Solution Tutorial 8: ML-Estimation

Exercises

1. Bernoulli

   (a) $l_i(\theta) = log[f(y_i, \theta)] = y_i * log(\theta) + (1 - y_i) * log(1 - \theta)$

   $L(\theta) = \sum_{i=1}^{N} l_i(\theta) = log(\theta) \sum_{i=1}^{N} y_i + log(1 - \theta) \sum_{i=1}^{N}(1 - y_i)$

   (b) $s_i(\theta) = \frac{\partial\, l_i(\theta)}{\partial\, \theta} = \frac{1}{\theta} y_i + \frac{1}{1-\theta}(1 - y_i)(-1) = \frac{1}{\theta} y_i - \frac{1}{1-\theta}(1 - y_i)$

   FOC:

$$\begin{aligned}
\sum_{i=1}^{N} s_i(\hat{\theta}) &= \frac{1}{\hat{\theta}} \sum_{i=1}^{N} y_i - \frac{1}{1-\hat{\theta}} \sum_{i=1}^{N}(1 - y_i) \stackrel{!}{=} 0 \\
&\Rightarrow \frac{1}{\hat{\theta}} \sum_{i=1}^{N} y_i - \frac{1}{1-\hat{\theta}} N + \frac{1}{1-\hat{\theta}} \sum_{i=1}^{N} y_i = 0 \\
&\Rightarrow \frac{1}{\hat{\theta}} \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} y_i - N + \sum_{i=1}^{N} y_i = 0 \\
&\Rightarrow \frac{1}{\hat{\theta}} \sum_{i=1}^{N} y_i = N \\
\hat{\theta}_{ML} &= \frac{1}{N} \sum_{i=1}^{N} y_i = \bar{y}
\end{aligned}$$

   (c) Bias: $E(\hat{\theta}) = E(\bar{y}) = \frac{1}{N} \sum_{i=1}^{N} E(y_i) = \frac{1}{N} \sum_{i=1}^{N} \theta = \theta$ unbiased

(d) Variance:

$$\begin{aligned}
Var(\hat{\theta}) &= E\Big[(\hat{\theta}-\theta)^2\Big] = E\Big[(\frac{1}{N}\sum_{i=1}^{N} y_i - \theta)^2\Big] = E\Big[\big(\frac{1}{N}\sum_{i=1}^{N}(y_i-\theta)\big)^2\Big] \\
&= \frac{1}{N^2}E\Big[\sum_{i=1}^{N}(y_i-\theta)^2 + \sum_{i=1}^{N}\sum_{j=1}^{N}(y_i-\theta)(y_j-\theta)\Big] \text{ , with i} \neq \text{j} \\
&= \frac{1}{N^2}\sum_{i=1}^{N}\underbrace{E(y_i-\theta)^2}_{Var(y_i)=\theta(1-\theta)} + \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\underbrace{E\big[(y_i-\theta)(y_j-\theta)\big]}_{Cov(y_i,y_j)=0 \text{ (random sample!)}} \\
&= \frac{1}{N^2}N\theta(1-\theta) = \frac{\theta(1-\theta)}{N}
\end{aligned}$$

or

$$\begin{aligned}
Var(\hat{\theta}) &= Var\Big[\frac{1}{N}\sum_{i=1}^{N} y_i\Big] = \frac{1}{N^2}Var\Big[\sum_{i=1}^{N} y_i\Big] \\
&\overset{iid}{=} \frac{1}{N^2}\sum_{i=1}^{N} Var\Big[y_i\Big] = \frac{1}{N^2}N\theta(1-\theta) = \frac{\theta(1-\theta)}{N}
\end{aligned}$$

(e) Note: We have a random sample and the random variable $v_i = y_i - \theta$ has mean zero and variance $Var(v_i) = Var(y_i) = \theta(1-\theta)$. Hence, $N^{-\frac{1}{2}}\sum_{i=1}^{N} v_i \xrightarrow{d} N(0, \theta(1-\theta))$

$$\begin{aligned}
N^{-\frac{1}{2}}\sum_{i=1}^{N} v_i &= N^{-\frac{1}{2}}\sum_{i=1}^{N}(y_i-\theta) = \frac{\sqrt{N}}{N}\sum_{i=1}^{N}(y_i-\theta) = \sqrt{N}\Big[\Big(\frac{1}{N}\sum_{i=1}^{N} y_i\Big) - \theta\Big] \\
&= \sqrt{N}(\hat{\theta}-\theta) \\
&\Rightarrow \sqrt{N}(\hat{\theta}-\theta) \xrightarrow{d} N(0, \theta(1-\theta)) \\
&\Rightarrow \hat{\theta} \xrightarrow{d} \mathcal{N}\Big(\theta, \frac{\theta(1-\theta)}{N}\Big)
\end{aligned}$$

