

第10章 推定量の求め方

10.1 最小二乗法

- ・ n 個のデータ (実現値) : x_1, x_2, \dots, x_n
 - ・ 背後に対応する確率変数を仮定 : X_1, X_2, \dots, X_n
 - ・ $E(X_i) = \mu, V(X_i) = \sigma^2$ を仮定
- 母数 (μ, σ^2) を推定する。

観測データ x_1, x_2, \dots, x_n をもとにして, μ の最小二乗推定値を求める。

$$\min_{\mu} \sum_{i=1}^n (x_i - \mu)^2$$

μ の解を $\hat{\mu}$ とすると,

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

となり, $\hat{\mu} \equiv \bar{x}$ を得る。

すなわち,

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

を μ について解く。

μ の最小二乗推定量はデータ x_i を対応する確率変数 X_i で置き換えて,

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

となり、 $\hat{\mu} \equiv \bar{X}$ を得る ($\hat{\mu}$ について、推定値と推定量は同じ記号を使っている)。

以上を回帰分析に応用すると、

$$\min_{\alpha, \beta} \sum_{i=1}^n (Y_i - \alpha - \beta X_i)^2$$

を解くことになる。

すなわち、

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

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

の連立方程式を α, β について解く。

10.2 最尤法

n 個の確率変数 X_1, X_2, \dots, X_n は互いに独立で、同じ確率分布 $f(x) \equiv f(x; \theta)$ とする。ただし、 θ は母数で、例えば、 $\theta = (\mu, \sigma^2)$ である。

X_1, X_2, \dots, X_n の結合分布は、互いに独立なので、

$$f(x_1, x_2, \dots, x_n; \theta) \equiv \prod_{i=1}^n f(x_i; \theta)$$

と表される。

観測データ x_1, x_2, \dots, x_n を与えたもとの、 $\prod_{i=1}^n f(x_i; \theta)$ は θ の関数として表される。すなわち、

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

となる。

$l(\theta)$ を尤度関数と呼ぶ。

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

となる θ を最尤推定値 $\hat{\theta} = \hat{\theta}(x_1, x_2, \dots, x_n)$ と呼ぶ。

データ x_1, x_2, \dots, x_n を確率変数 X_1, X_2, \dots, X_n で置き換えて, $\hat{\theta} = \hat{\theta}(X_1, X_2, \dots, X_n)$ を最尤推定量と呼ぶ。

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

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,

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

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,

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

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 $\hat{\mu}$ とすると、 μ の最尤推定値は、

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

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、

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

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

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

の連立方程式を解く。

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

$$\begin{aligned} \hat{\mu} &= \frac{1}{n} \sum_{i=1}^n x_i \equiv \bar{x} \\ \hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \hat{\mu})^2 \equiv \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \end{aligned}$$

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,

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

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

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,

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

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

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

$$\begin{aligned}
& \frac{\partial^2 \log f(X_i; \theta)}{\partial \theta \partial \theta'} \\
&= \begin{pmatrix} \frac{\partial^2 \log f(X_i; \theta)}{\partial \mu^2} & \frac{\partial^2 \log f(X_i; \theta)}{\partial \mu \partial \sigma^2} \\ \frac{\partial^2 \log f(X_i; \theta)}{\partial \sigma^2 \partial \mu} & \frac{\partial^2 \log f(X_i; \theta)}{\partial (\sigma^2)^2} \end{pmatrix} \\
&= \begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{1}{\sigma^4}(X_i - \mu) \\ -\frac{1}{\sigma^4}(X_i - \mu) & \frac{1}{2\sigma^4} - \frac{1}{\sigma^6}(X_i - \mu)^2 \end{pmatrix} \\
& \mathbb{E} \left[\frac{\partial^2 \log f(X_i; \theta)}{\partial \theta \partial \theta'} \right] \\
&= \begin{pmatrix} -\frac{1}{\sigma^2} & -\frac{1}{\sigma^4} \mathbb{E}(X_i - \mu) \\ -\frac{1}{\sigma^4} \mathbb{E}(X_i - \mu) & \frac{1}{2\sigma^4} - \frac{1}{\sigma^6} \mathbb{E}[(X_i - \mu)^2] \end{pmatrix} \\
&= \begin{pmatrix} -\frac{1}{\sigma^2} & 0 \\ 0 & -\frac{1}{2\sigma^4} \end{pmatrix}
\end{aligned}$$

