ST 793 : Solution of Homework-2

Salil Koner

September 2019

1

1.1 Problem 2.34

(a) We have given that

$$\mathbf{I}\left(\mu,\sigma\right) = \frac{1}{\sigma^2}\begin{bmatrix} 1 & -0.423 \\ -0.423 & 1.824 \end{bmatrix}$$

We know that the asymptotic covariance matrix of MLE based on a random sample of size n is given by

$$\text{Cov}\left(\hat{\mu},\hat{\sigma}\right) = \frac{1}{n}\left[\mathbf{I}\left(\mu,\sigma\right)\right]^{-1} = \frac{\sigma^2}{n}\begin{bmatrix} 1 & -0.423 \\ -0.423 & 1.824 \end{bmatrix}^{-1} = \frac{\sigma^2}{n}\begin{bmatrix} 1.1088 & 0.2571 \\ 0.2571 & 0.6079 \end{bmatrix}$$

In homework-1, we obtained the estimate of (μ,σ) as $\hat{\mu} = 4395.145, \hat{\sigma} = 1882.495$. So, an estimate of the asymptotic covariance of $(\hat{\mu},\hat{\sigma})$ is given by

$$\hat{\text{Cov}}\left(\hat{\mu},\hat{\sigma}\right) = \frac{\hat{\sigma}^2}{35}\begin{bmatrix} 1.1088 & 0.2571 \\ 0.2571 & 0.6079 \end{bmatrix} = \begin{bmatrix} 112,263.82 & 26,034.87 \\ & 61,548.15 \end{bmatrix}$$

(b) Estimate of the median of largest flow rate in 100 years is

$$\hat{Q} = 1882.495\left[-\log_e\left(-\log_e\left(0.993\right)\right)\right] + 4395.145 = 13,729.19$$

Because,

$$\hat{Q} = \begin{pmatrix} 1 & -\log_e\left(-\log_e\left(0.993\right)\right) \end{pmatrix}\begin{pmatrix} \hat{\mu} \\ \hat{\sigma} \end{pmatrix} = \begin{pmatrix} 1 & 4.9583 \end{pmatrix}\begin{pmatrix} \hat{\mu} \\ \hat{\sigma} \end{pmatrix}$$

(c) An estimate for the asymptotic variance of $\hat{Q}$ is

$$\begin{aligned} \text{Var}(\hat{Q}) &= \begin{pmatrix} 1 & 4.9583 \end{pmatrix}\hat{\text{Cov}}\left(\hat{\mu},\hat{\sigma}\right)\begin{pmatrix} 1 & 4.9583 \end{pmatrix}^\top \\ &= \begin{pmatrix} 1 & 4.9583 \end{pmatrix}\begin{bmatrix} 112,263.82 & 26,034.87 \\ & 61,548.15 \end{bmatrix}\begin{pmatrix} 1 & 4.9583 \end{pmatrix}^\top \\ &= 1,883,609 \end{aligned}$$

1.2 Problem 2.49

(a) Noticing that $(Y_i,Z_i)_{i=1}^n$ is the complete data in this situation, the complete data likelihood function is given by

$$\begin{aligned} L_c\left(\lambda,p|\left(Y_i,Z_i\right)_{i=1}^n\right) &= \Pi_{i=1}^n f\left(Y_i,Z_i|\lambda,p\right) \\ &= \Pi_{i=1}^n\left[pf_1\left(Y_i\right)\right]^{Z_i}\left[(1-p)f_2\left(Y_i\,;\lambda\right)\right]^{1-Z_i} \\ &= p^{\sum_{i=1}^n Z_i}(1-p)^{n-\sum_{i=1}^n Z_i}e^{-\lambda\sum_{i=1}^n(1-Z_i)}\lambda^{\sum_{i=1}^n Y_i(1-Z_i)}\frac{1}{\Pi_{i=1}^n Y_i!^{1-Z_i}} \end{aligned}$$