2. Poisson

(a)

$$
\begin{aligned}
l_i(\lambda) &= log[f(y_i, \lambda)] = -\lambda + y_i * log(\lambda) - log(y_i!) \\
\\
L(\lambda) &= \sum_{i=1}^{N} l_i(\lambda) = -\lambda N + \sum_{i=1}^{N} [y_i * log(\lambda)] - \sum_{i=1}^{N} log(y_i!) \\
&= -\lambda N + log(\lambda) \sum_{i=1}^{N} y_i - \sum_{i=1}^{N} log(y_i!)
\end{aligned}
$$

(b) $s_i(\lambda) = \frac{\partial\, l_i(\lambda)}{\partial\, \lambda} = -1 + \frac{y_i}{\lambda}$

$E[s_i(\lambda)] = -1 + \frac{1}{\lambda} E(y_i) = -1 + \frac{1}{\lambda}\lambda = 0$

$Var[s_i(\lambda)] = E[s_i(\lambda)^2] = E\left[\frac{y_i^2}{\lambda^2} - \frac{2y_i}{\lambda} + 1\right] = \frac{1}{\lambda^2} E(y_i^2) - \frac{2}{\lambda} E(y_i) + 1$
Note: $E(y_i^2) = Var(y_i) + E(y_i)^2 = \lambda + \lambda^2$

$\Rightarrow Var[s_i(\lambda)] = \frac{1}{\lambda^2}(\lambda + \lambda^2) - \frac{2}{\lambda}\lambda + 1 = \frac{1}{\lambda} + 1 - 2 + 1 = \frac{1}{\lambda}$

CLT: $N^{-\frac{1}{2}} \sum_{i=1}^{N} s_i(\lambda) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{\lambda}\right)$

(c) FOC: $\sum_{i=1}^{N} s_i(\hat{\lambda}) = -N + \frac{1}{\hat{\lambda}} \sum_{i=1}^{N} y_i \overset{!}{=} 0 \Rightarrow \hat{\lambda}_{ML} = \frac{1}{N} \sum_{i=1}^{N} y_i = \bar{y}$

(d) Bias: $E(\hat{\lambda}) = E(\bar{y}) = \frac{1}{N} \sum_{i=1}^{N} E(y_i) = \lambda$ unbiased

(e) Variance:

$$\begin{aligned}
Var(\hat{\lambda}) &= E\Big[(\hat{\lambda}-\lambda)^2\Big] = E\Big[(\frac{1}{N}\sum_{i=1}^{N} y_i - \lambda)^2\Big] = E\Big[\big(\frac{1}{N}\sum_{i=1}^{N}(y_i-\lambda)\big)^2\Big] \\
&= \frac{1}{N^2}E\Big[\sum_{i=1}^{N}(y_i-\lambda)^2 + \sum_{i=1}^{N}\sum_{j=1}^{N}(y_i-\lambda)(y_j-\lambda)\Big] \text{ , with i} \neq \text{j} \\
&= \frac{1}{N^2}\sum_{i=1}^{N}\underbrace{E(y_i-\lambda)^2}_{Var(y_i)=\lambda} + \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\underbrace{E\big[(y_i-\lambda)(y_j-\lambda)\big]}_{Cov(y_i,y_j)=0 \text{ (random sample!)}} \\
&= \frac{1}{N^2}\sum_{i=1}^{N}\lambda = \frac{\lambda}{N}
\end{aligned}$$

or

$$\begin{aligned}
Var(\hat{\lambda}) &= Var\Big[\frac{1}{N}\sum_{i=1}^{N} y_i\Big] = \frac{1}{N^2}Var\Big[\sum_{i=1}^{N} y_i\Big] \\
&\overset{iid}{=} \frac{1}{N^2}\sum_{i=1}^{N} Var\Big[y_i\Big] = \frac{1}{N^2}N\lambda = \frac{\lambda}{N}
\end{aligned}$$

(f) $H(y_i,\lambda) = \frac{\partial\, s_i(\lambda)}{\partial\, \lambda} = -\frac{y_i}{\lambda^2} \Rightarrow E[H(y_i,\lambda)] = -\frac{E(y_i)}{\lambda^2} = -\frac{\lambda}{\lambda^2} = -\frac{1}{\lambda}$

(g)
- Direct application of CLT
  $E(y_i-\lambda)=0$, $Var(y_i-\lambda)=Var(y_i)=\lambda$, y_i from random sample

$$\begin{aligned}
N^{-\frac{1}{2}}\sum_{i=1}^{N}(y_i-\lambda) &= \frac{\sqrt{N}}{N}\sum_{i=1}^{N}(y_i-\lambda) = \sqrt{N}\Big[\big(\frac{1}{N}\sum_{i=1}^{N} y_i\big) - \lambda\Big] \\
&= \sqrt{N}(\hat{\lambda}-\lambda) \xrightarrow{d} \mathcal{N}(0,\lambda)
\end{aligned}$$

$$\Rightarrow \hat{\lambda} \xrightarrow{d} \mathcal{N}\Big(\lambda, \frac{\lambda}{N}\Big)$$

