無情報事前分布による Lomax 分布のパラメータ推定

三國 憲太郎

東京理科大学 黒沢研究室

January 9, 2022

目次

目次

はじめに

Lomax 分布は Pareto Type II 分布とも呼ばれ, 経済, 保険などに用いられることが多い.
本論文では [1] の論文を元に実際にシミュレーションを実行し, その方法やパフォーマンスについて検討する.

ベイズ推定を行う上で, 事後分布が確率の公理を満たさないこと (improper となること) は望ましくない.
ここでは 2 つの無情報事前分布を考え, それぞれに対する事後分布が確率の公理を満たすかどうかを確認する.

目次

モデルの定義

Lomax 分布の pdf は

$$f(x|\beta,\alpha) = \frac{\alpha}{\beta}\left(1 + \frac{x}{\beta}\right)^{-(\alpha+1)} \quad (x \geq 0) \tag{1}$$

ただし $\alpha > 0, \beta > 0$ とする.

$$\mathrm{E}(X) = \frac{\beta}{\alpha - 1} \quad (\alpha > 1).$$

$$\mathrm{Var}(X) = \frac{\alpha\beta^2}{(\alpha-1)^2(\alpha-2)} \quad (\alpha > 2).$$

Lomax 分布は以下の二つの分布を用いて階層的に表現される.

$$X|\beta,\lambda \sim \mathrm{Exponential}\left(\frac{\lambda}{\beta}\right), \quad \lambda|\alpha \sim \mathrm{Gamma}(\alpha, 1).$$

X, λ の結合確率密度関数は

$$f(x|\beta,\lambda)f(\lambda|\alpha) = \frac{1}{\beta\Gamma(\alpha)}\lambda^{\alpha}\exp\left\{-\lambda\left(1+\frac{x}{\beta}\right)\right\}.$$

λ で積分すると

$$
\begin{aligned}
f(x|\beta,\alpha) &= \frac{1}{\beta\Gamma(\alpha)}\int_0^\infty \lambda^\alpha \exp\left\{-\lambda\left(1+\frac{x}{\beta}\right)\right\} d\lambda \\
&= \frac{1}{\beta\Gamma(\alpha)}\Gamma(\alpha+1)\left(1+\frac{x}{\beta}\right)^{-(\alpha+1)} \\
&= \frac{\alpha}{\beta}\left(1+\frac{x}{\beta}\right)^{-(\alpha+1)}.
\end{aligned}
$$

よって $X \sim \mathrm{Lomax}(\beta,\alpha)$.

$\boldsymbol{X} = (X_1, \ldots, X_n)$ を (1) からのランダム標本, $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ とする.

$$
\begin{aligned}
f(\boldsymbol{\lambda}|\boldsymbol{x}, \beta, \alpha) &= \frac{f(\boldsymbol{x}|\boldsymbol{\lambda}, \beta, \alpha) f(\boldsymbol{\lambda}, \beta, \alpha)}{f(\boldsymbol{x}, \beta, \alpha)} \\
&= \frac{f(\boldsymbol{x}|\beta, \boldsymbol{\lambda}) f(\boldsymbol{\lambda}|\beta, \alpha) \pi(\beta, \alpha)}{f(\boldsymbol{x}|\beta, \alpha) \pi(\beta, \alpha)} \\
&\propto f(\boldsymbol{x}|\beta, \boldsymbol{\lambda}) f(\boldsymbol{\lambda}|\alpha) \\
&\propto \prod_{i=1}^{n} \lambda_i \exp\left\{\frac{-\lambda_i x_i}{\beta}\right\} \prod_{i=1}^{n} \lambda_i^{\alpha-1} \exp\{-\lambda_i\} \\
&= \prod_{i=1}^{n} \lambda_i^{\alpha} \exp\left\{-\lambda_i \left(1 + \frac{x_i}{\beta}\right)\right\}.
\end{aligned}
$$

$$\therefore \lambda_i | x_i, \beta, \alpha \sim \text{Gamma}\left(\alpha + 1, 1 + \frac{x_i}{\beta}\right).$$

次に, α, β の条件付き分布を考える.