So, the complete data log-likelihood function is given by

$$\ell_c\left(\lambda, p \middle| \left(Y_i, Z_i\right)_{i=1}^n\right) = \left(\sum_{i=1}^n Z_i\right)\log_e p + \left(n - \sum_{i=1}^n Z_i\right)\log_e(1-p) - \left(n - \sum_{i=1}^n Z_i\right)\lambda$$
$$+ \left(\sum_{i=1}^n Y_i\left(1 - Z_i\right)\right)\log_e \lambda - \sum_{i=1}^n \log_e(Y_i!)\left(1 - Z_i\right)$$

Taking conditional expectation of $\ell_c\left(\lambda, p \middle| \left(Y_i, Z_i\right)_{i=1}^n\right)$ given $\left(Y_i\right)_{i=1}^n, \lambda^\nu, p^\nu$, we get

$$\begin{aligned}
Q\left(\lambda, p | \lambda^\nu, p^\nu\right) &= \mathrm{E}\left[\ell_c\left(\lambda, p\right) | \left(Y_i\right)_{i=1}^n, \lambda^\nu, p^\nu\right] \\
&= \left(\sum_{i=1}^n \mathrm{E}\left(Z_i | Y_i\right)\right)\log_e p + \left(n - \sum_{i=1}^n \mathrm{E}\left(Z_i | Y_i\right)\right)\log_e(1-p) - \left(n - \sum_{i=1}^n \mathrm{E}\left(Z_i | Y_i\right)\right)\lambda \\
&\quad + \left(\sum_{i=1}^n Y_i\left(1 - \mathrm{E}\left(Z_i | Y_i\right)\right)\right)\log_e \lambda - \sum_{i=1}^n \log_e(Y_i!)\left(1 - \mathrm{E}\left(Z_i | Y_i\right)\right) \\
&= \left(\sum_{i=1}^n w_i^\nu\right)\log_e p + \left(n - \sum_{i=1}^n w_i^\nu\right)\log_e(1-p) - \left(n - \sum_{i=1}^n w_i^\nu\right)\lambda \\
&\quad + \left(\sum_{i=1}^n Y_i\left(1 - w_i^\nu\right)\right)\log_e \lambda - \sum_{i=1}^n \log_e(Y_i!)\left(1 - w_i^\nu\right)
\end{aligned}$$

(b) Differentiating $Q\left(\lambda, p | \lambda^\nu, p^\nu\right)$ with respect to λ and p we get,

$$0 = \frac{\partial Q}{\partial p} = \left(\sum_{i=1}^n w_i^\nu\right)\frac{1}{p} - \left(n - \sum_{i=1}^n w_i^\nu\right)\frac{1}{1-p} \implies p^{\nu+1} = \frac{1}{n}\sum_{i=1}^n w_i^\nu \tag{1}$$

and

$$0 = \frac{\partial Q}{\partial \lambda} = -\left(n - \sum_{i=1}^n w_i^\nu\right) + \left(\sum_{i=1}^n Y_i\left(1 - w_i^\nu\right)\right)\frac{1}{\lambda} \implies \lambda^{\nu+1} = \frac{\sum_{i=1}^n Y_i\left(1 - w_i^\nu\right)}{\sum_{i=1}^n\left(1 - w_i^\nu\right)} \tag{2}$$

To make sure the maximization, we need to compute the Hessian and check whether it is negative definite at $(p, \lambda) = \left(p^{\nu+1}, \lambda^{\nu+1}\right)$.