- Application of the results for ML estimators
  $\sqrt{N}(\hat{\lambda} - \lambda) \xrightarrow{d} \mathcal{N}\big(0, A_0^{-1}\big)$(information matrix equality holds in this case)

  where $A_0 = -E[H(y_i, \lambda)] = \frac{1}{\lambda} \Rightarrow A_0^{-1} = \lambda$

  $\Rightarrow \sqrt{N}(\hat{\lambda} - \lambda) \xrightarrow{d} \mathcal{N}(0, \lambda)$

3.

(a) Note: $x_i = (x_{1i}\ x_{2i})$ for two regressors

$$\begin{aligned} l_i(\theta) &= -\mu(x_i) + y_i * log[\mu(x_i)] - log(y_i!) \\ &= -exp(x_i\theta) + y_i x_i \theta - log(y_i!) \end{aligned}$$

$$L(\theta) = \sum_{i=1}^{N} l_i(\theta) = \sum_{i=1}^{N} [-exp(x_i\theta) + y_i x_i \theta] - \sum_{i=1}^{N} log(y_i!)$$

(b) $s_i(\theta) = \nabla'_\theta\ l_i(\theta) = -exp(x_i\theta)x'_i + y_i x'_i = x'_i[y_i - exp(x_i\theta)]$

$E[s_i(\theta)|x_i] = x'_i[E(y_i|x_i) - exp(x_i\theta)] = x'_i[exp(x_i\theta) - exp(x_i\theta)] = 0$

(c) FOC:

$$\begin{aligned} \sum_{i=1}^{N} s_i(\hat{\theta}) &= \sum_{i=1}^{N} x'_i[y_i - exp(x_i\hat{\theta})] \overset{!}{=} 0 \\ &\Rightarrow \sum_{i=1}^{N} x'_i y_i = \sum_{i=1}^{N} x'_i exp(x_i\hat{\theta}) \end{aligned}$$

only numerically solvable, can get $\hat{\theta}$ from that

(d) Hessian: $H_i(\theta) = \nabla_\theta\ s_i(\theta) = -exp(x_i\theta)x'_i x_i$

(e)

$$E[s_i(\theta_0)s_i(\theta_0)'|x_i] = E\Big[x_i'x_i\big(y_i - exp(x_i\theta_0)\big)^2|x_i\Big] = x_i'x_i \; E\Big[\big(y_i - E(y_i|x_i)\big)^2|x_i\Big]$$
$$= x_i'x_i \; Var(y_i|x_i) = x_i'x_i \; E(y_i|x_i) = x_i'x_i \; exp(x_i\theta_0)$$

as well: $-E[H_i(\theta_0)|x_i] = x_i'x_i \; exp(x_i\theta_0)$

$\Rightarrow$ CIME holds, because $E[s_i(\theta_0)s_i(\theta_0)'|x_i] = -E[H_i(\theta_0)|x_i]$

(f) Remember: $A = -E[H_i(\theta_0)|x_i]$ and $A_0 = -E[H_i(\theta_0)]$

$$A_0 = -E[H_i(\theta_0)] = E[exp(x_i\theta_0)x_i'x_i]$$
$$Avar(\hat{\theta}) = \frac{A_0^{-1}}{N} = \frac{\Big(E\big[exp(x_i\theta_0)x_i'x_i\big]\Big)^{-1}}{N}$$
$$\widehat{Avar}(\hat{\theta}) = \frac{\Big(\frac{1}{N}\sum_{i=1}^N exp(x_i\hat{\theta})x_i'x_i\Big)^{-1}}{N} = \Big(\sum_{i=1}^N exp(x_i\hat{\theta})x_i'x_i\Big)^{-1}$$

plug in solution from c) and obtain estimate for asympt. variance for tests!

$\Rightarrow$ Asymptotic standard errors: square root of main diagonal of $\widehat{Avar}(\hat{\theta})$

4.

