

$$\max_{\theta} l(\theta)$$

と

$$\max_{\theta} \log l(\theta)$$

の θ の解はともに同じものであることに注意。 $\log l(\theta)$ を対数尤度関数と呼ぶ。

最尤推定量の性質： n が大きいとき、

$$\widehat{\theta} \sim N(\theta, \sigma_{\theta}^2)$$

ただし、

$$\begin{aligned} \sigma_{\theta}^2 &= \frac{1}{\sum_{i=1}^n \mathbb{E}\left[\left(\frac{d \log f(X_i; \theta)}{d\theta}\right)^2\right]} \\ &= \frac{1}{\sum_{i=1}^n \mathbb{E}\left[\frac{d^2 \log f(X_i; \theta)}{d\theta^2}\right]} \end{aligned}$$

θ がベクトル ($k \times 1$) の場合, n が大きいとき,

$$\widehat{\theta} \sim N(\theta, \Sigma_{\theta})$$

ただし,

$$\begin{aligned}\Sigma_{\theta} &= \left(\sum_{i=1}^n \mathbb{E} \left[\left(\frac{\partial \log f(X_i; \theta)}{\partial \theta} \right) \left(\frac{\partial \log f(X_i; \theta)}{\partial \theta} \right)' \right] \right)^{-1} \\ &= - \left(\sum_{i=1}^n \mathbb{E} \left[\frac{\partial^2 \log f(X_i; \theta)}{\partial \theta \partial \theta'} \right] \right)^{-1}\end{aligned}$$

例 1: 正規母集団 $N(\mu, \sigma^2)$ からの標本値 x_1, x_2, \dots, x_n を用いて,

- (1) σ^2 が既知のとき, μ の最尤推定値と最尤推定量
- (2) σ^2 が未知のとき, μ と σ^2 の最尤推定値と最尤推定量をそれぞれ求める。

[解] $N(\mu, \sigma^2)$ の密度関数は,

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right)$$

となる。したがって、互いに独立な X_1, X_2, \dots, X_n の結合分布は,

$$\begin{aligned} f(x_1, x_2, \dots, x_n; \mu, \sigma^2) &\equiv \prod_{i=1}^n f(x_i; \mu, \sigma^2) \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x_i - \mu)^2\right) \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right) \end{aligned}$$

となる。

(1) σ^2 が既知のとき、尤度関数 $l(\mu)$ は,

$$l(\mu) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right)$$

となる。

$l(\mu)$ を最大にする μ と $\log l(\mu)$ を最大にする μ は同じになる。

したがって、対数尤度関数は、

$$\log l(\mu) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$$

となり、

$$\frac{d \log l(\mu)}{d\mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0$$

となる μ を求める。 μ の解を $\widehat{\mu}$ とすると、 μ の最尤推定値は、

$$\widehat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i \equiv \bar{x}$$

を得る。

さらに、観測値 x_1, x_2, \dots, x_n をその確率変数 X_1, X_2, \dots, X_n で置き換えて、 μ の最尤推定量は、

$$\widehat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i \equiv \bar{X}$$

となる。

$\hat{\mu}$ の分散を求めるために、

$$\log f(X_i; \mu) = -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(X_i - \mu)^2$$

$$\frac{d \log f(X_i; \mu)}{d\mu} = \frac{1}{\sigma^2}(X_i - \mu)$$

$$\left(\frac{d \log f(X_i; \mu)}{d\mu}\right)^2 = \frac{1}{\sigma^4}(X_i - \mu)^2$$

$$E\left[\left(\frac{d \log f(X_i; \mu)}{d\mu}\right)^2\right] = \frac{1}{\sigma^4}E[(X_i - \mu)^2] = \frac{1}{\sigma^2}$$

