台と形状を母数とするベータ分布の母数推定と歴史的津波規模データへの適用

間野 修平　　数理・推論研究系 教授

0　本ポスターの内容

台(確率密度が0でない区間)の最尤推定は難しいことが知られている．本ポスターでは，台と形状を母数とするベータ分布について，台と形状の推定を分離し，経験ベイズ尤度に基づき台を推定する方法と，その歴史的津波規模データへの適用を紹介する．本発表は渋谷政昭氏(慶應義塾大学名誉教授)との共同研究に基づく．

1　台の最尤推定の難しさ

台を$(c,d)\in\mathbb{R}$とする位置尺度母数c,dのベータ分布の密度は

$$f(x;\theta,c,d)=\frac{1}{d-c}\,g\left(\frac{x-c}{d-c};\theta\right),\quad c<x<d,$$

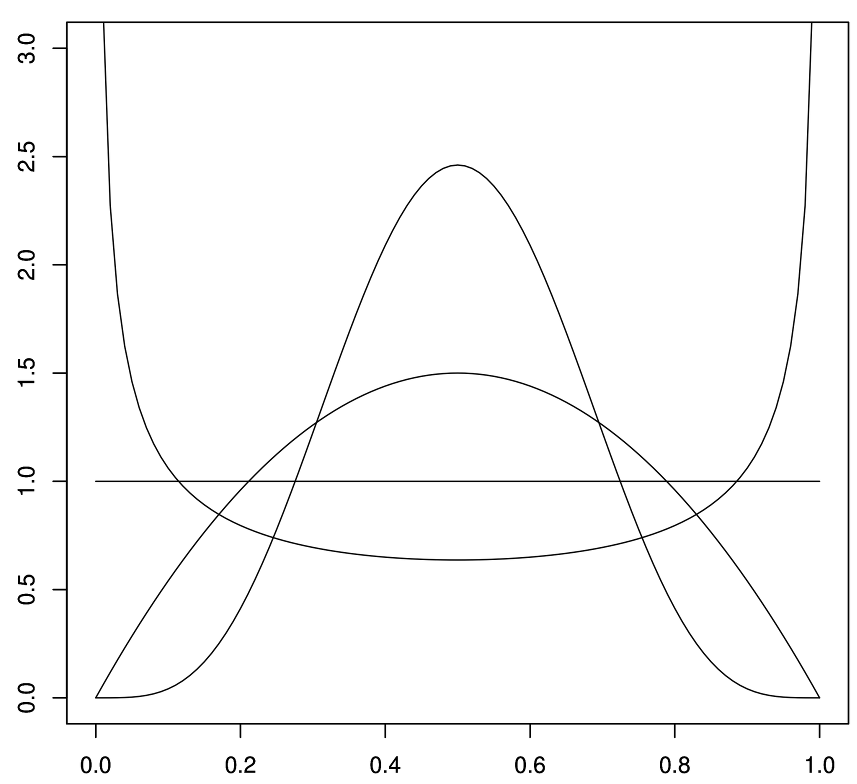
ここで$g(y;\theta)$は$B(a,b)$をベータ函数として

$$g(y;\theta)=\{B(a,b)\}^{-1}\,y^{a-1}(1-y)^{b-1},\quad 0<y<1$$

である．$\theta=(a,b),\ a>0,\ b>0$を形状母数という．台が母数に依存する最尤推定は標準的な正則条件を満たさない．尤度が有界でないので局所最大化を考える．形状母数を既知として，母集団最大値dの局所最尤推定量は

- $b\ge 2$ 漸近正規性がある($b=2$では収束レートが速い)
- $1<b<2$ 漸近正規性がない
- $b\le 1$ 存在しない(標本最大値)

漸近正規性がないと信頼区間の構成が難しく，外挿としての母集団最大値に興味があるのに標本最大値が得られても不毛だろう．分布の対称性から，母集団最小値cと形状母数aについても同様の関係がある．