(a) When u is cond. normally distributed $N(0, \sigma^2)$, then $y = x\beta + u$ is cond. normally distributed as well: $y_i \sim N(x\beta, \sigma^2)$. The pdf of this normal distribution is given as:

$$f(y_i|x_i, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_i - x_i\beta)^2}, \theta = (\beta', \sigma^2)', u_i = y_i - x_i\beta$$

$l_i(\theta) = -\frac{1}{2}log(2\pi) - \frac{1}{2}log(\sigma^2) - \frac{1}{2\sigma^2}(y_i - x_i\beta)^2$

$L(\theta) = -\frac{N}{2}log(2\pi) - \frac{N}{2}log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - x_i\beta)^2$

(b) Note: $\theta = \begin{bmatrix} \beta \\ \sigma^2 \end{bmatrix}$

$$\begin{aligned} s_i(\theta) &= \frac{\partial\, l_i(\theta)}{\partial\, \theta} \\ s_{1i}(\theta) &= \frac{\partial\, l_i(\theta)}{\partial\, \beta} = \frac{1}{\sigma^2}x_i'u_i \\ s_{2i}(\theta) &= \frac{\partial\, l_i(\theta)}{\partial\, \sigma^2} = -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}u_i^2 = \frac{u_i^2 - \sigma^2}{2\sigma^4}(suppose\sigma^2 = a, substitutionrule) \end{aligned}$$

$\Rightarrow s_i(\theta) = \begin{bmatrix} \frac{1}{\sigma^2}x_i'u_i \\ \frac{u_i^2 - \sigma^2}{2\sigma^4} \end{bmatrix}$

$E[s_{1i}(\theta_0)|x_i] = \frac{1}{\sigma_0^2}x_i'E(u_i|x_i) = 0$

$E[s_{2i}(\theta_0)|x_i] = -\frac{1}{2\sigma_0^2} + \frac{1}{2\sigma_0^4}\underbrace{E(u_i^2|x_i)}_{\sigma_0^2} = -\frac{1}{2\sigma_0^2} + \frac{1}{2\sigma_0^2} = 0$

(c) FOC's:

i) $\sum_{i=1}^{N} \frac{1}{\hat{\sigma}^2} x_i' \hat{u}_i \stackrel{!}{=} 0 \;\Rightarrow\; \sum_{i=1}^{N} (x_i' y_i - x_i' x_i \hat{\beta}) = 0 \Rightarrow \sum_{i=1}^{N} x_i' y_i = \Big(\sum_{i=1}^{N} x_i' x_i \Big) \hat{\beta}$

$$\Rightarrow \; \hat{\beta} = \Big(\sum_{i=1}^{N} x_i' x_i \Big)^{-1} \sum_{i=1}^{N} x_i' y_i$$

ii) $\sum_{i=1}^{N} \Big(\frac{\hat{u}_i^2}{2\hat{\sigma}^4} - \frac{1}{2\hat{\sigma}^2} \Big) \stackrel{!}{=} 0 \;\Rightarrow\; \frac{1}{2\hat{\sigma}^4} \sum_{i=1}^{N} \hat{u}_i^2 = \frac{N}{2\hat{\sigma}^2} \Rightarrow \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^{N} \hat{u}_i^2$

$$\Rightarrow \; \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^{N} (y_i - x_i \hat{\beta})^2 = \frac{1}{N} \sum_{i=1}^{N} \hat{u}_i{}^2$$

(d) $E(\hat{\beta}|X) = (X'X)^{-1} X' E(Y|X) = (X'X)^{-1} X' X \beta_0 = \beta_0$

$\Rightarrow E(\hat{\beta}) = E\big[E(\hat{\beta}|X)\big] = E[\beta_0] = \beta_0$

$E(\hat{\sigma}^2|X) = E\big(\frac{1}{N} \hat{U}'\hat{U}|X\big) = \frac{1}{N} E(U' M_X U|X) = \frac{\sigma_0^2}{N} E\Big(\frac{U' M_X U}{\sigma_0^2}|X\Big)$

Recall from the formulary: $\frac{U' M_X U}{\sigma_0^2}|X \sim \chi^2_{N-K}$

if $\nu \sim \chi^2_{N-K} \Rightarrow E(\nu) = N - K \Rightarrow E\Big(\frac{U' M_X U}{\sigma_0^2}|X\Big) = N - K$

$\Rightarrow E(\hat{\sigma}^2|X) = \frac{\sigma_0^2}{N}(N - K)$

$\Rightarrow E(\hat{\sigma}^2) = E\big[E(\hat{\sigma}^2|X)\big] = \sigma_0^2 \frac{N-K}{N} \neq \sigma_0^2$