と計算される。

最尤推定量の性質から、 n が大きいとき、

$$\hat{\mu} \sim N(\mu, \sigma_\mu^2)$$

ただし、

$$\sigma_\mu^2 = \frac{1}{\sum_{i=1}^n E\left[\left(\frac{d \log f(X_i; \mu)}{d\mu}\right)^2\right]} = \frac{\sigma^2}{n}$$

この場合は、 n の大きさに関わらず、 $\widehat{\mu} \sim N(\mu, \sigma_\mu^2)$ が成り立つ。

(2) σ^2 が未知のとき、 μ と σ^2 の尤度関数は、

$$l(\mu, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right)$$

となる。

対数尤度関数は、

$$\begin{aligned} \log l(\mu, \sigma^2) &= -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 \\ &\quad - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \end{aligned}$$

と表される。

μ と σ^2 について、最大化するためには、

$$\begin{aligned} \frac{\partial \log l(\mu, \sigma^2)}{\partial \mu} &= \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - \mu) = 0 \\ \frac{\partial \log l(\mu, \sigma^2)}{\partial \sigma^2} &= -\frac{n}{2} \frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2 = 0 \end{aligned}$$

の連立方程式を解く。

μ, σ^2 の解を $\widehat{\mu}, \widehat{\sigma}^2$ とすると、最尤推定値は、

$$\widehat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i \equiv \bar{x}$$
$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \widehat{\mu})^2 \equiv \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

となる。

観測値 x_1, x_2, \dots, x_n をその確率変数 X_1, X_2, \dots, X_n で置き換えて、 μ, σ^2 の最尤推定量は、

$$\widehat{\mu} = \frac{1}{n} \sum_{i=1}^n X_i \equiv \bar{X}$$
$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \widehat{\mu})^2 \equiv \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

となる。

σ^2 の最尤推定量 $\hat{\sigma}^2$ は, σ^2 の不偏推定量 $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ とは異なることに注意。

$\theta = (\mu, \sigma^2)'$ とする。 n が大きいとき,

$$\hat{\theta} \sim N(\theta, \Sigma_\theta)$$

ただし,

$$\Sigma_\theta = -\left(\sum_{i=1}^n \mathbb{E}\left[\frac{\partial^2 \log f(X_i; \theta)}{\partial \theta \partial \theta'}\right]\right)^{-1}$$

$$\log f(X_i; \theta) = -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} (X_i - \mu)^2$$

$$\frac{\partial \log f(X_i; \theta)}{\partial \theta} = \begin{pmatrix} \frac{\partial \log f(X_i; \theta)}{\partial \mu} \\ \frac{\partial \log f(X_i; \theta)}{\partial \sigma^2} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{\sigma^2}(X_i - \mu) \\ -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(X_i - \mu)^2 \end{pmatrix}$$

$$\begin{aligned} & \frac{\partial^2 \log f(X_i; \theta)}{\partial \theta \partial \theta'} \\ &= \begin{pmatrix} \frac{\partial^2 \log f(X_i; \theta)}{\partial \mu^2} & \frac{\partial^2 \log f(X_i; \theta)}{\partial \mu \partial \sigma^2} \\ \frac{\partial^2 \log f(X_i; \theta)}{\partial \sigma^2 \partial \mu} & \frac{\partial^2 \log f(X_i; \theta)}{\partial (\sigma^2)^2} \end{pmatrix} \\ &= \begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{1}{\sigma^4}(X_i - \mu) \\ -\frac{1}{\sigma^4}(X_i - \mu) & \frac{1}{2\sigma^4} - \frac{1}{\sigma^6}(X_i - \mu)^2 \end{pmatrix} \end{aligned}$$

$$\mathbb{E} \left[\frac{\partial^2 \log f(X_i; \theta)}{\partial \theta \partial \theta'} \right]$$

$$\begin{aligned}
&= \begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{1}{\sigma^4}E(X_i - \mu) \\ -\frac{1}{\sigma^4}E(X_i - \mu) & \frac{1}{2\sigma^4} - \frac{1}{\sigma^6}E[(X_i - \mu)^2] \end{pmatrix} \\
&= \begin{pmatrix} -\frac{1}{\sigma^2} & 0 \\ 0 & -\frac{1}{2\sigma^4} \end{pmatrix}
\end{aligned}$$