よって,

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

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,

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

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。

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

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

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

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

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

したがって、

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

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 $\hat{\lambda}$ は,

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

となる。

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

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

となる。

$\hat{\lambda}$ の分布を求める。

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

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

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

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

$$\begin{aligned} E(X_i) &= \sum_{x=0}^{\infty} x f(x; \lambda) \\ &= \sum_{x=0}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!} \\ &= \sum_{x=1}^{\infty} x \frac{\lambda^x e^{-\lambda}}{x!} \\ &= \sum_{x=1}^{\infty} \lambda \frac{\lambda^{x-1} e^{-\lambda}}{(x-1)!} \\ &= \sum_{x=0}^{\infty} \lambda \frac{\lambda^x e^{-\lambda}}{x!} \\ &= \lambda \\ &= \frac{1}{\sum_{i=1}^n E\left(\frac{d^2 \log f(X_i; \lambda)}{d\lambda^2}\right)} = \frac{\lambda}{n} \end{aligned}$$

したがって、

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

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,

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

となる。

$\log l(\lambda)$ を最大にする p を求める。

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

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

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

となる。

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

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

となる。

$\hat{\lambda}$ の分布を求める。

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

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

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

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

したがって,

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

を得る。

10.2.1 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。

10.2.2 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、

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

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

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

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,

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

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

σ^2 の最尤推定量は,

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

となり, s^2 とは異なる。

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

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

ただし,

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

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

Y_i の密度関数 $g(Y_i; \theta)$ の対数は,

$$\log g(Y_i; \theta) = -\frac{1}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2)$$

$$- \frac{1}{2\sigma^2} (Y_i - \alpha - \beta X_i)^2$$

となる。

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

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

ただし, $u_i = Y_i - \alpha - \beta X_i$

期待値をとると,

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

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

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、

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

となる。

→ 最小二乗推定量の分布と同じ。

10.2.3 自己回帰モデルの最尤推定法

p 次の自己回帰モデル, すなわち, AR(p) モデル (AutoRegressive model) :

$$Y_i = \phi_1 Y_{i-1} + \phi_2 Y_{i-2} + \cdots + \phi_p Y_{i-p} + u_i$$

AR(1) モデル:

$$Y_i = \phi_1 Y_{i-1} + u_i, \quad u_i \sim N(0, \sigma^2)$$

where $|\phi_1| < 1$ が仮定される。

Y_1, Y_2, \dots, Y_n の結合密度関数 $f(y_n, y_{n-1}, \dots, y_1)$ は次のように分解される。

$$f(y_n, y_{n-1}, \dots, y_1) = f(y_1) \prod_{i=2}^n f(y_i | y_{i-1}, \dots, y_1).$$

条件付き密度関数 $f(y_i | y_{i-1}, \dots, y_1)$ と条件なしの密度関数 $f(y_1)$ の積からなる。

$Y_i = \phi_1 Y_{i-1} + u_i$ から、条件付き Y_i の平均と分散は、

$E(Y_i | Y_{i-1}, \dots, Y_1) = \phi_1 Y_{i-1}$, and $V(Y_i | Y_{i-1}, \dots, Y_1) = \sigma^2$
として得られる。

条件付き分布 $f(y_i | y_{i-1}, \dots, y_1)$ は、

$$f(y_i | y_{i-1}, \dots, y_1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y_i - \phi_1 y_{i-1})^2\right).$$