(e) Hessian:

$$\frac{\partial^2\, l_i(\theta)}{\partial\, \beta\, \partial\, \beta'} = \frac{\partial}{\partial\, \beta'} \Big[\frac{1}{\sigma^2} x_i'(y_i - x_i\beta) \Big] = -\frac{1}{\sigma^2} x_i' x_i$$

$$\frac{\partial^2\, l_i(\theta)}{\partial\, \beta\, \partial\, \sigma^2} = \frac{\partial}{\partial\, \sigma^2} \Big[\frac{1}{\sigma^2} x_i'(y_i - x_i\beta) \Big] = -\frac{1}{\sigma^4} x_i' u_i$$

$$\frac{\partial^2\, l_i(\theta)}{\partial\, \sigma^2\, \partial\, \sigma^2} = \frac{\partial}{\partial\, \sigma^2} \Big[-\frac{1}{2\sigma^2} + \frac{u_i^2}{2\sigma^4} \Big] = \frac{1}{2\sigma^4} - \frac{u_i^2}{\sigma^6} = \frac{\sigma^2 - 2u_i^2}{2\sigma^6}$$

$$H_i(\theta) = \begin{bmatrix} -\frac{1}{\sigma^2} x_i' x_i & -\frac{1}{\sigma^4} x_i' u_i \\ -\frac{1}{\sigma^4} x_i u_i & \frac{\sigma^2 - 2u_i^2}{2\sigma^6} \end{bmatrix}$$

(f) $-E\big[H_i(\theta_0)|x_i\big] = \begin{bmatrix} \frac{1}{\sigma_0^2} x_i' x_i & 0 \\ 0 & \frac{1}{2\sigma_0^4} \end{bmatrix}$, Note: $E\Big[\frac{\sigma^2 - 2u_i^2}{2\sigma^6}|x_i\Big] = \frac{-\sigma^2}{2\sigma^6} = -\frac{1}{2\sigma_0^4}$

remember: $s_i(\theta) = \begin{bmatrix} \frac{1}{\sigma^2} x_i' u_i \\ \frac{u_i^2 - \sigma^2}{2\sigma^4} \end{bmatrix}$

$\Rightarrow E(s_i(\theta_0)s_i(\theta_0)'|x_i) = E\left\{ \begin{bmatrix} \frac{1}{\sigma_0^4} x_i' x_i u_i^2 & \frac{u_i^2 - \sigma_0^2}{2\sigma_0^4} \cdot \frac{1}{\sigma_0^2} x_i' u_i \\ \frac{u_i^2 - \sigma_0^2}{2\sigma_0^4} \cdot \frac{1}{\sigma_0^2} x_i u_i & \frac{(u_i^2 - \sigma_0^2)^2}{4\sigma_0^8} \end{bmatrix} \Bigg| x_i \right\}$

(i) $E\Big[\frac{1}{\sigma_0^4} x_i' x_i u_i^2 \Big| x_i\Big] = \frac{1}{\sigma_0^2} x_i' x_i$

(ii) $E\Big[\frac{u_i^3 - u_i \sigma_0^2}{2\sigma_0^6} x_i' \Big| x_i\Big] = \frac{E(u_i^3|x_i) - \sigma_0^2 E(u_i|x_i)}{2\sigma_0^6} x_i' = 0$, because $E(u_i^3|x_i) = 0$

(iii) $E\Big[\frac{u_i^4 - 2\sigma_0^2 u_i^2 + \sigma_0^4}{2 \cdot 2\sigma_0^8} \Big| x_i\Big] = \frac{E(u_i^4|x_i) - 2\sigma_0^2 E(u_i^2|x_i) + \sigma_0^4}{2 \cdot 2\sigma_0^8} = \frac{3\sigma_0^4 - 2\sigma_0^4 + \sigma_0^4}{4\sigma_0^8}$
$= \frac{2\sigma_0^4}{4\sigma_0^8} = \frac{1}{2\sigma_0^4}$

$\Rightarrow E\big[s_i(\theta_0)s_i(\theta_0)'|x_i\big] = \begin{bmatrix} \frac{1}{\sigma_0^2} x_i' x_i & 0 \\ 0 & \frac{1}{2\sigma_0^4} \end{bmatrix}$

$\Rightarrow$ CIME holds.

(g) $A(x_i, \theta_0) = -E\big[H(x_i, \theta_0)|x_i\big] = \begin{bmatrix} \frac{1}{\sigma_0^2} x_i' x_i & 0 \\ 0 & \frac{1}{2\sigma_0^4} \end{bmatrix}$

$\hat{A} = \frac{1}{N} \sum_{i=1}^N A(x_i, \hat{\theta})$

$\Rightarrow \hat{A} = \begin{bmatrix} \frac{1}{\hat{\sigma}^2} \frac{1}{N} \sum_i x_i' x_i & 0 \\ 0 & \frac{1}{2\hat{\sigma}^4} \end{bmatrix}$