よって,

$$\begin{aligned}
\Sigma_{\theta} &= -\left(\sum_{i=1}^n E\left[\frac{\partial^2 \log f(X_i; \theta)}{\partial \theta \partial \theta'}\right]\right)^{-1} \\
&= \begin{pmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{pmatrix}
\end{aligned}$$

まとめると, μ , σ^2 の最尤推定量 $\hat{\mu} = (1/n) \sum_{i=1}^n X_i$, $\hat{\sigma}^2 = (1/n) \sum_{i=1}^n (X_i - \bar{X})^2$

の分布は, n が大きいとき,

$$\begin{pmatrix} \widehat{\mu} \\ \widehat{\sigma^2} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}, \begin{pmatrix} \frac{\sigma^2}{n} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{pmatrix}\right)$$

となる。

例 2: X_1, X_2, \dots, X_n は互いに独立で, それぞれパラメータ p を持ったベルヌイ分布に従うものとする。すなわち, X_i の確率関数は,

$$f(x; p) = p^x(1-p)^{1-x} \quad x = 0, 1$$

となる。

このとき尤度関数は,

$$l(p) = \prod_{i=1}^n f(x_i; p) = \prod_{i=1}^n p^{x_i}(1-p)^{1-x_i}$$

となり、対数尤度関数は、

$$\begin{aligned}\log l(p) &= \sum_{i=1}^n \log f(x_i; p) \\ &= \log(p) \sum_{i=1}^n x_i + \log(1-p) \sum_{i=1}^n (1-x_i) \\ &= \log(p) \sum_{i=1}^n x_i + \log(1-p)(n - \sum_{i=1}^n x_i)\end{aligned}$$

となる。

$\log l(p)$ を最大にする p を求める。

$$\frac{d \log l(p)}{dp} = \frac{1}{p} \sum_{i=1}^n x_i - \frac{1}{1-p} (n - \sum_{i=1}^n x_i) = 0$$

したがって、 p について解くと、 p の最尤推定値 \hat{p} は、

$$\hat{p} = \frac{1}{n} \sum_{i=1}^n x_i$$

となる。

さらに、 x_i を X_i で置き換えて、 p の最尤推定量 \hat{p} は、

$$\hat{p} = \frac{1}{n} \sum_{i=1}^n X_i$$

となる。

\hat{p} の分布を求める。

$$\log f(X_i; p) = X_i \log(p) + (1 - X_i) \log(1 - p)$$

$$\frac{d \log f(X_i; p)}{dp} = \frac{X_i}{p} - \frac{1 - X_i}{1 - p} = \frac{X_i - p}{p(1 - p)}$$

$$E\left[\left(\frac{d \log f(X_i; p)}{dp}\right)^2\right] = \frac{E[(X_i - p)^2]}{p^2(1 - p)^2}$$

$$E[(X_i - p)^2] = \sum_{x_i=0}^1 (x_i - p)^2 f(x_i; p)$$

$$\begin{aligned}
&= \sum_{x_i=0}^1 (x_i - p)^2 p^{x_i} (1 - p)^{1-x_i} \\
&= p^2(1 - p) + (1 - p)^2 p = p(1 - p)
\end{aligned}$$

$$\frac{1}{\sum_{i=1}^n \mathbb{E}\left[\left(\frac{d \log f(X_i; p)}{dp}\right)^2\right]} = \frac{p(1 - p)}{n}$$

したがって、

$$\widehat{p} \sim N\left(p, \frac{p(1 - p)}{n}\right)$$

を得る。

例 3: X_1, X_2, \dots, X_n は互いに独立で、それぞれパラメータ λ を持ったポアソン分布に従うものとする。すなわち、 X_i の確率関数は、

$$f(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \quad x = 0, 1, 2, \dots$$