$$f(\alpha|\boldsymbol{x},\boldsymbol{\lambda},\beta) \propto f(\boldsymbol{\lambda}|\alpha)\pi(\beta,\alpha) \propto [\Gamma(\alpha)]^{-n}\left(\prod_{i=1}^{n}\lambda_i\right)^{\alpha-1}\pi(\beta,\alpha) \tag{2}$$

$$\begin{aligned}
\because f(\alpha|\boldsymbol{x},\boldsymbol{\lambda},\beta) &= \frac{f(\boldsymbol{\lambda}|\alpha,\beta,\boldsymbol{x})f(\alpha,\beta,\boldsymbol{x})}{f(\boldsymbol{x},\boldsymbol{\lambda},\beta)} \\
&= \frac{f(\boldsymbol{\lambda}|\alpha)f(\boldsymbol{x}|\alpha,\beta)\pi(\beta,\alpha)}{f(\boldsymbol{x},\boldsymbol{\lambda},\beta)} \\
&= \frac{f(\boldsymbol{\lambda}|\alpha)f(\boldsymbol{x}|\beta)\pi(\beta,\alpha)}{f(\boldsymbol{x},\boldsymbol{\lambda},\beta)}.
\end{aligned}$$

$$f(\beta|\boldsymbol{x}, \boldsymbol{\lambda}, \alpha) \propto f(\boldsymbol{x}|\beta, \boldsymbol{\lambda})\pi(\beta, \alpha) \propto \beta^{-n}\exp\left\{-\frac{1}{\beta}\sum_{i=1}^{n}\lambda_i x_i\right\}\pi(\beta, \alpha) \tag{3}$$

$$\begin{aligned}
\because f(\beta|\boldsymbol{x}, \boldsymbol{\lambda}, \alpha) &= \frac{f(\boldsymbol{x}|\beta, \boldsymbol{\lambda}, \alpha)f(\beta, \boldsymbol{\lambda}, \alpha)}{f(\boldsymbol{x}, \boldsymbol{\lambda}, \alpha)} \\
&= \frac{f(\boldsymbol{x}|\beta, \boldsymbol{\lambda})f(\boldsymbol{\lambda}|\beta, \alpha)\pi(\beta, \alpha)}{f(\boldsymbol{x}, \boldsymbol{\lambda}, \alpha)} \\
&= \frac{f(\boldsymbol{x}|\beta, \boldsymbol{\lambda})f(\boldsymbol{\lambda}|\alpha)\pi(\beta, \alpha)}{f(\boldsymbol{x}, \boldsymbol{\lambda}, \alpha)}.
\end{aligned}$$

目次

無情報事前分布

Jeffreys の事前分布は以下のように定義される.

$$\pi_J \propto |I(\beta,\alpha)|^{1/2}.$$

ここで

$$I(\beta,\alpha) = n\begin{bmatrix} \dfrac{\alpha}{\beta^2(\alpha+2)} & -\dfrac{1}{\beta(\alpha+1)} \\ -\dfrac{1}{\beta(\alpha+1)} & \dfrac{1}{\alpha^2} \end{bmatrix}.$$

よって

$$\pi_J(\beta,\alpha) \propto \frac{1}{\beta(\alpha+1)\alpha^{1/2}(\alpha+2)^{1/2}} \quad (\beta,\alpha>0). \tag{4}$$

パラメータの事前独立性を仮定した場合の Jeffreys の事前分布は

$$\pi_{IJ}(\beta, \alpha) \propto \pi(\beta)\pi(\alpha) = \frac{1}{\beta\alpha} \quad (\beta, \alpha > 0) \tag{5}$$

となる。

定理 1

(4) の Jeffreys の事前分布のもとで, 事後分布は proper となる.

証明:
(4) の Jeffreys の事前分布のもとで, β, α の事後分布は

$$\pi(\beta, \alpha|\boldsymbol{x}) \propto \frac{\alpha^{n-1/2}\beta^{-(n+1)}}{(\alpha+1)(\alpha+2)^{1/2}} \prod_{i=1}^{n} \left(1 + \frac{x_i}{\beta}\right)^{-(\alpha+1)}.$$

この式の積分が有限であることを示す.

まず,

$$\int_0^\infty \beta^{-(n+1)} \prod_{i=1}^{n} \left(1 + \frac{x_i}{\beta}\right)^{-(\alpha+1)} d\beta \tag{6}$$

を考える.

