

2006. 4. 6.

武庫川流域委員会 委員長 松本 誠 様

委員 法西 浩

意 見 書

いろいろとお世話になっております。委員会に基本高水流量に関する意見書を提出します。

基本高水流量算定に^尤度を！

- 基本高水を $4000\text{m}^3/\text{s}$ までとする案と、 $4500\text{m}^3/\text{s}$ あるいはそれ以上との案が併立したまま、整備計画を検討されてきた。
- 整備計画では多方面の総合治水を熱心に、真摯に論議され、実りが多かった。
- いよいよ答申書を提出する段階になり、両論併記の声も聞えてくるようになった。
- 前回流域委員会で、傍聴者の方から、両論併記でなく、1本にしぼるべきだ、との発言があった。
- 私も同感である。
- では、どうするか。私は、自然科学の理論にたえるべき、最も妥当な数値を選定すべきである、と考える。
- 「^尤度を使った選定法あれこれ」が、久保拓弥氏のKuboWeb から出ていたので、それをお届けする。（別紙参照）
- また、委員会への意見書として、しばしば畑委員の意見書を、もっと吟味せよ、と述べている方があり、畑氏の意見書にもとずいて、図を作った。（別紙参照）
- もう時間がないので、両論を併記した、ということのないように、お願いしたい。

尤度を使った推定法あれこれ

久保拓弥 kubo@ees.hokudai.ac.jp

最尤推定 (さいゆうすいてい) — 「最ももっともらしい」パラメーターの推定

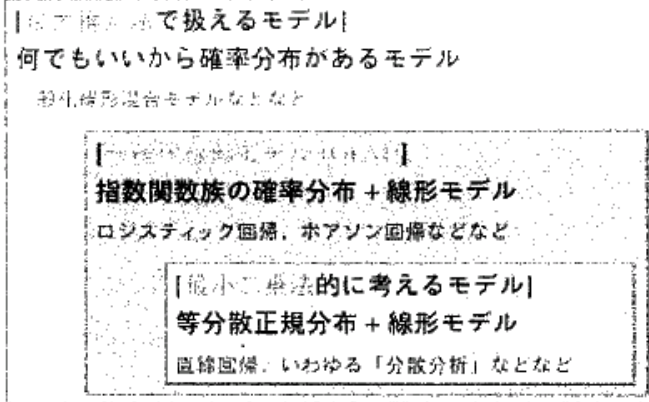
- 「尤」の音読みは「ゆう」、訓読みは「もっともらしい (尤もらしい)」です。
- 尤度とは、ある確率論的モデルを仮定しているときに、その観測データが得られる確率 (あるいは確率密度) …… 簡単には、ある観測データに (あるパラメーターのもとで) 確率論的モデルが「どれくらいあてはまっているか」を表す尺度です
- 最尤推定とは、尤度を「手持ちの観測データのもとで、あるパラメーター値が得られる確率」とみなして (つまり尤度が未知パラメーターの関数とみなして)、尤度を最大化するようなパラメーター値を探索する推定方法です
- 最尤推定法を使う手順は
 1. 尤度方程式を作る: 確率論的モデルを作り (データがどういう確率分布に従うか、確率分布のパラメーターの関数型はどうなってるか)、それを数式として定義する …… これが尤度方程式である
 2. 尤度最大化によって最尤推定値を計算する: そのモデルで定義される「尤度 (ゆうど)」を最大化させるパラメーター推定値を計算する …… これが最尤推定値 (maximum likelihood estimate; MLE)

別のいいかたをすると、確率論的モデルのパラメーターを変えていって、観測データにもっともよく「あてはまる」ところを探索していく方法です。

- 簡単な例
 - あるモデル: コインを投げたときにオモテがでる確率は $p = 0.6$ である
 - 観測データ: 10 枚のコインを投げたら 7 枚がオモテだった
 - 尤度 (二項分布の計算): $(10! / (7! 3!)) \times p^7 \times (1 - p)^3 = 0.21499$
 - 最尤推定: $p = 0.6$ は観測された現象を説明できる良いモデルなのだろうか? 尤度を最大化する p を計算したい (この例だと尤度を最大化するのは $p = 0.7$)
- 実際の計算では尤度ではなく対数尤度を最大化して最尤推定値を計算します。これはそのほうが計算が格段に簡単になるためです。
- 下の図は
 - 最小二乗法は一般化線形モデルの一部
 - 一般化線形モデルは最尤推定法で考える 統計モデルの一部ということをあらわしています。

最尤推定法 — 現代的な統計学の基本

尤度 (後述) つまり「[モデルの尤度関数](#)」を最大にするようなパラメーターを探しまわる、という推定方法



- 2006 年 1 月の [統計学授業](#) の 3 回目でごく簡単な事例について説明しています。
- 2005 年 3 月の [生態学会大会 自由集会](#) の [前口上](#) で最尤推定法のごく簡単な説明しています。

最尤推定と R

- もっとも簡単に最尤推定を行うには free な統計ソフトウェア R を使うのがよいでしょう (参照: [RjpWiki](#) — “オープンソースの統計解析システム R に関する日本語による情報交換を目的とした Wiki”)
- R で使える最尤推定法の関数の一部
 - `glm()`: 一般化線形モデルを計算する関数 -- logistic 回帰や Poisson 回帰などが簡単にできる; [統計学授業 \(2005\)](#) や [計算生態学](#) を参照
 - `nls()` など: 等分散正規分布を仮定している線形・非線型の混合モデルを最尤推定法
 - `optim()`: 究極のわざ, 自分で定義した尤度関数を最大化できる
 - (関連わざ) `stepAIC()`: `glm` など推定した結果に対して AIC を使ってモデル選択できる (つまり変数選択)
- R に関する他のリンクは [計算生態学ページ](#) にあります。

Perl スクリプトなどを使った計算例 (古い内容)

- 「父親として成功するシウリザクラの条件」を最尤推定する方法の [解説メモ \(PDF, 57KB\)](#) 2003 年度の計算

- [KKPA](#) —— 個体ごとに観測期間がばらばらな場合の(一定速度)死亡率推定法の Perl スクリプト
- 確率論的死亡モデル(二項分布と logistic な関数型を仮定)のパラメーターを最尤推定する方法のすごく初歩的な解説([PDF ファイル](#)) 2000.12.08 更新

確率
回帰
年

2000
1900
1800
1700
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
550
500
450
400
350
300
250
200
150
100
50
0

250 500 750 1000 1250 1500 1750 2000 2250 2500 2750 3000 3250 3500 3750 4000 4250 4500 4750 5000 m/s

図 超過確率1/100は、雨量を247mmとしたときの
適切なサンプリングした18個のデータが表1に
示されている。(資料設定1表1参照)
 $(1/100) \times (100 - \text{カバー率}/100) = 1/r_L$
この式で求めた値を図にプロットした。i:1~18
番目まで
(この式は畑委員の意見書からの引用である。)
図の○内の数字は、表1の番号に相応する。

たて軸左は確率・年
右はカバー率%
よこ軸はピーク流量m/s

カバー率

95
90
80
50
0%