となる。

このとき尤度関数は,

$$l(\lambda) = \prod_{i=1}^n f(x_i; \lambda) = \prod_{i=1}^n \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

となり, 対数尤度関数は,

$$\begin{aligned} \log l(\lambda) &= \sum_{i=1}^n \log f(x_i; \lambda) \\ &= \log(\lambda) \sum_{i=1}^n x_i - n\lambda - \sum_{i=1}^n \log(x_i!) \end{aligned}$$

となる。

$\log l(\lambda)$ を最大にする p を求める。

$$\frac{d \log l(\lambda)}{d\lambda} = \frac{1}{\lambda} \sum_{i=1}^n x_i - n = 0$$

したがって, λ について解くと, λ の最尤推定値 $\widehat{\lambda}$ は,

$$\widehat{\lambda} = \frac{1}{n} \sum_{i=1}^n x_i$$

となる。

さらに、 x_i を X_i で置き換えて、 λ の最尤推定量 $\widehat{\lambda}$ は、

$$\widehat{\lambda} = \frac{1}{n} \sum_{i=1}^n X_i$$

となる。

$\widehat{\lambda}$ の分布を求める。

$$\log f(X_i; \lambda) = X_i \log(\lambda) - \lambda - \log(X_i!)$$

$$\frac{d \log f(X_i; \lambda)}{d\lambda} = \frac{X_i}{\lambda} - 1$$

$$\frac{d^2 \log f(X_i; \lambda)}{d\lambda^2} = -\frac{X_i}{\lambda^2}$$

$$E\left(\frac{d^2 \log f(X_i; \lambda)}{d\lambda^2}\right) = \frac{E(X_i)}{\lambda^2}$$

$$\begin{aligned}
E(X_i) &= \sum_{x=0}^{\infty} x f(x; \lambda) \\
&= \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!} \\
&= \sum_{x=1}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!} \\
&= \sum_{x=1}^{\infty} \lambda \frac{\lambda^{x-1} e^{-\lambda}}{(x-1)!} \\
&= \sum_{x=0}^{\infty} \lambda \frac{\lambda^x e^{-\lambda}}{x!} \\
&= \lambda
\end{aligned}$$

$$-\frac{1}{\sum_{i=1}^n E\left(\frac{d^2 \log f(X_i; \lambda)}{d\lambda^2}\right)} = \frac{\lambda}{n}$$

したがって,

$$\widehat{\lambda} \sim N\left(\lambda, \frac{\lambda}{n}\right)$$

を得る。

例 4: X_1, X_2, \dots, X_n は互いに独立で, それぞれパラメータ λ を持った指数分布に従うものとする。すなわち, X_i の密度関数は,

$$f(x; \lambda) = \lambda e^{-\lambda x} \quad x > 0$$

となる。

このとき尤度関数は,

$$l(\lambda) = \prod_{i=1}^n f(x_i; \lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i}$$

となり, 対数尤度関数は,

$$\log l(\lambda) = \sum_{i=1}^n \log f(x_i; \lambda)$$

$$= n \log \lambda - \lambda \sum_{i=1}^n x_i$$

となる。

$\log l(\lambda)$ を最大にする p を求める。

$$\frac{d \log l(\lambda)}{d \lambda} = \frac{n}{\lambda} - \sum_{i=1}^n x_i = 0$$