This is much simpler and uses more "structure" than to base it on $-H_i(\sigma)$ which yields ($\hat{A} = \frac{1}{N} \sum_{i=1}^N -H_i(\hat{\theta})$)

$$\hat{A} = \begin{bmatrix} \frac{1}{\hat\sigma^2}\frac{1}{N}\sum_{i=1}^N x_i'x_i & \frac{1}{\hat\sigma^4}\frac{1}{N}\sum_{i=1}^N x_i'\hat{u}_i \\ +\frac{1}{\hat\sigma^4}\frac{1}{N}\sum_{i=1}^N x_i\hat{u}_i & -\frac{1}{2\hat\sigma^4}+\frac{1}{N}\sum_{i=1}^N \frac{\hat{u}_i^2}{\hat\sigma^6} \end{bmatrix}$$

However, this can be simplified because

(i) $\sum_{i=1}^N x_i'\hat{u}_i = 0$

(ii) $\sum_{i=1}^N \hat{u}_i^2/N = \hat\sigma^2 \Rightarrow -\frac{1}{2\hat\sigma^4}+\frac{\hat\sigma^2}{\hat\sigma^6} = \frac{-\hat\sigma^2+2\hat\sigma^2}{2\hat\sigma^6} = +\frac{\hat\sigma^2}{2\hat\sigma^6} = +\frac{1}{2\hat\sigma^4}$

$$\Rightarrow \hat{A} = \begin{bmatrix} \frac{1}{\hat\sigma^2}\frac{1}{N}\sum_i x_i'x_i & 0 \\ 0 & \frac{1}{2\hat\sigma^4} \end{bmatrix}$$

Hence, numerically both approaches lead to the same result.

(h)

$$\begin{aligned}
A_0 &= -E[H_i(\theta_0)] \overset{LIE}{=} -E\big[E[H_i(\theta_0|x_i)]\big] \\
&= E\big[-E[H_i(\theta_0)]\big] = E[A(x_i,\theta_0)] \text{ (see g)} \\
&= \begin{bmatrix} \frac{1}{\sigma_0^2}E(x_i'x_i) & 0 \\ 0 & \frac{1}{2\sigma_0^4} \end{bmatrix} \\
V &= A_0^{-1} = \begin{bmatrix} \sigma_0^2\big[E(x_i'x_i)\big]^{-1} & 0 \\ 0 & 2\sigma_0^4 \end{bmatrix} \\
Avar(\hat\beta,\hat\sigma^2) &= \frac{A_0^{-1}}{N} = \begin{bmatrix} \frac{\sigma_0^2}{N}\big[E(x_i'x_i)\big]^{-1} & 0 \\ 0 & \frac{2}{N}\sigma_0^4 \end{bmatrix} \\
\Rightarrow \widehat{Avar}(\hat\beta,\hat\sigma^2) &= \begin{bmatrix} \hat\sigma^2\big[(\sum_{i=1}^N x_i'x_i)\big]^{-1} & 0 \\ 0 & \frac{2}{N}\hat\sigma^4 \end{bmatrix}
\end{aligned}$$

$\Rightarrow$ Asymptotic standard error of $\hat\beta$: square root of main diagonal elements of $\hat\sigma^2\big[(\sum_{i=1}^N x_i'x_i)\big]^{-1}$

$\Rightarrow$ Asymptotic standard error of $\hat\sigma^2$: $\sqrt{\frac{2}{n}}\,\hat\sigma^2$

(i) Generalized Gauss-Newton:

$$\theta^{\{g+1\}} = \theta^{\{g\}} + \left[\sum_i A(x_i, \theta^{\{g\}})\right]^{-1} \left[\sum_i s_i(\theta^{\{g\}})\right]$$

Model: $y_i = \beta_0 + u_i, \qquad x_i = 1$

$$\Rightarrow \sum_{i=1}^N x_i' u_i = \sum_{i=1}^N (y_i - \beta_0) = N\bar{y} - N\beta_0$$

$$\sum_{i=1}^N u_i^2 = \sum_{i=1}^N (y_i - \beta_0)^2 = N\bar{y^2} - 2N\beta_0\bar{y} + N\beta_0^2$$

Score:

$$\sum_{i=1}^N s_i(\theta) = \begin{bmatrix} \frac{1}{\sigma^2}\sum_i x_i' u_i \\ \frac{1}{2\sigma^4}\sum_i (u_i^2 - \sigma^2) \end{bmatrix} = \begin{bmatrix} \frac{1}{\sigma^2}(N\bar{y} - N\beta_0) \\ \frac{1}{2\sigma^4}(N\bar{y^2} - 2N\beta_0\bar{y} + N\beta_0^2 - N\sigma^2) \end{bmatrix}$$

