$$= -\frac{1}{\sum_{i=1}^{n} \mathrm{E}\Big[\frac{\mathrm{d}^2 \log f(X_i;\theta)}{\mathrm{d}\theta^2}\Big]}$$

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，

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

例 1：　正規母集団 $\mathrm{N}(\mu, \sigma^2)$ からの標本値 $x_1, x_2, \cdots, x_n$ 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] $\mathrm{N}(\mu, \sigma^2)$ 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 $X_1, X_2, \cdots, X_n$ の結合分布は,

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

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,

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

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,

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

を得る。

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

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

となる。

$\hat{\mu}$ の分散を求めるために,

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

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

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

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

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，

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

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，

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

と表される。

μ と σ^2 について，最大化するためには,

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

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

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,

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

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

となる。

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

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

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

となる。

σ^2 の最尤推定量 $\hat{\sigma}^2$ は，σ^2 の不偏推定量 $S^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(X_i - \overline{X})^2$ 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,

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

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

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

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

$$\frac{\partial^2 \log f(X_i;\theta)}{\partial\theta\partial\theta'}$$

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

$$
= \begin{pmatrix} -\dfrac{1}{\sigma^2} & -\dfrac{1}{\sigma^4}(X_i-\mu) \\ -\dfrac{1}{\sigma^4}(X_i-\mu) & \dfrac{1}{2\sigma^4}-\dfrac{1}{\sigma^6}(X_i-\mu)^2 \end{pmatrix}
$$

$$
\mathrm{E}\Big[\frac{\partial^2 \log f(X_i;\theta)}{\partial\theta\partial\theta'}\Big]
$$

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

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

よって，

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

$$
= \begin{pmatrix} \dfrac{\sigma^2}{n} & 0 \\ 0 & \dfrac{2\sigma^4}{n} \end{pmatrix}
$$

まとめると，μ，σ^2 の最尤推定量 $\hat{\mu} = (1/n)\sum_{i=1}^{n} X_i$，$\hat{\sigma}^2 = (1/n)\sum_{i=1}^{n}(X_i - \overline{X})^2$ の分布は，n が大きいとき，

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

となる。

例 2： $X_1, X_n, \cdots, X_n$ 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,

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

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。

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

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。

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

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

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

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

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

したがって，

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

を得る。

例 3： $X_1, X_n, \cdots, X_n$ 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,

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

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。

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

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。

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

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

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

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

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

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

　したがって，

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

を得る。

例 4：　$X_1, X_n, \cdots, X_n$ 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，

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

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。

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

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。

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

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

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

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

したがって，

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

を得る。

変数変換

確率変数 X の密度関数を $f(x)$，分布関数を $F(x) \equiv \mathrm{P}(X < x)$ 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。

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

分布関数と密度関数との関係は,

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = f(x) \qquad \frac{\mathrm{d}G(x)}{\mathrm{d}x} = g(x)$$

であるので，Y の密度関数は,

$$\begin{aligned} g(y) &= \frac{\mathrm{d}G(y)}{\mathrm{d}y} \\ &= \begin{cases} \dfrac{1}{a} f\Bigl(\dfrac{y-b}{a}\Bigr), & a > 0 \text{ のとき} \\ -\dfrac{1}{a} f\Bigl(\dfrac{y-b}{a}\Bigr), & a < 0 \text{ のとき} \end{cases} \\ &= \Bigl|\frac{1}{a}\Bigr| f\Bigl(\frac{y-b}{a}\Bigr) \end{aligned}$$

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。

回帰分析への応用

回帰モデル

$$Y_i = \alpha + \beta X_i + u_i \qquad i = 1, 2, \cdots, n$$

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

u_i の密度関数は,

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

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,

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

となる。

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

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

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,

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

となる。

対数尤度関数は,

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

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，

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

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

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,

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

Y_i の密度関数 $g(Y_i;\theta)$ の対数は,

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

となる。

$$\frac{\partial \log g(Y_i;\theta)}{\partial \theta} = \begin{pmatrix} \dfrac{\partial \log g(Y_i;\theta)}{\partial \alpha} \\ \dfrac{\partial \log g(Y_i;\theta)}{\partial \beta} \\ \dfrac{\partial \log g(Y_i;\theta)}{\partial \sigma^2} \end{pmatrix}$$

$$= \begin{pmatrix} \dfrac{1}{\sigma^2}(Y_i - \alpha - \beta X_i) \\ \dfrac{1}{\sigma^2}X_i(Y_i - \alpha - \beta X_i) \\ -\dfrac{1}{2\sigma^2} + \dfrac{1}{2\sigma^4}(Y_i - \alpha - \beta X_i)^2 \end{pmatrix}$$