Y_i の条件なしの密度関数 $f(y_i)$ を求めるために、 Y_i を次のように書き換える。

$$\begin{aligned} y_i &= \phi_1 y_{i-1} + u_i \\ &= \phi_1^2 y_{i-2} + u_i + \phi_1 u_{i-1} \\ &\vdots \\ &= \phi_1^j y_{i-j} + u_i + \phi_1 u_{i-1} + \dots + \phi_1^j u_{i-j} \\ &\vdots \\ &= u_i + \phi_1 u_{i-1} + \phi_1^2 u_{i-2} + \dots, \quad j \text{ が無限大になるとき,} \end{aligned}$$

The unconditional expectation and variance of Y_i の条件なしの平均と分散は、

$$E(Y_i) = 0, \quad V(Y_i) = \sigma^2(1 + \phi_1^2 + \phi_1^4 + \dots) = \frac{\sigma^2}{1 - \phi_1^2}.$$

Therefore, the unconditional distribution of Y_i の条件なし分布は、

$$f(y_i) = \frac{1}{\sqrt{2\pi\sigma^2/(1 - \phi_1^2)}} \exp\left(-\frac{1}{2\sigma^2/(1 - \phi_1^2)} y_i^2\right).$$

となる。

よって, Y_1, Y_2, \dots, Y_n の結合密度関数は,

$$\begin{aligned} f(y_n, y_{n-1}, \dots, y_1) &= f(y_1) \prod_{i=2}^n f(y_i | y_{i-1}, \dots, y_1) \\ &= \frac{1}{\sqrt{2\pi\sigma^2/(1-\phi_1^2)}} \exp\left(-\frac{1}{2\sigma^2/(1-\phi_1^2)} y_1^2\right) \\ &\quad \times \prod_{i=2}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (y_i - \phi_1 y_{i-1})^2\right) \end{aligned}$$

したがって, 尤度関数は,

$$\begin{aligned} \log L(\phi_1, \sigma^2; y_i, y_{i-1}, \dots, y_1) &= -\frac{1}{2} \log(2\pi\sigma^2/(1-\phi_1^2)) - \frac{1}{2\sigma^2/(1-\phi_1^2)} y_1^2 \\ &\quad - \frac{n-1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=2}^n (y_i - \phi_1 y_{i-1})^2. \end{aligned}$$

ϕ_1 と σ^2 について, 尤度関数 $\log L$ を最大化する。

最大化問題:

- ニュートン・ラブソン法 (Newton-Raphson Method) やスコアリング法 (Method of Scoring)
- 単純探索法 (Simple Grid Search), すなわち, $-1 < \phi_1 < 1$ の範囲内で様々な ϕ_1 の値について尤度関数を計算し最大になるものを見つける。

10.2.4 誤差項に系列相関がある場合

回帰モデル

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

$$u_i = \rho u_{i-1} + \epsilon_i$$

$\epsilon_1, \epsilon_2, \dots, \epsilon_n$ 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。条件付き密度関数 $f(Y_i|Y_{i-1}, Y_{i-2}, \dots, Y_1; \theta)$ は,

$$\begin{aligned} \log f(Y_i|Y_{i-1}, Y_{i-2}, \dots, Y_1; \theta) \\ = -\frac{1}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \left((Y_i - \rho Y_{i-1}) - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1}) \right)^2 \end{aligned}$$

となる。なぜなら、 $E(Y_i|Y_{i-1}, Y_{i-2}, \dots, Y_1) = \rho Y_{i-1} + \alpha(1 - \rho) + \beta(X_i - \rho X_{i-1})$,
 $V(Y_i|Y_{i-1}, Y_{i-2}, \dots, Y_1) = \sigma^2$ となる。

条件なし密度関数 $f(Y_i; \theta)$ は,

$$\log f(Y_i; \theta) = -\frac{1}{2} \log(2\pi\sigma^2/(1 - \rho^2)) - \frac{1}{2\sigma^2/(1 - \rho^2)} (Y_i - \alpha - \beta X_i)^2$$

となる。なぜなら、 $E(Y_i) = \alpha + \beta X_i$, $V(Y_i) = \sigma^2/(1 - \rho^2)$ となる。 $E(u_i) = 0$,
 $V(u_i) = \sigma^2/(1 - \rho^2)$ に注意。