したがって、 λ について解くと、 λ の最尤推定値 $\widehat{\lambda}$ は、

$$\widehat{\lambda} = \frac{n}{\sum_{i=1}^n x_i}$$

となる。

さらに、 x_i を X_i で置き換えて、 λ の最尤推定量 $\widehat{\lambda}$ は、

$$\widehat{\lambda} = \frac{n}{\sum_{i=1}^n X_i}$$

となる。

$\widehat{\lambda}$ の分布を求める。

$$\begin{aligned}\log f(X_i; \lambda) &= \log \lambda - \lambda X_i \\ \frac{d \log f(X_i; \lambda)}{d\lambda} &= \frac{1}{\lambda} - X_i \\ \frac{d^2 \log f(X_i; \lambda)}{d\lambda^2} &= -\frac{1}{\lambda^2} \\ -\frac{1}{\sum_{i=1}^n E\left(\frac{d^2 \log f(X_i; \lambda)}{d\lambda^2}\right)} &= \frac{\lambda^2}{n}\end{aligned}$$

したがって,

$$\widehat{\lambda} \sim N\left(\lambda, \frac{\lambda^2}{n}\right)$$

を得る。

12.2.1 変数変換

確率変数 X の密度関数を $f(x)$, 分布関数を $F(x) \equiv P(X < x)$ とする。 $Y = aX + b$ とするとき, Y の密度関数 $g(y)$ を求める。

Y の分布関数を $G(y)$ として、次のように変形できる。

$$\begin{aligned} G(y) &= P(Y < y) = P(aX + b < y) \\ &= \begin{cases} P\left(X < \frac{y-b}{a}\right), & a > 0 \text{ のとき} \\ P\left(X > \frac{y-b}{a}\right), & a < 0 \text{ のとき} \end{cases} \\ &= \begin{cases} P\left(X < \frac{y-b}{a}\right), & a > 0 \text{ のとき} \\ 1 - P\left(X < \frac{y-b}{a}\right), & a < 0 \text{ のとき} \end{cases} \\ &= \begin{cases} F\left(\frac{y-b}{a}\right), & a > 0 \text{ のとき} \\ 1 - F\left(\frac{y-b}{a}\right), & a < 0 \text{ のとき} \end{cases} \end{aligned}$$

分布関数と密度関数との関係は、

$$\frac{dF(x)}{dx} = f(x) \qquad \frac{dG(x)}{dx} = g(x)$$

であるので、 Y の密度関数は、

$$g(y) = \frac{dG(y)}{dy}$$

$$\begin{aligned}
&= \begin{cases} \frac{1}{a}f\left(\frac{y-b}{a}\right), & a > 0 \text{ のとき} \\ -\frac{1}{a}f\left(\frac{y-b}{a}\right), & a < 0 \text{ のとき} \end{cases} \\
&= \left|\frac{1}{a}\right|f\left(\frac{y-b}{a}\right)
\end{aligned}$$

と表される。

一般に、確率変数 X の密度関数を $f(x)$ とする。単調変換 $X = h(Y)$ とするとき、 Y の密度関数 $g(y)$ は、

$$g(y) = |h'(y)|f(h(y))$$

となる。

12.2.2 回帰分析への応用

回帰モデル

$$Y_i = \alpha + \beta X_i + u_i \quad i = 1, 2, \dots, n$$

u_1, u_2, \dots, u_n は互いに独立で、すべての i について $u_i \sim N(0, \sigma^2)$ を仮定する。

u_i の密度関数は,

$$f(u_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}u_i^2\right)$$

となる。

Y_i の密度関数 $g(Y_i)$ は,

$$g(Y_i) = |h'(Y_i)|f(h(Y_i))$$

によって求められる。

この場合, $h(Y_i) = Y_i - \alpha - \beta X_i$ なので, $h'(Y_i) = 1$ となる。

したがって, Y_i の密度関数は,

$$\begin{aligned} g(Y_i) &= |h'(Y_i)|f(h(Y_i)) \\ &= f(h(Y_i)) \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(Y_i - \alpha - \beta X_i)^2\right) \end{aligned}$$

となる。

u_1, u_2, \dots, u_n は互いに独立であれば, Y_1, Y_2, \dots, Y_n も互いに独立になるので, Y_1, Y_2, \dots, Y_n の結合密度関数は,

$$\begin{aligned} g(Y_1, Y_2, \dots, Y_n) &= \prod_{i=1}^n g(Y_i) \\ &= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2\right) \end{aligned}$$