$y = \min\{x_1, \ldots, x_n\}$ とすると,

$$\left(1 + \frac{x_i}{\beta}\right)^{\alpha+1} \geq \left(1 + \frac{y}{\beta}\right)^{\alpha+1}.$$

$\alpha \geq 0,\ i = 1, \ldots, n$ に対し,

$$\prod_{i=1}^{n} \left(1 + \frac{x_i}{\beta}\right)^{-(\alpha+1)} \leq \left(1 + \frac{y}{\beta}\right)^{-n(\alpha+1)}.$$

よって

$$\int_0^\infty \beta^{-(n+1)} \prod_{i=1}^{n} \left(1 + \frac{x_i}{\beta}\right)^{-(\alpha+1)} d\beta \leq \int_0^\infty \beta^{-(n+1)} \left(1 + \frac{y}{\beta}\right)^{-n(\alpha+1)} d\beta.$$

変数変換 $u = \frac{y}{\beta},\ du = -\frac{y}{\beta^2}d\beta$ **により**

$$
\begin{aligned}
\int_0^\infty \beta^{-(n+1)} \left(1 + \frac{y}{\beta}\right)^{-n(\alpha+1)} d\beta &= \frac{1}{y^n} \int_0^\infty \frac{u^{n-1}}{(1+u)^{n\alpha+n}} du \\
&= \frac{1}{y^n} B(n, n\alpha) \\
&= \frac{1}{y^n} \frac{\Gamma(n)\Gamma(n\alpha)}{\Gamma(n\alpha+n)} \\
&= \frac{(n-1)!}{y^n} \frac{1}{\prod_{j=0}^{n-1}(n\alpha+j)}
\end{aligned}
$$

よって

$$\int_0^\infty \int_0^\infty \frac{\alpha^{n-1/2}\beta^{-(n+1)}}{(\alpha+1)(\alpha+2)^{1/2}} \prod_{i=1}^{n} \left(1 + \frac{x_i}{\beta}\right)^{-(\alpha+1)} d\beta d\alpha$$

$$\leq \frac{(n-1)!}{y^n} \int_0^\infty \frac{\alpha^{n-1/2}}{(\alpha+1)(\alpha+2)^{1/2}} \frac{1}{\prod_{j=0}^{n-1}(n\alpha+j)} d\alpha$$

$$= \frac{(n-1)!}{y^n} \left(\int_0^1 f(\alpha)d\alpha + \int_1^\infty f(\alpha)d\alpha \right),$$

ただし,

$$f(\alpha) = \frac{\alpha^{n-1/2}}{(\alpha+1)(\alpha+2)^{1/2}} \frac{1}{\prod_{j=0}^{n-1}(n\alpha+j)}.$$

ここで, $j \geq 1,\ \alpha > 0$ に対し $(n\alpha + j) \geq 1$ なので

$$\frac{1}{\prod_{j=0}^{n-1}(n\alpha + j)} \leq \frac{1}{n\alpha}.$$

また, $\dfrac{1}{\alpha+1} \leq 1,\ \dfrac{1}{(\alpha+2)^{1/2}} < \dfrac{1}{\sqrt{2}}$ より

$$\begin{aligned}\int_0^1 f(\alpha)d\alpha &\leq \frac{1}{n\sqrt{2}}\int_0^1 \alpha^{n-3/2}d\alpha \\ &= \frac{1}{n(n-1/2)\sqrt{2}}.\end{aligned}$$

一方, $(n\alpha + j) \geq n\alpha$ なので

$$\frac{1}{\prod_{j=1}^{n-1}(n\alpha + j)} \leq \frac{1}{n^n}\frac{1}{\alpha^n}.$$

また, $\frac{1}{\alpha+1} \leq \frac{1}{\alpha}$, $\frac{1}{(\alpha+2)^{1/2}} \leq \frac{1}{\alpha^{1/2}}$ より

$$\begin{aligned}\int_1^\infty f(\alpha)d\alpha &\leq \frac{1}{n^n}\int_1^\infty \alpha^{-2}d\alpha \\ &= \frac{1}{n^n}.\end{aligned}$$

以上より

$$
\begin{aligned}
&\frac{(n-1)!}{y^n}\left(\int_0^1 f(\alpha)d\alpha + \int_1^\infty f(\alpha)d\alpha\right) \\
\leq\ &\frac{(n-1)!}{y^n}\left(\frac{1}{n(n-1/2)\sqrt{2}} + \frac{1}{n^n}\right) \\
<\ &\infty.
\end{aligned}
$$

∴ Jeffreys の事前分布を用いると事後分布は proper となる. □

定理 2

(5) のパラメータ間の事前独立性を仮定した Jeffreys の事前分布のもとで, 事後分布は improper となる.

証明:

(5) のパラメータ間の事前独立性を仮定した Jeffreys の事前分布のもとで, β, α の結合事後分布は

$$\pi(\beta, \alpha | \boldsymbol{x}) \propto \alpha^{n-1} \beta^{-(n+1)} \prod_{i=1}^{n} \left(1 + \frac{x_i}{\beta}\right)^{-(\alpha+1)}.$$

この式の積分が有限でないことを示す.