よって、尤度関数は,

$$\begin{aligned} \log l(\theta) &= \log f(Y_1; \theta) + \sum_{i=2}^n \log f(Y_i|Y_{i-1}, Y_{i-2}, \dots, Y_1; \theta) \\ &= -\frac{1}{2} \log(2\pi\sigma^2/(1 - \rho^2)) - \frac{1}{2\sigma^2/(1 - \rho^2)} (Y_1 - \alpha - \beta X_1)^2 \\ &\quad - \frac{n-1}{2} \log(2\pi) - \frac{n-1}{2} \log(\sigma^2) \\ &\quad - \frac{1}{2\sigma^2} \sum_{i=2}^n \left((Y_i - \rho Y_{i-1}) - \alpha(1 - \rho) - \beta(X_i - \rho X_{i-1}) \right)^2 \\ &= -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) + \frac{1}{2} \log(1 - \rho^2) \\ &\quad - \frac{1}{2\sigma^2/(1 - \rho^2)} (Y_i - \alpha - \beta X_i)^2 \end{aligned}$$

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

となる。ただし,

$$\begin{aligned}
Y_i^* &= Y_i \sqrt{1 - \rho^2}, & z_i &= \sqrt{1 - \rho^2}, & X_i^* &= X_i \sqrt{1 - \rho^2}, & i &= 1 \text{ のとき,} \\
Y_i^* &= Y_i - \rho Y_{i-1}, & z_i &= 1 - \rho, & X_i^* &= X_i - \rho X_{i-1}, & i &\neq 1 \text{ のとき,}
\end{aligned}$$

とする。

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。

$$\frac{\partial \log l(\theta)}{\partial \alpha} = \frac{1}{\sigma^2} \sum_{i=1}^n z_i (Y_i^* - \alpha z_i - \beta X_i^*) = 0 \quad (10.1)$$

$$\frac{\partial \log l(\theta)}{\partial \beta} = \frac{1}{\sigma^2} \sum_{i=1}^n X_i^* (Y_i^* - \alpha z_i - \beta X_i^*) = 0 \quad (10.2)$$

$$\frac{\partial \log l(\theta)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (Y_i^* - \alpha z_i - \beta X_i^*)^2 = 0 \quad (10.3)$$

$$\frac{\partial \log l(\theta)}{\partial \rho} = -\frac{\rho}{1 - \rho^2} - \frac{1}{2\sigma^2} \sum_{i=1}^n \frac{\partial (Y_i^* - \alpha z_i - \beta X_i^*)^2}{\partial \rho} = 0 \quad (10.4)$$

ただし,

$$\frac{\partial(Y_i^* - \alpha z_i - \beta X_i^*)^2}{\partial \rho} = \begin{cases} (-Y_i + \alpha + \beta X_i) \frac{\rho}{\sqrt{1-\rho^2}}, & i = 1 \text{ のとき} \\ -Y_{i-1} + \alpha + \beta X_{i-1}, & i = 2, 3, \dots, n \text{ のとき} \end{cases}$$

とする。

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、 $\rho = 0$ とする。

(ii) (10.1) と (10.2) から、 $\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n z_i^2 & \sum_{i=1}^n z_i X_i^* \\ \sum_{i=1}^n z_i X_i^* & \sum_{i=1}^n X_i^{*2} \end{pmatrix}^{-1} \begin{pmatrix} \sum_{i=1}^n z_i Y_i^* \\ \sum_{i=1}^n X_i^* Y_i^* \end{pmatrix}$ を計算すれば、 α と β が求められる。

(iii) ステップ (ii) で求めた α と β と (10.3) から、 $\sigma^2 = \frac{1}{n} \sum_{i=1}^n (Y_i^* - \alpha z_i - \beta X_i^*)^2$ によって、 σ^2 が求まる。

(iv) ステップ (ii), (iii) で求めた α , β , σ^2 を代入して、(10.4) を満たす ρ を求める。すなわち、非線形方程式を解く。

(v) ステップ (ii) ~ (iv) を、収束するまで繰り返し計算する。収束先が最尤推定量 $\hat{\alpha}$, $\hat{\beta}$, $\hat{\sigma}^2$, $\hat{\rho}$ となる。