となる。これは α, β, σ^2 の関数となっている。

よって, 尤度関数は,

$$l(\alpha, \beta, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2\right)$$

となる。

対数尤度関数は,

$$\begin{aligned} \log l(\alpha, \beta, \sigma^2) &= -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) \\ &\quad - \frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2 \end{aligned}$$

となる。

$\log l(\alpha, \beta, \sigma^2)$ を最大にするために,

$$\frac{\partial \log l(\alpha, \beta, \sigma^2)}{\partial \alpha} = \frac{1}{\sigma^2} \sum_{i=1}^n (Y_i - \alpha - \beta X_i) = 0$$

$$\frac{\partial \log l(\alpha, \beta, \sigma^2)}{\partial \beta} = \frac{1}{\sigma^2} \sum_{i=1}^n X_i (Y_i - \alpha - \beta X_i) = 0$$

$$\frac{\partial \log l(\alpha, \beta, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2} \frac{1}{\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2 = 0$$

の連立方程式を解く。

上2つの式は σ^2 に依存していない。 α , β の最尤推定量は最小二乗推定量と同じになる。

すなわち,

$$\widehat{\beta} = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

$$\widehat{\alpha} = \bar{Y} - \widehat{\beta} \bar{X}$$

σ^2 の最尤推定量は,

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - \widehat{\alpha} - \widehat{\beta}X_i)^2$$

となり, s^2 とは異なる。

$\widehat{\theta} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\sigma}^2)'$, $\theta = (\alpha, \beta, \sigma^2)'$ とする。 n が大きいとき,

$$\widehat{\theta} \sim N(\theta, \Sigma_\theta)$$

ただし,

$$\begin{aligned} \Sigma_\theta &= \left(\sum_{i=1}^n E \left[\left(\frac{\partial \log g(Y_i; \theta)}{\partial \theta} \right) \left(\frac{\partial \log g(Y_i; \theta)}{\partial \theta} \right)' \right] \right)^{-1} \\ &= - \left(\sum_{i=1}^n E \left[\frac{\partial^2 \log g(Y_i; \theta)}{\partial \theta \partial \theta'} \right] \right)^{-1} \end{aligned}$$

Y_i の密度関数 $g(Y_i; \theta)$ の対数は,

$$\begin{aligned} \log g(Y_i; \theta) &= -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2) \\ &\quad - \frac{1}{2\sigma^2} (Y_i - \alpha - \beta X_i)^2 \end{aligned}$$

となる。

$$\begin{aligned}\frac{\partial \log g(Y_i; \theta)}{\partial \theta} &= \begin{pmatrix} \frac{\partial \log g(Y_i; \theta)}{\partial \alpha} \\ \frac{\partial \log g(Y_i; \theta)}{\partial \beta} \\ \frac{\partial \log g(Y_i; \theta)}{\partial \sigma^2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{\sigma^2}(Y_i - \alpha - \beta X_i) \\ \frac{1}{\sigma^2}X_i(Y_i - \alpha - \beta X_i) \\ -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(Y_i - \alpha - \beta X_i)^2 \end{pmatrix}\end{aligned}$$