$w = \max\{x_1, \ldots, x_n\}$ とすると

$$\int_0^\infty \beta^{-(n+1)} \prod_{i=1}^{n} \left(1 + \frac{x_i}{\beta}\right)^{-n(\alpha+1)} d\beta$$

$$\geq \int_0^\infty \beta^{-(n+1)} \left(1 + \frac{w}{\beta}\right)^{-n(\alpha+1)} d\beta.$$

定理 1 の証明と同じように, $v = \dfrac{w}{\beta},\ dv = -\dfrac{w}{\beta^2} d\beta$ の変数変換を行うと

$$
\begin{aligned}
\int_0^{\infty} \beta^{-(n+1)} \left(1 + \frac{w}{\beta}\right)^{-n(\alpha+1)} d\beta &= \frac{1}{w^n} \int_0^{\infty} \frac{v^{n-1}}{(1+v)^{n\alpha+n}} dv \\
&= \frac{1}{w^n} \frac{\Gamma(n)\Gamma(n\alpha)}{\Gamma(n\alpha+n)} \\
&= \frac{(n-1)!}{w^n} \frac{1}{\prod_{j=0}^{n-1}(n\alpha+j)}.
\end{aligned}
$$

これを用いて

$$
\begin{aligned}
&\int_0^\infty \int_0^\infty \alpha^{n-1}\beta^{-(n+1)} \prod_{i=1}^{n}\left(1+\frac{x_i}{\beta}\right)^{-(\alpha+1)} d\beta d\alpha \\
\geq\ & \frac{(n-1)!}{w^n}\int_0^\infty \frac{\alpha^{n-1}}{\prod_{j=0}^{n-1}(n\alpha+j)} d\alpha \\
\geq\ & \frac{(n-1)!}{w^n}\int_1^\infty \frac{\alpha^{n-1}}{\prod_{j=0}^{n-1}(n\alpha+j)} d\alpha .
\end{aligned}
$$

ここで, $\alpha \geq 1$, $j < n$ **に対し**, $n\alpha + j < n\alpha + n = n(\alpha + 1) \leq 2n\alpha$ **となることから**

$$\frac{1}{\prod_{j=0}^{n-1}(n\alpha + j)} \geq \frac{1}{(2n)^n}\frac{1}{\alpha^n}. \tag{7}$$

よって

$$\frac{(n-1)!}{w^n}\int_1^\infty \frac{\alpha^{n-1}}{\prod_{j=0}^{n-1}(n\alpha + j)}d\alpha \geq \frac{(n-1)!}{(2wn)^n}\int_1^\infty \alpha^{-1}d\alpha = \infty$$

$\therefore$ **独立な** Jeffreys **の事前分布を用いると事後分布は** improper **となる。** □

目次

シミュレーション

通常の Jeffreys の事前分布を用いると事後分布は proper となり, パラメータ間の事前独立性を仮定した Jeffreys の事前分布を用いると, 事後分布は improper となる. よって, 通常の Jeffreys の事前分布を用いてシミュレーションを行う.

α の事後分布は (2) に Jeffreys の事前分布を代入して

$$f(\alpha|\boldsymbol{x},\boldsymbol{\lambda},\beta) \propto \frac{1}{(\alpha+1)\alpha^{1/2}(\alpha+2)^{1/2}[\Gamma(\alpha)]^n}\left(\prod_{i=1}^{n}\lambda_i\right)^{\alpha-1}.$$

メトロポリス法によりこの分布に従う乱数を発生させる.

同様に β の事後分布も (3) に Jeffreys の事前分布を代入して

$$f(\beta|\boldsymbol{x}, \boldsymbol{\lambda}, \alpha) \propto \beta^{-(n+1)} \exp\left\{-\frac{1}{\beta}\sum_{i=1}^{n}\lambda_i x_i\right\},$$

つまり,

$$\beta|\boldsymbol{x}, \boldsymbol{\lambda}, \alpha \sim \mathrm{IG}\left(n, \sum_{i=1}^{n}\lambda_i x_i\right).$$

推定の評価指標として平均相対誤差 (MRE) と平均二乗誤差 (MSE) を用いる.

$$\mathrm{MRE} = \frac{1}{N}\sum_{j=1}^{N}\frac{\hat{\theta}^{(j)}}{\theta}, \quad \mathrm{MSE} = \frac{1}{N}\sum_{j=1}^{N}(\hat{\theta}^{(j)} - \theta)^2.$$

真のパラメータを $\alpha = 1.5, \beta = 2$ とし, サイズ $n = (100, 110, \ldots, 250)$ のサンプルをそれぞれ 1000 個発生させてシミュレーションを行なう.