Notice that,

$$\frac{\partial^2 Q}{\partial p^2}\Big|_{p=p^{\nu+1}} = -\left(\sum_{i=1}^n w_i^\nu\right)\frac{1}{\left(p^{\nu+1}\right)^2} - \left(n - \sum_{i=1}^n w_i^\nu\right)\frac{1}{\left(1 - p^{\nu+1}\right)^2} = -\frac{n^2}{\left(\sum_{i=1}^n w_i^\nu\right)\left(n - \sum_{i=1}^n w_i^\nu\right)} < 0$$

because, w_i^ν is conditional expectation of a random variable (Z_i) that lies between 0 and 1, which implies, $0 \le w_i^\nu \le 1$ for all $i = 1, 2 \ldots, n$. and

$$\frac{\partial^2 Q}{\partial \lambda^2}\Big|_{\lambda=\lambda^{\nu+1}} = -\left(\sum_{i=1}^n Y_i\left(1 - w_i^\nu\right)\right)\frac{1}{\left(\lambda^{\nu+1}\right)^2} \le 0 \qquad \text{as } Y_i \ge 0 \text{ and } 0 \le w_i^\nu \le 1$$

and as the likelihood is completely separable with respect to λ and p,

$$\frac{\partial^2 Q}{\partial p \partial \lambda} = 0$$

This implies that the Hessian in negative definite and the update of $p^{\nu+1}$ and $\lambda^{\nu+1}$ are as in (1) and (2) respectively.

1.3 Problem 2.51

(a) Since we only observed $|Y_i|, i = q+1, \dots, n$, to obtain the observed the data likelihood, we need to find the distribution of $|Y|$ when $Y \sim \text{Normal}(\mu, \sigma^2)$

$$\begin{aligned} P(|Y| \le y) &= P(-y \le Y \le y) \\ &= P(Y \le y) - P(Y \le -y) \\ &= \Phi\left(\frac{y-\mu}{\sigma}\right) - \Phi\left(\frac{-y-\mu}{\sigma}\right) \quad y > 0 \end{aligned}$$

After differentiating with respect to y, we get the density of $|Y|$ is given by

$$f_{|Y|}\left(y|\mu, \sigma^2\right) = \frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right) + \frac{1}{\sigma}\phi\left(\frac{y+\mu}{\sigma}\right) \qquad [\text{because } \ \phi(x) = \phi(-x) \ \ \forall\ x]$$

where $\phi(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right)$ is the density of standard normal distribution. If Y_i and Y_j are independently distributed as Y_i and $|Y_j|$. So, the observed data likelihood is given by

$$\begin{aligned} L\left(\mu, \sigma^2 | Y_1, Y_2, \dots, Y_q, |Y_{q+1}|, \dots, |Y_n|\right) &= \Pi_{i=1}^q f_{Y_i}(y_i) \Pi_{i=q+1}^n f_{|Y_i|}(y_i) \\ &= \Pi_{i=1}^q \left(\frac{1}{\sigma}\phi\left(\frac{y_i-\mu}{\sigma}\right)\right) I(y_i \in \mathbb{R}) \quad \times \\ &\quad \Pi_{i=q+1}^n \left(\frac{1}{\sigma}\phi\left(\frac{y_i-\mu}{\sigma}\right) + \frac{1}{\sigma}\phi\left(\frac{y_i+\mu}{\sigma}\right)\right) I(y_i > 0) \end{aligned}$$

(b) If $Z_i = 2I(Y_i > 0) - 1$, then

$$Z_i = \begin{cases} 1 & \text{if } \ Y_i > 0 \\ -1 & \text{if } \ Y_i \le 0 \end{cases}$$

This implies that Z_i carries the sign of Y_i and $P\left(Z_i^2 = 1\right) = 1$ for all $i = (q+1), \dots, n$. Thus,

$$Y_i = |Y_i|\, Z_i \qquad \forall\ i = q+1, \dots, n$$

Thus, if we would have the information on $Z_i, i = q+1, \dots, n$, then the complete data would have been available to us. So, considering $Y_1, Y_2, \dots, Y_q, |Y_{q+1}|, \dots, |Y_n|, Z_{q+1}, Z_{q+2}, \dots, Z_n$ as our complete data, the complete data likelihood is given by

$$\begin{aligned} L_c\left(\mu, \sigma^2 | \{Y_i\}_{i=1}^n, \{|Y_i|\}_{i=q+1}^n, \{Z_i\}_{i=q+1}^n\right) &= \Pi_{i=1}^n f_{Y_i}(y_i|\mu, \sigma^2) \\ &= \left(2\pi\sigma^2\right)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^q (Y_i - \mu)^2\right) \exp\left(-\frac{1}{2\sigma^2}\sum_{i=q+1}^n \left(|Y_i|\, Z_i - \mu\right)^2\right) \end{aligned}$$