Hessian:

$$\left[\sum_{i=1}^N A(x_i, \theta)\right]^{-1} = \begin{bmatrix} \sigma^2(\sum_i x_i' x_i)^{-1} & 0 \\ 0 & 2\frac{\sigma^4}{N} \end{bmatrix} = \begin{bmatrix} \frac{\sigma^2}{N} & 0 \\ 0 & \frac{2\sigma^4}{N} \end{bmatrix}$$

$$\Rightarrow \left[\sum_{i=1}^N A(x_i, \theta)\right]^{-1} \cdot \left[\sum_{i=1}^N s_i(\theta)\right] = \begin{bmatrix} \frac{\sigma^2}{N}\frac{1}{\sigma^2}(N\bar{y} - N\beta_0) \\ 2\frac{\sigma^4}{N}\frac{1}{2\sigma^4}N(\bar{y^2} - 2\beta_0\bar{y} + \beta_0^2 - \sigma^2) \end{bmatrix} = \begin{bmatrix} (\bar{y} - \beta_0) \\ (\bar{y^2} - 2\beta_0\bar{y} + \beta_0^2 - \sigma^2) \end{bmatrix}$$

$$\Rightarrow \begin{pmatrix} \beta_0^{(g+1)} \\ (\sigma^2)^{(g+1)} \end{pmatrix} = \begin{pmatrix} \beta_0^{(g)} \\ (\sigma^2)^{(g)} \end{pmatrix} + \begin{pmatrix} \bar{y} - \beta_0^{(g)} \\ \bar{y^2} - 2\beta_0^{(g)}\bar{y} + (\beta_0^{(g)})^2 - (\sigma^2)^{(g)} \end{pmatrix}$$

$$= \begin{pmatrix} \bar{y} \\ \bar{y^2} - 2\beta_0^{(g)}\bar{y} + (\beta_0^{(g)})^2 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \beta_0^{(1)} \\ (\sigma^2)^{(1)} \end{pmatrix} = \begin{pmatrix} 1.5 \\ 6.25 - 2 \cdot 0 \cdot 1.5 + 0^2 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 6.25 \end{pmatrix}$$

$$\begin{pmatrix} \beta_0^{(2)} \\ (\sigma^2)^{(2)} \end{pmatrix} = \begin{pmatrix} 1.5 \\ 6.25 - 2 \cdot 1.5 \cdot 1.5 + 1.5^2 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 6.25 - 2.25 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 4 \end{pmatrix}$$

No further change: finale estimate!

$$\begin{pmatrix} \beta_0^{(3)} \\ (\sigma^2)^{(3)} \end{pmatrix} = \begin{pmatrix} 1.5 \\ 6.25 - 2 \cdot 1.5 \cdot 1.5 + 1.5^2 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 6.25 - 2.25 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 4 \end{pmatrix}$$

5.

(a) LR-Test

$$y = x_1\beta_{10} + x_2\beta_{20} + u \qquad\qquad H_0 : \beta_{20} = 0$$

$\Rightarrow H_0 : y = x_1\beta_{10} + u \quad H_1 : y = x_1\beta_{10} + x_2\beta_{20} + u$

$l_i(\theta) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}(y_i - x_i\beta)^2$

$L(\theta) = -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - x_i\beta)^2$

- ML-estimator under $H_0 : \tilde{\theta} = \begin{bmatrix} \tilde{\beta}_1 \\ \tilde{\sigma}^2 \end{bmatrix}$

$$\tilde{\beta}_1 = (x_1'x_1)^{-1}x_1'y \qquad \tilde{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - x_{1i}\tilde{\beta}_1)^2$$

(see ex. 4 FOC)

$$\begin{aligned} L(\tilde{\theta}) &= -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log(\tilde{\sigma}^2) - \frac{1}{2\tilde{\sigma}^2}\underbrace{\sum_{i=1}^{N}(y_i - x_{1i}\tilde{\beta}_1)^2}_{N\tilde{\sigma}^2} \\ &= -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log(\tilde{\sigma}^2) - \frac{N}{2} \end{aligned}$$

- ML-estimator under $H_1 : \hat{\theta} = \begin{bmatrix} \hat{\beta} \\ \hat{\sigma}^2 \end{bmatrix}$

$$\hat{\beta} = (x'x)^{-1}x'y \qquad \hat{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - x_i\hat{\beta})^2$$

$$L(\hat{\theta}) = -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log(\hat{\sigma}^2) - \frac{N}{2}$$

- LR statistic

$$\begin{aligned} LR &= 2\big[L(\hat{\theta}) - L(\tilde{\theta})\big] = 2\Big[-\frac{N}{2}\log(\hat{\sigma}^2) + \frac{N}{2}\log(\tilde{\sigma}^2)\Big] \\ &= N\cdot\log(\tilde{\sigma}^2) - N\cdot\log(\hat{\sigma}^2) = N\cdot\log\left(\frac{\tilde{\sigma}^2}{\hat{\sigma}^2}\right) \end{aligned}$$