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

$$= \begin{pmatrix} \dfrac{\partial^2 \log g(Y_i;\theta)}{\partial \alpha^2} & \dfrac{\partial^2 \log g(Y_i;\theta)}{\partial \alpha \partial \beta} & \dfrac{\partial^2 \log g(Y_i;\theta)}{\partial \alpha \partial \sigma^2} \\ \dfrac{\partial^2 \log g(Y_i;\theta)}{\partial \beta \partial \alpha} & \dfrac{\partial^2 \log g(Y_i;\theta)}{\partial \beta^2} & \dfrac{\partial^2 \log g(Y_i;\theta)}{\partial \beta \partial \sigma^2} \\ \dfrac{\partial^2 \log g(Y_i;\theta)}{\partial \sigma^2 \partial \alpha} & \dfrac{\partial^2 \log g(Y_i;\theta)}{\partial \sigma^2 \partial \beta} & \dfrac{\partial^2 \log g(Y_i;\theta)}{\partial (\sigma^2)^2} \end{pmatrix}$$

$$= \begin{pmatrix} -\dfrac{1}{\sigma^2} & -\dfrac{X_i}{\sigma^2} & -\dfrac{u_i}{\sigma^4} \\ -\dfrac{X_i}{\sigma^2} & -\dfrac{X_i^2}{\sigma^2} & -\dfrac{X_i u_i}{\sigma^4} \\ -\dfrac{u_i}{\sigma^4} & -\dfrac{X_i u_i}{\sigma^4} & \dfrac{1}{2\sigma^4} - \dfrac{u_i^2}{\sigma^6} \end{pmatrix}$$

ただし，$u_i = Y_i - \alpha - \beta X_i$

期待値をとると，

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

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

となる。

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

$$= \begin{pmatrix} \dfrac{n}{\sigma^2} & \dfrac{\sum_{i=1}^n X_i}{\sigma^2} & 0 \\ \dfrac{\sum_{i=1}^n X_i}{\sigma^2} & \dfrac{\sum_{i=1}^n X_i^2}{\sigma^2} & 0 \\ 0 & 0 & \dfrac{n}{2\sigma^4} \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} \sigma^2 \begin{pmatrix} n & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix}^{-1} & 0 \\ 0 & \dfrac{2\sigma^4}{n} \end{pmatrix}$$

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。

$\longrightarrow$ 最小二乗推定量の分布と同じ。

誤差項に系列相関がある場合

回帰モデル

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

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

$\epsilon_2, \epsilon_3, \cdots, \epsilon_n$ は互いに独立で，すべての i について $\epsilon_i \sim \mathrm{N}(0, \sigma^2)$ 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。

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

尤度関数は,

$$\log l(\theta) = \sum_{i=2}^{n} \log f(Y_i; \theta) = -\frac{n-1}{2}\log(2\pi) - \frac{n-1}{2}\log(\sigma^2)$$
$$-\frac{1}{2\sigma^2}\sum_{i=2}^{n}\Big((Y_i - \rho Y_{i-1}) - \alpha(1-\rho) - \beta(X_i - \rho X_{i-1})\Big)^2$$

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。

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

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

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

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

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

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。
$\longrightarrow$ 下記のように収束計算によって求める。

(i) 初期段階では，$\hat{\rho} = 0$ とする。

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

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

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

(iv) $\hat{\alpha} = \dfrac{\widetilde{\alpha}}{1-\hat{\rho}}$

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

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

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

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

10.0.3　尤度比検定

n 個の確率変数 $X_1, X_2, \cdots, X_n$ 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 $\widetilde{\theta}$，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。

$\dfrac{l(\widetilde{\theta})}{l(\hat{\theta})}$ 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，

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

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。

$$\int \cdots \int \prod_{i=1}^{n} f(x_i; \widetilde{\theta}) \mathrm{d}x_1 \cdots \mathrm{d}x_n = \alpha$$

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，$n \longrightarrow \infty$ のとき，

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

となる。

　この検定を尤度比検定と呼ぶ。

例 1：　正規母集団 $\mathrm{N}(\mu, \sigma^2)$ からの標本値 $x_1, x_2, \cdots, x_n$ 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,

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

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,

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

となる。

尤度比検定統計量は,

$$\frac{l(\mu_0)}{l(\overline{X})} = \frac{\exp\Bigl(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \mu_0)^2\Bigr)}{\exp\Bigl(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i - \overline{X})^2\Bigr)}$$

$$= \exp\Bigl(-\frac{1}{2\sigma^2/n}(\overline{X} - \mu_0)^2\Bigr) < c$$

となる c を求める。

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

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

すなわち，

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

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：　$X_1, X_n, \cdots, X_n$ 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，

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

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。

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

$\rightarrow$ 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，$n \longrightarrow \infty$ のとき,

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

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

$\chi^2(1)$ 分布の上側 100 α% 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

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

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

について，$\beta_1, \cdots, \beta_k$ 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

$\theta = (\beta_1, \cdots, \beta_k, \sigma^2)$ とする。
尤度関数は,

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

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

となる。
H_0 の制約つき最尤推定量を $\widetilde{\theta} = (\widetilde{\beta}_1, \cdots, \widetilde{\beta}_k, \widetilde{\sigma}^2)$ とする。この仮設に含まれる制約数を G とする。