Complete data log-likelihood is given by

$$\begin{aligned} \ell_c\left(\mu, \sigma^2\right) &= -\frac{n}{2}\left(\log_e 2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\left(\sum_{i=1}^q (Y_i - \mu)^2 + \sum_{i=q+1}^n \left(|Y_i|\, Z_i - \mu\right)^2\right) \\ &= -\frac{n}{2}\left(\log_e 2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\left[\left(\sum_{i=1}^q Y_i^2 - 2\mu\sum_{i=1}^q Y_i + q\mu^2\right) + \left(\sum_{i=q+1}^n |Y_i|^2 Z_i^2 - 2\mu\sum_{i=q+1}^n |Y_i|\, Z_i + (n-q)\mu^2\right)\right] \\ &= -\frac{n}{2}\left(\log_e 2\pi\sigma^2\right) - \frac{1}{2\sigma^2}\left(\sum_{i=1}^n Y_i^2 - 2\mu\sum_{i=1}^q Y_i + n\mu^2 - 2\mu\sum_{i=q+1}^n |Y_i|\, Z_i\right) \quad \left[\text{because } \ P\left(Z_i^2 = 1\right) = 1\right] \end{aligned}$$

(c) Because, $Z_i = 2I(Y_i > 0) - 1$, $\mathrm{E}\left(Z_i \middle| \left|Y_i\right|, \mu, \sigma\right) = 2w_i\left(\mu, \sigma\right) - 1$. Then the E-step is given by

$$\begin{aligned} Q\left(\mu, \sigma^2 | \mu^\nu, \sigma^\nu\right) &= \mathrm{E}\left(\ell_c\left(\mu, \sigma^2\right) \middle| \{Y_i\}_{i=1}^n, \{|Y_i|\}_{i=q+1}^n, \mu^\nu, \sigma^\nu\right) \\ &= -\frac{n}{2}\left(\log_e \sigma^2 + \log_e 2\pi\right) - \frac{1}{2\sigma^2}\left(\sum_{i=1}^n Y_i^2 - 2\mu\sum_{i=1}^q Y_i + n\mu^2 - 2\mu\sum_{i=q+1}^n |Y_i|\,\mathrm{E}\left(Z_i \middle| |Y_i|, \mu^\nu, \sigma^\nu\right)\right) \\ &= -\frac{n}{2}\left(\log_e \sigma^2 + \log_e 2\pi\right) - \frac{1}{2\sigma^2}\left(\sum_{i=1}^n Y_i^2 - 2\mu\sum_{i=1}^q Y_i + n\mu^2 - 2\mu\sum_{i=q+1}^n |Y_i|\left(2w_i\left(\mu^\nu, \sigma^\nu\right) - 1\right)\right) \end{aligned}$$

(d) Differentiating $Q\left(\mu, \sigma^2 | \mu^\nu, \sigma^\nu\right)$ with respect to μ we get,

$$\frac{\partial Q\left(\mu, \sigma^2 | \mu^\nu, \sigma^\nu\right)}{\partial \mu} = -\frac{1}{2\sigma^2}\left(-2\sum_{i=1}^q Y_i + 2n\mu - 2\sum_{i=q+1}^n |Y_i|\left(2w_i\left(\mu^\nu, \sigma^\nu\right) - 1\right)\right) = 0$$

This gives the update formula for μ as

$$\mu^{\nu+1} = \frac{1}{n}\left(\sum_{i=1}^q Y_i + \sum_{i=q+1}^n |Y_i|\left(2w_i\left(\mu^\nu, \sigma^\nu\right) - 1\right)\right) \tag{3}$$

To make sure that we reach towards the maximizer, the second derivative,

$$\frac{\partial^2 Q\left(\mu, \sigma^2 | \mu^\nu, \sigma^\nu\right)}{\partial^2 \mu} = -\frac{n}{2\sigma^2} < 0$$

So, the update formula for μ is as in (3).