10.3 尤度比検定

n 個の確率変数 X_1, X_2, \dots, X_n は互いに独立で、同じ確率分布 $f(x) \equiv f(x; \theta)$ とする。

尤度関数は、

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

となる。

θ の制約つき最尤推定量を $\tilde{\theta}$, 制約無し最尤推定量を $\hat{\theta}$ とする。
制約の数を G 個とする。

$\frac{l(\tilde{\theta})}{l(\hat{\theta})}$ を尤度比と呼ぶ

検定方法 1: 尤度比がある値より小さいときに, 帰無仮説を棄却する。すなわち,

$$\frac{l(\tilde{\theta})}{l(\hat{\theta})} < c$$

となるときに, 帰無仮説を棄却する。この場合, c を次のようにして求める必要がある。

$$\int \cdots \int \prod_{i=1}^n f(x_i; \tilde{\theta}) dx_1 \cdots dx_n = \alpha$$

ただし, α は有意水準 (帰無仮説が正しいときに, 帰無仮説を棄却する確率) を表す。

検定方法 2 (大標本検定): または, $n \rightarrow \infty$ のとき,

$$-2 \log \frac{l(\tilde{\theta})}{l(\hat{\theta})} \rightarrow \chi^2(G)$$

となる。

この検定を尤度比検定と呼ぶ。

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, σ^2 が既知のとき, 帰無仮説 $H_0: \mu = \mu_0$, $H_1: \mu \neq \mu_0$ の尤度比検定を行う。

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。

μ の最尤推定量は、

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

となる。

尤度比検定統計量は、

$$\begin{aligned} \frac{l(\mu_0)}{l(\bar{X})} &= \frac{\exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - \mu_0)^2\right)}{\exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (X_i - \bar{X})^2\right)} \\ &= \exp\left(-\frac{1}{2\sigma^2/n} (\bar{X} - \mu_0)^2\right) < c \end{aligned}$$

となる c を求める。

H_0 が正しいときに、 $\sqrt{n}(\bar{X} - \mu_0)/\sigma \sim N(0, 1)$ となるので、

$$P\left(\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| > z_{\alpha/2}\right) = \alpha$$

すなわち、

$$P\left(\exp\left(-\frac{1}{2\sigma^2/n} (\bar{X} - \mu_0)^2\right) < \exp\left(-\frac{1}{2} z_{\alpha/2}^2\right)\right) = \alpha$$

と変形できる。したがって、

$$c = \exp\left(-\frac{1}{2} z_{\alpha/2}^2\right)$$

とすればよい。

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、

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

となる。

p の最尤推定量 \hat{p} は、

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

である。

次の仮説検定を考える。

$$H_0: p = p_0 \quad H_1: p \neq p_0$$

→ 制約数は 1 つ。 ($G = 1$)

尤度比は、

$$\frac{l(p_0)}{l(\hat{p})} = \frac{\prod_{i=1}^n p_0^{X_i}(1-p_0)^{1-X_i}}{\prod_{i=1}^n \hat{p}^{X_i}(1-\hat{p})^{1-X_i}}$$

したがって、 $n \rightarrow \infty$ のとき、

$$-2 \log \frac{l(p_0)}{l(\hat{p})} = -2 \log \frac{p_0}{\hat{p}} \sum_{i=1}^n X_i - 2 \log \frac{1-p_0}{1-\hat{p}} \sum_{i=1}^n (1-X_i)$$

