\log(1-\pi)\right] \\
&= \exp\left[y\log\pi - y\log(1-\pi) + n\log(1-\pi) + \log\binom{n}{y}\right] \\
&= \exp\left[y\log\frac{\pi}{1-\pi} + n\log(1-\pi) + \log\binom{n}{y}\right] \\
&= \exp\left[a(y)b(\pi) + c(\pi) + d(y)\right]
\end{aligned}
$$

指数型分布族としての二項分布

- パラメータπ
- 正準形 $a(y) = y$
- 自然パラメータ $b(\pi) = \log \frac{\pi}{1-\pi}$
  - 局外パラメータn
- その他の項 $c(\pi) = n \log(1-\pi)$

$$d(y) = \log \binom{n}{y}$$