(e) Note that, by Bayes theorem,

$$\begin{aligned} w_i\left(\mu, \sigma\right) = P(Y_i > 0 \,|\, |Y_i| = y_i) &= \frac{f_{|Y_i|}(y_i | Y_i > 0)P(Y_i > 0)}{f_{|Y_i|}(y_i | Y_i > 0)P(Y_i > 0) + f_{|Y_i|}(y_i | Y_i \le 0)P(Y_i \le 0)} \\ &= \frac{f_{|Y_i|}(y_i | Y_i > 0)\Phi\left(\frac{\mu}{\sigma}\right)}{f_{|Y_i|}(y_i | Y_i > 0)\Phi\left(\frac{\mu}{\sigma}\right) + f_{|Y_i|}(y_i | Y_i \le 0)\Phi\left(-\frac{\mu}{\sigma}\right)} \end{aligned} \tag{4}$$

Now, CDF of $|Y_i| \,|Y_i > 0$ is

$$F_{|Y_i|}(y | Y_i > 0) = \begin{cases} 0 & \text{if } \ y < 0 \\ \frac{P(0 < Y_i \le y)}{P(Y_i > 0)} = \frac{\Phi\left(\frac{y-\mu}{\sigma}\right) - \Phi\left(-\frac{\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} & \text{if } \ y \ge 0 \end{cases}$$

So, the density of $|Y_i| \,|Y_i > 0$ is

$$f_{|Y_i|}(y | Y_i > 0) = \frac{\frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)}{\Phi\left(\frac{\mu}{\sigma}\right)} I(y \ge 0) \tag{5}$$

Similarly, we can show that the density of $|Y_i| \,|Y_i \le 0$ is

$$f_{|Y_i|}(y | Y_i \le 0) = \frac{\frac{1}{\sigma}\phi\left(\frac{y+\mu}{\sigma}\right)}{\Phi\left(-\frac{\mu}{\sigma}\right)} I(y \ge 0) \tag{6}$$

Putting (5) and (6) in (4) we get,

$$\begin{aligned} w_i\left(\mu, \sigma\right) &= \frac{\frac{1}{\sigma}\phi\left(\frac{y_i-\mu}{\sigma}\right)}{\frac{1}{\sigma}\phi\left(\frac{y_i-\mu}{\sigma}\right) + \frac{1}{\sigma}\phi\left(\frac{y_i+\mu}{\sigma}\right)} \\ &= \left(1 + \exp\left(-\frac{2\mu y_i}{\sigma^2}\right)\right)^{-1} \end{aligned}$$

which is the simplified expression for $w_i\left(\mu, \sigma\right)$.

1.4 Problem 2.54

(a) Since, Y_i's are iid, then the chance of Y_i lying in any of the intervals is same for all Y_i and they are independent. So, this fits an appropriate set-up of multinomial distribution with

$$p_i = P\left(Y_1 \in [a_{i-1}, a_i)\right) = F(a_i) - F(a_{i-1}) \quad i = 1, 2 \ldots, k$$

Assuming that the support of the distribution of Y_i is $[a_0, a_k)$, note that, $\sum_{i=1}^k p_i = 1$. Then, $N_1, N_2, \ldots, N_k$ follows multinomial $(n, p_1, p_2, \ldots, p_k)$. As our only observed data is $N_1, N_2, \ldots, N_k$, the observed data likelihood is given by

$$L\left(p_1, p_2, \ldots, p_k | N_1, N_2, \ldots, N_k\right) = \begin{cases} \frac{n!}{N_1! N_2! \ldots, N_k!} p_1^{N_1} p_2^{N_2}, \ldots, p_k^{N_k} & \text{if } \sum_{i=1}^k N_i = n \\ 0 & \text{otherwise} \end{cases}$$