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

尤度比

$$
\begin{aligned}
\frac{l(\widetilde{\theta})}{l(\hat{\theta})} &= \frac{(2\pi\widetilde{\sigma}^2)^{-\frac{n}{2}} \exp\Bigl(-\frac{1}{2\widetilde{\sigma}^2} \sum_{i=1}^{n} (Y_i - \widetilde{\beta}_1 X_{1i} - \cdots - \widetilde{\beta}_k X_{ki})^2\Bigr)}{(2\pi\hat{\sigma}^2)^{-\frac{n}{2}} \exp\Bigl(-\frac{1}{2\hat{\sigma}^2} \sum_{i=1}^{n} (Y_i - \hat{\beta}_1 X_{1i} - \cdots - \hat{\beta}_k X_{ki})^2\Bigr)} \\
&= \frac{(\widetilde{\sigma}^2)^{-\frac{n}{2}} \exp\Bigl(-\frac{n-G}{2}\Bigr)}{(\hat{\sigma}^2)^{-\frac{n}{2}} \exp\Bigl(-\frac{n-k}{2}\Bigr)} \\
&= \left(\frac{\frac{1}{n-G} \sum_{i=1}^{n} \widetilde{u}_i^2}{\frac{1}{n-k} \sum_{i=1}^{n} \hat{u}_i^2} \right)^{-n/2} \exp\Bigl(-\frac{k-G}{2}\Bigr) \\
&= \exp\Bigl(-\frac{k-G}{2}\Bigr) \left(\frac{n-k}{n-G}\right)^{-n/2} \left(\frac{\sum_{i=1}^{n} \widetilde{u}_i^2}{\sum_{i=1}^{n} \hat{u}_i^2}\right)^{-n/2}
\end{aligned}
$$

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

のとき仮説を棄却する。

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

を利用すると c が求まる。

ただし，途中で以下を利用

$$\widetilde{\sigma}^2 = \frac{1}{n-G}\sum_{i=1}^{n}(Y_i - \widetilde{\beta}_1 X_{1i} - \cdots - \widetilde{\beta}_k X_{ki})^2$$
$$= \frac{1}{n-G}\sum_{i=1}^{n}\widetilde{u}_i^2$$

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

近似的には,

$$-2\log\frac{l(\widetilde{\theta})}{l(\hat{\theta})} = -2\log\frac{(\widetilde{\sigma}^2)^{-\frac{n}{2}}\exp\left(-\frac{n-G}{2}\right)}{(\hat{\sigma}^2)^{-\frac{n}{2}}\exp\left(-\frac{n-k}{2}\right)}$$

$$
\begin{aligned}
&= n\log\Big(\frac{\widetilde{\sigma}^2}{\hat{\sigma}^2}\Big) + (k-G) \\
&\longrightarrow \chi^2(G)
\end{aligned}
$$

例 4：　回帰モデル

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

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,

$$
\log l(\theta) = \sum_{i=2}^{n} \log f(Y_i;\theta) = -\frac{n-1}{2}\log(2\pi) - \frac{n-1}{2}\log(\sigma^2)
$$

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

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,

$$\log l(\theta) = \sum_{i=2}^{n} \log f(Y_i;\theta) = -\frac{n-1}{2}\log(2\pi) - \frac{n-1}{2}\log(\sigma^2)$$
$$-\frac{1}{2\sigma^2}\sum_{i=2}^{n}(Y_i - \alpha - \beta X_i)^2$$

となる。

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 $\widetilde{\theta} = (\widetilde{\alpha}, \widetilde{\beta}, \widetilde{\sigma}^2, 0)$ が得られる。

すなわち,

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

$\log l(\hat{\theta})$ は，$\hat{\sigma}^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=2}^{n}\Big((Y_i - \hat{\rho}Y_{i-1}) - \hat{\alpha}(1-\hat{\rho}) - \hat{\beta}(X_i - \hat{\rho}X_{i-1})\Big)^2$ に注意して,

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

となる。

同様に，$\log l(\widetilde{\theta})$ は，$\widetilde{\sigma}^2 = \dfrac{1}{n-1}\displaystyle\sum_{i=2}^{n}(Y_i - \widetilde{\alpha} - \widetilde{\beta}X_i)^2$ に注意して,

$$\log l(\widetilde{\theta}) = -\frac{n-1}{2}\log(2\pi) - \frac{n-1}{2}\log(\widetilde{\sigma}^2)$$

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

となる。

　したがって，尤度比検定統計量

$$
-2\log\frac{l(\widetilde{\theta})}{l(\hat{\theta})}=(n-1)\log\frac{\widetilde{\sigma}^2}{\hat{\sigma}^2}
$$

は，n が大きくなると，$\chi^2(1)$ 分布に近づく。