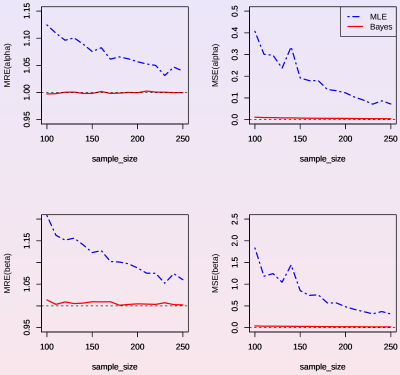

Figure: シミュレーションの結果

ベイズ推定は最尤推定よりも精度が優れているという結果になった.

目次

問題点と改善策の提案

先のシミュレーションにおいて λ_i を用いたが, これは未知のパラメータ α に依存する分布からの確率変数であり, 通常これを得ることができない.
これに対する改善案を二つ考え, シミュレーションを行なった.

提案 1

まずは λ_i を用いずに事後分布を求める.
β, α の同時事後分布は

$$\pi(\beta, \alpha | \boldsymbol{x}) \propto \frac{\alpha^{n-1/2}\beta^{-(n+1)}}{(\alpha+1)(\alpha+2)^{1/2}} \prod_{i=1}^{n} \left(1 + \frac{x_i}{\beta}\right)^{-(\alpha+1)} .$$

これを用いて 2 変数のメトロポリス法により事後分布からの乱数を生成する.

提案 1

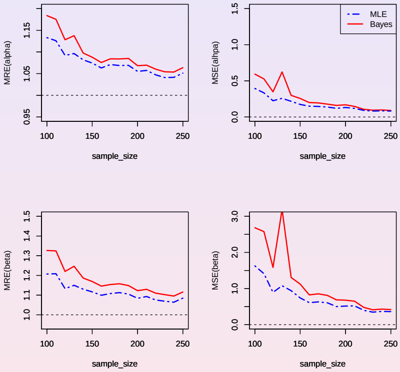

Figure: 提案 1 のシミュレーションの結果

提案 1

提案 1 では λ_i を用いずに推定することができるが, 最尤推定と比べて精度が低下した.

提案 2

次に λ_i を推定し, それを用いてパラメータ推定を行う.
λ_i の条件付き分布は

$$\lambda_i | x_i, \beta, \alpha \sim \mathrm{Gamma}\left(\alpha + 1, 1 + \frac{x_i}{\beta}\right)$$

であった. これを最尤推定値 $\hat{\beta}$, $\hat{\alpha}$ を用いて

$$\lambda_i | x_i, \hat{\beta}, \hat{\alpha} \sim \mathrm{Gamma}\left(\hat{\alpha} + 1, 1 + \frac{x_i}{\hat{\beta}}\right)$$

として各 λ_i を推定する.

提案 2

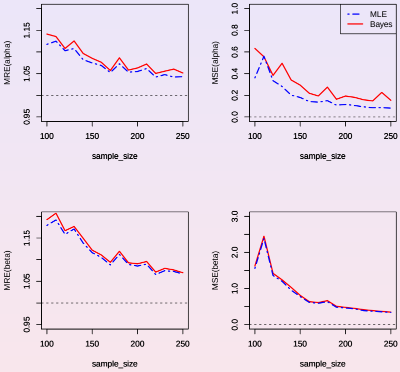

Figure: 提案 2 のシミュレーションの結果

提案 2

提案 2 では λ_i が得られなくても推定が可能であるが, 提案 1 と同様に最尤推定と比べて精度が低下した.

目次

まとめ

[1] のシミュレーションの再現ではベイズ推定の方が精度が良いことが確認できたが, 問題点があり, それについて検討した.
今回提案した二つの手法は現実的に用いることは可能であるが, 推定の精度は最尤推定より劣っていて, 実用的ではない.
今後は λ_i を別の方法で生成, または λ_i を使わない方法によってより精度の高い推定法を考えることが課題である.

参考文献

Paulo H. Ferreira, Eduardo Ramos, Pedro L. Ramos, Jhon F.B. Gonzales, Vera L.D. Tomazella, Ricardo S. Ehlers, Eveliny B. Silva, and Francisco Louzada. Objective bayesian analysis for the lomax distribution. *Statistics and Probability Letters*, 2019.