$$\rightarrow \chi^2(1)$$

$\chi^2(1)$ 分布の上側 100 $\alpha\%$ 点を $\chi^2_{\alpha}(1)$ とするとき、

$$-2 \log \frac{l(p_0)}{l(\hat{p})} > \chi^2_{\alpha}(1)$$

のとき、帰無仮説 $H_0: p = p_0$ を棄却する。

例3： 回帰モデル

$$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + u_i$$

$$u_i \sim N(0, \sigma^2) \quad i = 1, 2, \dots, n$$

について、 β_1, \dots, β_k に関する仮説の尤度比検定を行う。

例えば、

$$H_0 : \beta_1 = 0$$

$$H_0 : \beta_1 + \beta_2 = 1$$

$$H_0 : \beta_1 = \beta_2 = \beta_3 = 0$$

などのような仮説検定

$\theta = (\beta_1, \dots, \beta_k, \sigma^2)$ とする。

尤度関数は、

$$l(\theta) = \prod_{i=1}^n f(Y_i; \theta)$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i - \beta_1 X_{1i} - \cdots - \beta_k X_{ki})^2\right)$$

となる。

H_0 の制約つき最尤推定量を $\tilde{\theta} = (\tilde{\beta}_1, \dots, \tilde{\beta}_k, \tilde{\sigma}^2)$ とする。この仮説に含まれる制約数を G とする。

制約なし最尤推定量を $\hat{\theta} = (\hat{\beta}_1, \dots, \hat{\beta}_k, \hat{\sigma}^2)$ とする。

尤度比

$$\frac{l(\tilde{\theta})}{l(\hat{\theta})} = \frac{(2\pi\tilde{\sigma}^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\tilde{\sigma}^2} \sum_{i=1}^n (Y_i - \tilde{\beta}_1 X_{1i} - \cdots - \tilde{\beta}_k X_{ki})^2\right)}{(2\pi\hat{\sigma}^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\hat{\sigma}^2} \sum_{i=1}^n (Y_i - \hat{\beta}_1 X_{1i} - \cdots - \hat{\beta}_k X_{ki})^2\right)}$$

$$= \frac{(\tilde{\sigma}^2)^{-\frac{n}{2}} \exp\left(-\frac{n-G}{2}\right)}{(\hat{\sigma}^2)^{-\frac{n}{2}} \exp\left(-\frac{n-k}{2}\right)}$$

$$\begin{aligned}
&= \left(\frac{1}{n-G} \sum_{i=1}^n \tilde{u}_i^2 \right)^{-n/2} \exp\left(-\frac{k-G}{2}\right) \\
&= \exp\left(-\frac{k-G}{2}\right) \left(\frac{n-k}{n-G}\right)^{-n/2} \left(\frac{\sum_{i=1}^n \tilde{u}_i^2}{\sum_{i=1}^n \hat{u}_i^2}\right)^{-n/2} \\
&= \exp\left(-\frac{k-G}{2}\right) \left(\frac{n-k}{n-G}\right)^{-n/2} \\
&\quad \times \left(1 + \frac{\sum_{i=1}^n \tilde{u}_i^2 - \sum_{i=1}^n \hat{u}_i^2}{\sum_{i=1}^n \hat{u}_i^2}\right)^{-n/2} \\
&= \exp\left(-\frac{k-G}{2}\right) \left(\frac{n-k}{n-G}\right)^{-n/2} \\
&\quad \times \left(1 + \frac{G}{n-k} \frac{(\sum_{i=1}^n \tilde{u}_i^2 - \sum_{i=1}^n \hat{u}_i^2)/G}{\sum_{i=1}^n \hat{u}_i^2/(n-k)}\right)^{-n/2} \\
&< c
\end{aligned}$$

のとき仮説を棄却する。

$$\frac{(\sum_{i=1}^n \tilde{u}_i^2 - \sum_{i=1}^n \hat{u}_i^2)/G}{\sum_{i=1}^n \hat{u}_i^2/(n-k)} \sim F(G, n-k)$$

を利用すると c が求まる。

ただし、途中で以下を利用

$$\begin{aligned}
\tilde{\sigma}^2 &= \frac{1}{n-G} \sum_{i=1}^n (Y_i - \tilde{\beta}_1 X_{1i} - \cdots - \tilde{\beta}_k X_{ki})^2 \\
&= \frac{1}{n-G} \sum_{i=1}^n \tilde{u}_i^2 \\
\hat{\sigma}^2 &= \frac{1}{n-k} \sum_{i=1}^n (Y_i - \hat{\beta}_1 X_{1i} - \cdots - \hat{\beta}_k X_{ki})^2 \\
&= \frac{1}{n-k} \sum_{i=1}^n \hat{u}_i^2
\end{aligned}$$