$$\frac{\partial^2 \log g(Y_i; \theta)}{\partial \theta \partial \theta'}$$

$$\begin{aligned}
&= \begin{pmatrix} \frac{\partial^2 \log g(Y_i; \theta)}{\partial \alpha^2} & \frac{\partial^2 \log g(Y_i; \theta)}{\partial \alpha \partial \beta} & \frac{\partial^2 \log g(Y_i; \theta)}{\partial \alpha \partial \sigma^2} \\ \frac{\partial^2 \log g(Y_i; \theta)}{\partial \beta \partial \alpha} & \frac{\partial^2 \log g(Y_i; \theta)}{\partial \beta^2} & \frac{\partial^2 \log g(Y_i; \theta)}{\partial \beta \partial \sigma^2} \\ \frac{\partial^2 \log g(Y_i; \theta)}{\partial \sigma^2 \partial \alpha} & \frac{\partial^2 \log g(Y_i; \theta)}{\partial \sigma^2 \partial \beta} & \frac{\partial^2 \log g(Y_i; \theta)}{\partial (\sigma^2)^2} \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{\sigma^2} & -\frac{X_i}{\sigma^2} & -\frac{u_i}{\sigma^4} \\ -\frac{X_i}{\sigma^2} & -\frac{X_i^2}{\sigma^2} & -\frac{X_i u_i}{\sigma^4} \\ -\frac{u_i}{\sigma^4} & -\frac{X_i u_i}{\sigma^4} & \frac{1}{2\sigma^4} - \frac{u_i^2}{\sigma^6} \end{pmatrix}
\end{aligned}$$

ただし, $u_i = Y_i - \alpha - \beta X_i$
期待値をとると,

$$\mathbb{E}\left(\frac{\partial^2 \log g(Y_i; \theta)}{\partial \theta \partial \theta'}\right) = \mathbb{E}\begin{pmatrix} \frac{1}{\sigma^2} & -\frac{X_i}{\sigma^2} & -\frac{u_i}{\sigma^4} \\ -\frac{X_i}{\sigma^2} & -\frac{X_i^2}{\sigma^2} & -\frac{X_i u_i}{\sigma^4} \\ -\frac{u_i}{\sigma^4} & -\frac{X_i u_i}{\sigma^4} & \frac{1}{2\sigma^4} - \frac{u_i^2}{\sigma^6} \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{X_i}{\sigma^2} & 0 \\ -\frac{X_i}{\sigma^2} & -\frac{X_i^2}{\sigma^2} & 0 \\ 0 & 0 & -\frac{1}{2\sigma^4} \end{pmatrix}$$

となる。

$$\begin{aligned} \Sigma_{\theta} &= -\left(\sum_{i=1}^n \mathbb{E} \left[\frac{\partial^2 \log g(Y_i; \theta)}{\partial \theta \partial \theta'} \right] \right)^{-1} \\ &= \begin{pmatrix} \frac{n}{\sigma^2} & \frac{\sum_{i=1}^n X_i}{\sigma^2} & 0 \\ \frac{\sum_{i=1}^n X_i}{\sigma^2} & \frac{\sum_{i=1}^n X_i^2}{\sigma^2} & 0 \\ 0 & 0 & \frac{n}{2\sigma^4} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \sigma^2 \left(\begin{matrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{matrix} \right)^{-1} & 0 \\ 0 & \frac{2\sigma^4}{n} \end{pmatrix} \end{aligned}$$

したがって,

$$\begin{pmatrix} \widehat{\alpha} \\ \widehat{\beta} \end{pmatrix} \sim N\left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \sigma^2 \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix}^{-1}\right)$$

となる。

→ 最小二乗推定量の分布と同じ。

12.2.3 誤差項に系列相関がある場合

回帰モデル

$$Y_i = \alpha + \beta X_i + u_i$$

$$u_i = \rho u_{i-1} + \epsilon_i \quad i = 2, 3, \dots, n$$

$\epsilon_2, \epsilon_3, \dots, \epsilon_n$ は互いに独立で, すべての i について $\epsilon_i \sim N(0, \sigma^2)$ を仮定する。

u_i を消去すると,

$$(Y_i - \alpha - \beta X_i) = \rho(Y_{i-1} - \alpha - \beta X_{i-1}) + \epsilon_i$$

または

$$(Y_i - \rho Y_{i-1}) = \alpha(1 - \rho) + \beta(X_i - \rho X_{i-1}) + \epsilon_i$$

