and $\Big(\dfrac{1}{n}\dfrac{\partial^2 \log L(\theta;X)}{\partial\theta\partial\theta'}\Big)^{-1}\Big(\dfrac{1}{\sqrt{n}}\dfrac{\partial \log L(\theta;X)}{\partial\theta}\Big)$ has the same asymptotic distribution as $\Sigma^{-1}\Big(\dfrac{1}{\sqrt{n}}\dfrac{\partial \log L(\theta;X)}{\partial\theta}\Big)$.

11. **Optimization (最適化):**

MLE of θ results in the following maximization problem:

$$\max_{\theta} \quad \log L(\theta;x).$$

We often have the case where the solution of θ is not derived in closed form.

$\Longrightarrow$ Optimization procedure

$$0=\frac{\partial \log L(\theta;x)}{\partial\theta}=\frac{\partial \log L(\theta^*;x)}{\partial\theta}+\frac{\partial^2 \log L(\theta^*;x)}{\partial\theta\partial\theta'}(\theta-\theta^*).$$

Solving the above equation with respect to θ, we obtain the following:

$$\theta=\theta^*-\left(\frac{\partial^2 \log L(\theta^*;x)}{\partial\theta\partial\theta'}\right)^{-1}\frac{\partial \log L(\theta^*;x)}{\partial\theta}.$$

Replace the variables as follows:

$$\theta \longrightarrow \theta^{(i+1)}, \qquad\qquad \theta^* \longrightarrow \theta^{(i)}.$$

Then, we have:

$$\theta^{(i+1)} = \theta^{(i)} - \left(\frac{\partial^2 \log L(\theta^{(i)};x)}{\partial\theta\partial\theta'}\right)^{-1} \frac{\partial \log L(\theta^{(i)};x)}{\partial\theta}.$$

$\Longrightarrow$ **Newton-Raphson method (ニュートン・ラプソン法)**

Replacing $\dfrac{\partial^2 \log L(\theta^{(i)};x)}{\partial\theta\partial\theta'}$ by $\mathrm{E}\left(\dfrac{\partial^2 \log L(\theta^{(i)};x)}{\partial\theta\partial\theta'}\right)$, we obtain the following optimization algorithm:

$$\begin{aligned}\theta^{(i+1)} &= \theta^{(i)} - \left(\mathrm{E}\left(\frac{\partial^2 \log L(\theta^{(i)};x)}{\partial\theta\partial\theta'}\right)\right)^{-1} \frac{\partial \log L(\theta^{(i)};x)}{\partial\theta} \\ &= \theta^{(i)} + \left(I(\theta^{(i)})\right)^{-1} \frac{\partial \log L(\theta^{(i)};x)}{\partial\theta}\end{aligned}$$

$\Longrightarrow$ **Method of Scoring (スコア法)**

3　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,

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

であるので，Y の密度関数は,

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

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。

4　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,

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

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,

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

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,

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

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,

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

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。一階微分は,

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

となる。二階微分は,

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

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

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。

したがって,

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

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

したがって,

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

となる。

$\longrightarrow$ 最小二乗推定量の分布と同じ。

5　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，

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

すなわち，

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

と変形できる。したがって，

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

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 $100\,\alpha\%$ 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

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

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

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

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

$$= \frac{(\widetilde{\sigma}^2)^{-\frac{n}{2}} \exp\Bigl(-\frac{n-G}{2}\Bigr)}{(\hat{\sigma}^2)^{-\frac{n}{2}} \exp\Bigl(-\frac{n-k}{2}\Bigr)} = \left(\frac{\frac{1}{n-G}\sum_{i=1}^{n}\widetilde{u}_i^2}{\frac{1}{n-k}\sum_{i=1}^{n}\hat{u}_i^2}\right)^{-n/2} \exp\Bigl(-\frac{k-G}{2}\Bigr)$$

$$= \exp\Bigl(-\frac{k-G}{2}\Bigr)\left(\frac{n-k}{n-G}\right)^{-n/2}\left(\frac{\sum_{i=1}^{n}\widetilde{u}_i^2}{\sum_{i=1}^{n}\hat{u}_i^2}\right)^{-n/2}$$

$$= \exp\Bigl(-\frac{k-G}{2}\Bigr)\left(\frac{n-k}{n-G}\right)^{-n/2}\left(1+\frac{\sum_{i=1}^n \widetilde{u}_i^2 - \sum_{i=1}^n \hat{u}_i^2}{\sum_{i=1}^n \hat{u}_i^2}\right)^{-n/2}$$

$$= \exp\Bigl(-\frac{k-G}{2}\Bigr)\left(\frac{n-k}{n-G}\right)^{-n/2}\left(1+\frac{G}{n-k}\frac{(\sum_{i=1}^n \widetilde{u}_i^2 - \sum_{i=1}^n \hat{u}_i^2)/G}{\sum_{i=1}^n \hat{u}_i^2/(n-k)}\right)^{-n/2} < c$$

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,

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