(b) LM-Test $\qquad H_0 : \beta_{20} = 0$

- score for model under H_1: (known from 4b)
  $s_i(\theta) = \frac{1}{\sigma^2} \begin{bmatrix} x_i' u_i \\ \frac{u_i^2}{2\sigma^2} - \frac{1}{2} \end{bmatrix}$
  $\sum_{i=1}^N s_i(\theta) = \frac{1}{\sigma^2} \begin{bmatrix} \sum_{i=1}^N x_i' u_i \\ \frac{1}{2\sigma^2} \sum_{i=1}^N u_i^2 - \frac{N}{2} \end{bmatrix}$

- evaluate under H_0 $\qquad \tilde{u}_i = y_i - x_{1i}\tilde{\beta}_1 \qquad \tilde{\sigma}^2 = \frac{1}{N}\sum_{i=1}^N \tilde{u}_i^2$

$$\begin{aligned} \sum_{i=1}^N s_i(\tilde{\theta}) &= \frac{1}{\tilde{\sigma}^2} \begin{bmatrix} \sum_{i=1}^N x_i' \tilde{u}_i \\ \frac{1}{2\tilde{\sigma}^2} \sum_{i=1}^N \tilde{u}_i^2 - \frac{N}{2} \end{bmatrix} \\ &= \frac{1}{\tilde{\sigma}^2} \begin{bmatrix} \sum_i x_i' \tilde{u}_i \\ \frac{N}{2} - \frac{N}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{\tilde{\sigma}^2} \sum_i x_i' \tilde{u}_i \\ 0 \end{bmatrix} \end{aligned}$$

- A matrix for model under H_1:
  (known from 4g)

$$A(x_i, \theta_0) = \begin{bmatrix} \frac{1}{\sigma_0^2} x_i' x_i & 0 \\ 0 & \frac{1}{2\sigma_0^4} \end{bmatrix}$$

- Evaluate matrix under H_0:

$$\tilde{A} = N^{-1} \sum_{i=1}^N \begin{bmatrix} \frac{1}{\tilde{\sigma}^2} x_i' x_i & 0 \\ 0 & \frac{1}{2\tilde{\sigma}^4} \end{bmatrix} = \begin{bmatrix} \frac{1}{N\tilde{\sigma}^2} \sum_{i=1}^N x_i' x_i & 0 \\ 0 & \frac{1}{2\tilde{\sigma}^4} \end{bmatrix}$$

- Compute the LM-statistic:

$$\begin{aligned} LM &= \frac{1}{N} \sum_{i=1}^N s_i(\tilde{\theta})' \tilde{A}^{-1} \sum_{i=1}^N s_i(\tilde{\theta}) \\ LM &= \frac{1}{N} \Big[\frac{1}{\tilde{\sigma}^2} \big(\sum_{i=1}^N x_i' \tilde{u}_i \big)' \;\; 0 \Big] \begin{bmatrix} \tilde{\sigma}^2 N \big(\sum_{i=1}^N x_i' x_i \big)^{-1} & 0 \\ 0 & 2\tilde{\sigma}^4 \end{bmatrix} \begin{bmatrix} \frac{1}{\tilde{\sigma}^2} \sum_{i=1}^N x_i' \tilde{u}_i \\ 0 \end{bmatrix} \\ &= \frac{\big(\sum_{i=1}^N x_i' \tilde{u}_i \big)' \big(\sum_{i=1}^N x_i' x_i \big)^{-1} \big(\sum_{i=1}^N x_i' \tilde{u}_i \big)}{\tilde{\sigma}^2} \end{aligned}$$

(c) Transform LM-statistic:

$$LM = \frac{\tilde{u}'x(x'x)^{-1}x'\tilde{u}}{\tilde{u}'\tilde{u}/N} = \frac{\tilde{u}'P_x\tilde{u}}{\tilde{u}'\tilde{u}} \cdot N = \frac{\tilde{u}'P_x'P_x\tilde{u}}{\tilde{u}'\tilde{u}} \cdot N = \frac{(P_x\tilde{u})'(P_x\tilde{u})}{\tilde{u}'\tilde{u}} \cdot N$$

Note that $P_x\tilde{u}$ are the fitted values ($\hat{\tilde{u}}$) of the auxiliary regression

$$\tilde{u} = x\gamma + v$$

since $\hat{\tilde{u}} = x\hat{\gamma} = x(x'x)^{-1}x'\tilde{u} = P_x\tilde{u}$.

This yields

$$LM = \frac{\hat{\tilde{u}}'\hat{\tilde{u}}}{\tilde{u}'\tilde{u}} \cdot N = N \cdot R^2$$

where R^2 is the uncentered R^2 of the auxiliary regression.
The centered R^2 in this case would be

$$\frac{\hat{\tilde{u}}'\hat{\tilde{u}}}{\tilde{u}'\tilde{u} - N(\bar{\tilde{u}})^2}$$