と書き直すことが出来る。

$\theta = (\alpha, \beta, \sigma^2, \rho)$ とする。

$$\begin{aligned}\log f(Y_i; \theta) &= -\frac{1}{2} \log(2\pi\sigma^2) \\ &\quad - \frac{1}{2\sigma^2} \left((Y_i - \rho Y_{i-1}) - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1}) \right)^2\end{aligned}$$

尤度関数は,

$$\begin{aligned}\log l(\theta) &= \sum_{i=2}^n \log f(Y_i; \theta) = -\frac{n-1}{2} \log(2\pi) - \frac{n-1}{2} \log(\sigma^2) \\ &\quad - \frac{1}{2\sigma^2} \sum_{i=2}^n \left((Y_i - \rho Y_{i-1}) - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1}) \right)^2\end{aligned}$$

となる。

尤度関数をそれぞれ α , β , σ^2 , ρ について微分し, ゼロとおく。

$$\begin{aligned}\frac{\partial \log l(\theta)}{\partial \alpha} &= \frac{1-\rho}{\sigma^2} \sum_{i=2}^n \left((Y_i - \rho Y_{i-1}) \right. \\ &\quad \left. - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1}) \right) = 0\end{aligned}$$

$$\frac{\partial \log l(\theta)}{\partial \beta} = \frac{1}{\sigma^2} \sum_{i=2}^n (X_i - \rho X_{i-1}) \left((Y_i - \rho Y_{i-1}) - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1}) \right) = 0$$

$$\frac{\partial \log l(\theta)}{\partial \sigma^2} = -\frac{n-1}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=2}^n \left((Y_i - \rho Y_{i-1}) - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1}) \right)^2 = 0$$

$$\frac{\partial \log l(\theta)}{\partial \rho} = \frac{1}{\sigma^2} \sum_{i=2}^n (Y_{i-1} - \alpha - \beta X_{i-1}) \left((Y_i - \alpha - \beta X_i) - \rho(Y_{i-1} - \alpha - \beta X_{i-1}) \right) = 0$$

$(Y_i - \alpha - \beta X_i) - \rho(Y_{i-1} - \alpha - \beta X_{i-1})$ は
 $(Y_i - \rho Y_{i-1}) - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1})$ を書き直したものの。

4つの連立方程式を解いて、最尤推定量 $\hat{\alpha}$, $\hat{\beta}$, $\hat{\sigma}^2$, $\hat{\rho}$ が得られる。
 → 下記のように収束計算によって求める。

(i) 初期段階では、 $\widehat{\rho} = 0$ とする。

$$(ii) X_i^* = X_i - \widehat{\rho}X_{i-1}$$

$$Y_i^* = Y_i - \widehat{\rho}Y_{i-1}$$

$$(iii) \begin{pmatrix} \widetilde{\alpha} \\ \widetilde{\beta} \end{pmatrix} = \begin{pmatrix} n-1 & \sum_{i=2}^n X_i^* \\ \sum_{i=2}^n X_i^* & \sum_{i=2}^n X_i^{*2} \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=2}^n Y_i^* \\ \sum_{i=2}^n X_i^* Y_i^* \end{pmatrix}$$

$$(iv) \widehat{\alpha} = \frac{\widetilde{\alpha}}{1 - \widehat{\rho}}$$

$$(v) \widehat{u}_i = Y_i - \widehat{\alpha} - \widehat{\beta}X_i$$

$$(vi) \widehat{\sigma}^2 = \frac{1}{n-1} \sum_{i=2}^n (\widehat{u}_i - \widehat{\rho}\widehat{u}_{i-1})^2$$

$$(vii) \widehat{\rho} = \frac{\sum_{i=2}^n \widehat{u}_i \widehat{u}_{i-1}}{\sum_{i=2}^n \widehat{u}_{i-1}^2}$$

(viii) ステップ (ii) ~ (vii) を，収束するまで繰り返し計算する。