ベータ分布の密度
$(a,b)=$
$(5,5),(2,2),(1,1),(0.5,0.5)$,
$(c,d)=(0,1)$

2　提案法

標本を昇順に並べて$x_{(1)}<x_{(2)}<\cdots<x_{(n)}$とする．Nagatsukaら(2013)は，$(x_{(i)}-x_{(1)})/(x_{(n)}-x_{(1)}),\ i\in\{1,\dots,n\}$の尤度が位置尺度母数によらないことを利用した形状母数の最尤推定を提案した．まずこれを用いる．Hall-Wang (2005)は，$b\le 1$でも推定値が得られるように，経験ベイズ尤度の最大化を提案した．$p(d)=\frac{d-x_{(n)}}{d-x_{(n-1)}}$を尤度に掛けたもので，$d$が標本最大値に近づくことに対する対数尤度への罰則項になるので非正則性が緩和される．上の形状母数の推定値を代入した尤度に調整係数をかけた罰則項を加え，同様の罰則項をcについても加えたものを経験ベイズ尤度とし，それを最大化する位置尺度母数を位置尺度母数の推定値とする．サイズ100のシミュレーションデータに，モーメント推定に基づく反復法，台と形状の推定を分離しない経験ベイズ尤度の最大化，提案法を適用することを1000回行った結果を2つの表に示す．台と形状の推定を分離しない経験ベイズ尤度が推定値を与えないことが多いのは，局所最大値への収束に失敗するからである($\min(a,b)>2$では漸近的に最尤法に等しい)．提案法は，$\min(a,b)\le 1$と$\max(a,b)>2$の推定を両立し，$\max(a,b)>2$での精度が高い．形状母数が既知であれば位置尺度母数の推定量の正規でない極限分布を得ることはできるが，提案法の信頼区間の構成は難しいので，ブートストラップ法を用いる．

(a,b)	モーメント法	経験ベイズ尤度	提案法
$(0.5,0.5)$	0	2	0
$(1,1)$	0	2	0
$(2,2)$	1	1	0
$(5,5)$	22	49	8
$(1,0.5)$	0	1	0
$(2,0.5)$	0	0	0
$(5,0.5)$	14	163	1
$(2,1)$	0	2	0
$(5,1)$	26	47	8
$(5,2)$	54	10	6