(b) Our complete data here is $Y_1, Y_2, \ldots, Y_n$ and they are iid with density $f(y|\theta)$. So, the complete data log-likelihood is given by

$$\ell_c(\theta | y_1, y_2, \ldots, y_n) = \log_e L_c(\theta | Y_1, Y_2, \ldots, Y_n) = \log_e \Pi_{i=1}^n f_{Y_i}\left(y_i | \theta\right) = \sum_{i=1}^n \log_e f_{Y_i}\left(y_i | \theta\right)$$

$$\begin{aligned} P\left(Y_1 \leq y | Y_1 \in [a_{i-1}, a_i)\right) &= \begin{cases} 0 & \text{if } y < a_{i-1} \\ \frac{P(a_{i-1} \leq Y_1 \leq y)}{P(a_{i-1} \leq Y_1 \leq a_i)} & \text{if } a_{i-1} \leq y < a_i \\ 1 & \text{if } y \geq a_i \end{cases} \\ &= \begin{cases} 0 & \text{if } y < a_{i-1} \\ \frac{F_\theta(y) - F_\theta(a_{i-1})}{F_\theta(a_i) - F_\theta(a_{i-1})} & \text{if } a_{i-1} \leq y < a_i \\ 1 & \text{if } y \geq a_i \end{cases} \end{aligned}$$

where $F_\theta(x) = \int_{-\infty}^x f(y;\theta) dy$ is the CDF of Y_1. Differentiating the above CDF with respect to y, we get the density of Y_1 given $Y_1 \in [a_{i-1}, a_i)$ as

$$f_{Y_1}\left(y | Y_1 \in [a_{i-1}, a_i)\right) = \frac{f_\theta(y)}{F_\theta(a_i) - F_\theta(a_{i-1})} I\left(a_{i-1} \leq y < a_i\right)$$

2

The log-likelihood function of this location-scale family based on a single observation is given by

$$\ell(\mu, \sigma\,; y) = -\log_e \sigma + \log_e f_0\left(\frac{y-\mu}{\sigma}\right)$$

This implies,

$$\frac{\partial \ell}{\partial \mu} = -\frac{1}{\sigma} \frac{f_0'\left(\frac{y-\mu}{\sigma}\right)}{f_0\left(\frac{y-\mu}{\sigma}\right)} \qquad \frac{\partial \ell}{\partial \mu} = -\frac{1}{\sigma}\left[1 + \frac{f_0'\left(\frac{y-\mu}{\sigma}\right)}{f_0\left(\frac{y-\mu}{\sigma}\right)}\left(\frac{y-\mu}{\sigma}\right)\right]$$

This implies,

$$\begin{aligned} \mathrm{E}\left(\frac{\partial \ell}{\partial \mu}\right)^2 &= \int_{-\infty}^{\infty} \frac{1}{\sigma^2}\left(\frac{f_0'\left(\frac{y-\mu}{\sigma}\right)}{f_0\left(\frac{y-\mu}{\sigma}\right)}\right)^2 \frac{1}{\sigma} f_0\left(\frac{y-\mu}{\sigma}\right) dy \\ &= \frac{1}{\sigma^2} \int_{-\infty}^{\infty} \left[\frac{f_0'(x)}{f_0(x)}\right]^2 f_0(x) dx \end{aligned}$$