近似的には,

$$\begin{aligned} -2 \log \frac{l(\tilde{\theta})}{l(\hat{\theta})} &= -2 \log \frac{(\tilde{\sigma}^2)^{-\frac{n}{2}} \exp\left(-\frac{n-G}{2}\right)}{(\hat{\sigma}^2)^{-\frac{n}{2}} \exp\left(-\frac{n-k}{2}\right)} \\ &= n \log\left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2}\right) + (k-G) \\ &\rightarrow \chi^2(G) \end{aligned}$$

例 4 : 回帰モデル

$$\begin{aligned} Y_i &= \alpha + \beta X_i + u_i \\ u_i &= \rho u_{i-1} + \epsilon_i \\ \epsilon_i &\sim N(0, \sigma^2) \quad i = 2, 3, \dots, n \end{aligned}$$

について, $H_0 : \rho = 0$, $H_1 : \rho \neq 0$ の尤度比検定を行う。

$\theta = (\alpha, \beta, \sigma^2, \rho)$ とする。対数尤度関数は,

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

となる。

対数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。4本の連立方程式を解いて, 制約なし最尤推定量 $\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2, \hat{\rho})$ が得られる。

$\rho = 0$ と制約をおく。 $\theta = (\alpha, \beta, \sigma^2, 0)$ とする。対数尤度関数は,

$$\begin{aligned} \log l(\theta) &= \log f(Y_1) + \sum_{i=2}^n \log f(Y_i | Y_{i-1}, Y_{i-2}, \dots, Y_1; \theta) \\ &= -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) + \frac{1}{2} \log(1-\rho^2) \end{aligned}$$

$$\begin{aligned}
& -\frac{1}{2\sigma^2/(1-\rho^2)}(Y_i - \alpha - \beta X_i)^2 \\
& -\frac{1}{2\sigma^2} \sum_{i=2}^n \left((Y_i - \rho Y_{i-1}) - \alpha(1-\rho) - \beta(X_i - \rho X_{i-1}) \right)^2 \\
& = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) + \frac{1}{2} \log(1-\rho^2) \\
& -\frac{1}{2\sigma^2} \sum_{i=1}^n (Y_i^* - \alpha z_i - \beta X_i^*)^2
\end{aligned}$$

となる。ただし、

$$\begin{aligned}
Y_i^* &= Y_i \sqrt{1-\rho^2}, & z_i &= \sqrt{1-\rho^2}, & X_i^* &= X_i \sqrt{1-\rho^2}, & i &= 1 \text{ のとき}, \\
Y_i^* &= Y_i - \rho Y_{i-1}, & z_i &= 1 - \rho, & X_i^* &= X_i - \rho X_{i-1}, & i &\neq 1 \text{ のとき},
\end{aligned}$$

とする。

上記の対数尤度関数をそれぞれ α , β , σ^2 について微分し、ゼロとおく。3本の連立方程式を解いて、 $\rho = 0$ の制約付き最尤推定量 $\tilde{\theta} = (\tilde{\alpha}, \tilde{\beta}, \tilde{\sigma}^2, 0)$ が得られる。

すなわち、

$$\frac{\max_{\alpha, \beta, \sigma^2} l(\alpha, \beta, \sigma^2, 0)}{\max_{\alpha, \beta, \sigma^2, \rho} l(\alpha, \beta, \sigma^2, \rho)} = \frac{l(\tilde{\alpha}, \tilde{\beta}, \tilde{\sigma}^2, 0)}{l(\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2, \hat{\rho})} = \frac{l(\tilde{\theta})}{l(\hat{\theta})}$$

したがって、尤度比検定統計量

$$-2 \log \frac{l(\tilde{\theta})}{l(\hat{\theta})}$$

は、 n が大きくなると、 $\chi^2(1)$ 分布に近づく。