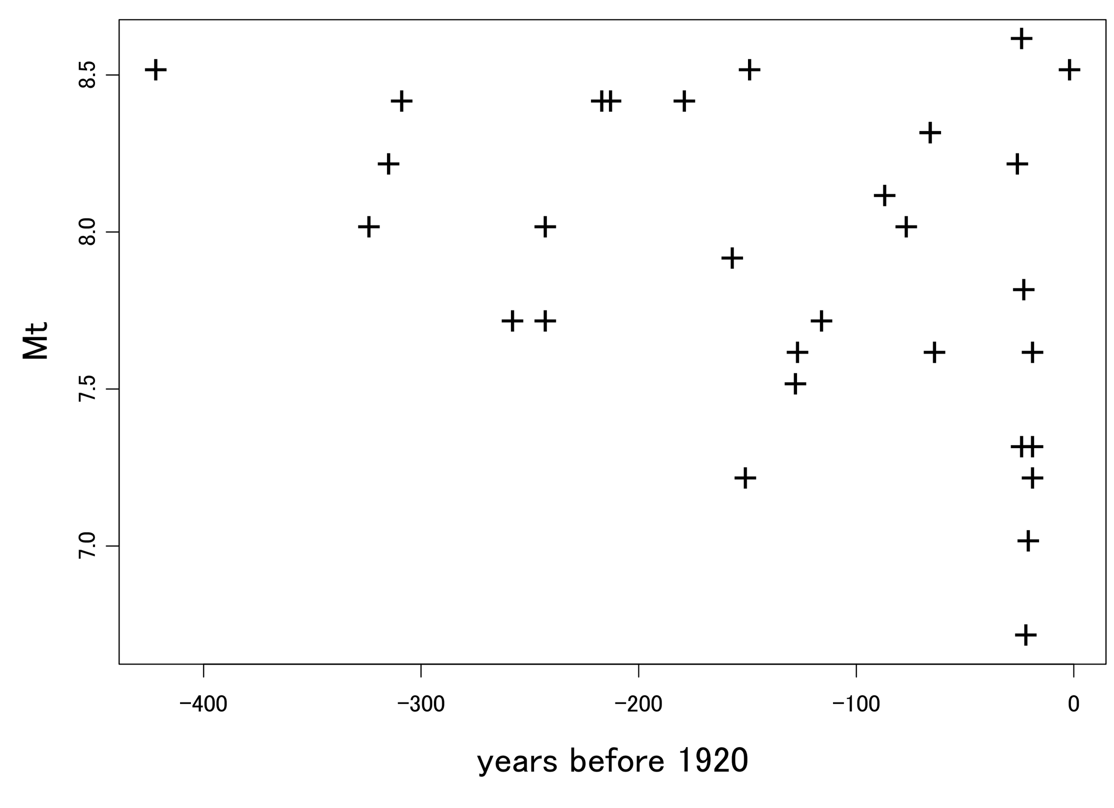
推定値を得られなかった概数

		$\hat{a}$		$\hat{b}$		$\hat{c}$		$\hat{d}$	
(a,b)	方法	Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
$(0.5,0.5)$	モーメント法	0.001	0.084	0.001	0.084	-0.000	0.001	0.000	0.001
	経験ベイズ尤度	0.016	0.069	0.016	0.069	0.000	0.001	-0.000	0.001
	提案法	0.016	0.074	0.017	0.074	0.000	0.001	-0.000	0.001
$(1,1)$	モーメント法	0.007	0.182	0.004	0.184	-0.001	0.012	0.001	0.012
	経験ベイズ尤度	-0.004	0.161	-0.008	0.167	0.004	0.011	-0.004	0.012
	提案法	0.045	0.181	0.040	0.184	0.002	0.012	-0.002	0.012
$(2,2)$	モーメント法	0.006	0.659	-0.011	0.641	-0.004	0.055	0.001	0.053
	経験ベイズ尤度	0.002	0.695	-0.018	0.652	0.005	0.058	-0.008	0.053
	提案法	0.045	0.631	0.025	0.585	0.003	0.055	-0.005	0.049
$(5,5)$	モーメント法	1.358	8.757	1.354	9.552	-0.033	0.274	0.031	0.287
	経験ベイズ尤度	0.829	5.034	0.780	5.049	-0.014	0.205	0.011	0.199
	提案法	0.130	3.507	0.104	4.210	-0.002	0.151	-0.002	0.173

各手法のバイアスと平均2乗誤差の平方根

3　歴史的津波規模データへの適用

阿部勝征氏(東京大学地震研究所名誉教授・故人)は，津波の遡上高から歴史的津波の規模を推定する方法を開発し，1498年以降の歴史的津波の規模の推定値を求めた．津波の規模の分布は一般化パレート分布でモデル化できるが(渋谷・高橋2012)，歴史的津波は，古く小さい津波ほど記録・遺跡が失われやすいため，そのままでは適用できない．年数と規模を与えたときに記録が残る確率が年数の函数と規模の函数の積で与えられ，規模の函数を冪と仮定すれば，歴史的津波の規模の分布はベータ分布でモデル化できることがわかる．1920年までのデータによる津波の規模の母集団最大値の95% 信頼区間は$[8.388, 8.842]$で，標本最大値の8.6(明治三陸地震, 1896) を超える確率は0.006であった．1920年までの100年を見ると歴史的津波は10年に一度程度は起きていて，1920年からの100年に8.6を超える確率は小さいとは言えない．実際，2011年の東北震災の津波は9.1であった．



歴史的津波規模データ
縦軸は規模

参考文献

Shuhei Mano, Masaaki Sibuya. Parameter estimation of generalized beta distributions and its application to a historical tsunami magnitude dataset. to appear in *Pioneering Works on Extreme Value Theory, In Honor of Professor Masaaki Sibuya*, Springer