$$
\begin{aligned}
\mathrm{E}\left(\frac{\partial \ell}{\partial \sigma}\right)^2 &= \int_{-\infty}^{\infty} \frac{1}{\sigma^2}\left[1+\frac{f_0'\left(\frac{y-\mu}{\sigma}\right)}{f_0\left(\frac{y-\mu}{\sigma}\right)}\left(\frac{y-\mu}{\sigma}\right)\right]^2 \frac{1}{\sigma} f_0\left(\frac{y-\mu}{\sigma}\right) dy \\
&= \frac{1}{\sigma^2}\int_{-\infty}^{\infty}\left[1+x\frac{f_0'(x)}{f_0(x)}\right]^2 f_0(x)dx
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}\left[\left(\frac{\partial \ell}{\partial \mu}\right)\left(\frac{\partial \ell}{\partial \sigma}\right)\right] &= \int_{-\infty}^{\infty} \frac{1}{\sigma^2}\left(\frac{f_0'\left(\frac{y-\mu}{\sigma}\right)}{f_0\left(\frac{y-\mu}{\sigma}\right)}\right)^2\left(\frac{y-\mu}{\sigma}\right)\frac{1}{\sigma} f_0\left(\frac{y-\mu}{\sigma}\right) dy + \int_{-\infty}^{\infty} \frac{1}{\sigma^2}\left[\frac{f_0'\left(\frac{y-\mu}{\sigma}\right)}{f_0\left(\frac{y-\mu}{\sigma}\right)}\right]\frac{1}{\sigma} f_0\left(\frac{y-\mu}{\sigma}\right) dy \\
&= \frac{1}{\sigma^2}\left[\int_{-\infty}^{\infty} x\left[\frac{f_0'(x)}{f_0(x)}\right]^2 f_0(x)dx + \int_{-\infty}^{\infty} f_0'(x)dx\right] \\
&= \frac{1}{\sigma^2}\left[\int_{-\infty}^{\infty} x\left[\frac{f_0'(x)}{f_0(x)}\right]^2 f_0(x)dx + \int_{-\infty}^{\infty} \frac{d}{dx} f_0(x)dx\right] \\
&= \frac{1}{\sigma^2}\left[\int_{-\infty}^{\infty} x\left[\frac{f_0'(x)}{f_0(x)}\right]^2 f_0(x)dx + \frac{d}{dx}\int_{-\infty}^{\infty} f_0(x)dx\right] \quad \text{[provided this interchange is valid for } f_0\text{]} \\
&= \frac{1}{\sigma^2}\left[\int_{-\infty}^{\infty} x\left[\frac{f_0'(x)}{f_0(x)}\right]^2 f_0(x)dx + \frac{d}{dx}\int_{-\infty}^{\infty} 1\right] \\
&= \frac{1}{\sigma^2}\int_{-\infty}^{\infty} x\left[\frac{f_0'(x)}{f_0(x)}\right]^2 f_0(x)dx
\end{aligned}
$$

The information matrix is given by

$$
\mathbf{I}(\mu,\sigma) = \frac{1}{\sigma^2}\begin{bmatrix} \int_{-\infty}^{\infty}\left[\frac{f_0'(x)}{f_0(x)}\right]^2 f_0(x)dx & \int_{-\infty}^{\infty} x\left[\frac{f_0'(x)}{f_0(x)}\right]^2 f_0(x)dx \\ & \int_{-\infty}^{\infty}\left[1+x\frac{f_0'(x)}{f_0(x)}\right]^2 f_0(x)dx \end{bmatrix}
$$

Notice, that, if the base density $f_0(x)$ is symmetric about 0, then $f_0(x) = f_0(-x)$, then f_0 is an even function, which means f_0' is an odd function. This implies, the integrand in the off-diagonal element of $\mathbf{I}(\mu,\sigma)$, $g(x) = x\left[\frac{f_0'(x)}{f_0(x)}\right]^2 f_0(x)$ is an odd function in x, resulting $\int_{-\infty}^{\infty} g(x)dx = 0$. Then, the information matrix would be diagonal. Otherwise, for general function f_0, the information matrix is a dense matrix.

Thus, we can say that if the base density is symmetric then there is no asymptotic variance inflation for estimating μ when σ is unknown. Otherwise, the variance would be inflated if we want to estimate μ assuming σ unknown compared to the case when σ is known.