MLE of Logistic Distribution

November 1, 2021

In this note, we derive asymptotic normality of MLE $\hat{\eta} = (\hat{\theta}, \hat{\sigma})$ of $\eta = (\theta, \sigma)$ for Logistic(θ,σ), where $\theta \in \mathbb{R}$ and $\sigma > 0$. Recall the following regularity conditions.

- **(R0)** The parameter θ is identifiable in Θ.
- **(R1)** The density $f(\cdot;\theta)$ have common support $\mathfrak{X}$.
- **(R2)** The parameter space is open in $\mathbb{R}^d$.
- **(R3)** The log-density pdf $\log f(x;\theta)$ is twice differentable as a function of θ for all $x \in \mathfrak{X}$.
- **(R4)** For any statistic $u(X_1, \dots, X_n)$ with finite expectation, the integral

  $$\mathrm{E}_\theta(u(X_1,\dots,X_n)) = \int_{\mathfrak{X}^n} u(x_1,\dots,x_n) \prod_{i=1}^n f(x_i;\theta) \prod_{i=1}^n dx_i$$

  is twice differentiable under the integral sign.
- **(R5)** The Fisher Information $I(\theta)$ exists and is invertible for all $\theta \in \Theta$.
- **(R6)** The likelihood equation $\dot{l}(\theta) = 0$ has the unique solution $\hat{\theta}$ and the solution is a consistent estimator of θ.
- **(R7)** For all $\theta \in \Theta$, there exists a function $M(\cdot)$ with $\mathrm{E}_\theta M(X_1) < \infty$ such that

  $$\max_{\theta\in\Theta} \max_{1\le h,i,j\le d} |\frac{\partial^3}{\partial\theta_h \partial\theta_j \partial\theta_j} \log f(X_1;\theta)| \le M(X_1), \mathrm{E}_\theta M(X_1) < \infty.$$

We verify each condition and derive asymptotic normality of $\hat{\eta}$. For convenience, we consider $\eta^* = (r,t)$ instead of η, where $r = \theta/\sigma$ and $t = 1/\sigma$. This transformation makes the computation of likelihood much easier. Note that the map $g : \mathbb{R} \times \mathbb{R}_+ \mapsto \mathbb{R} \times \mathbb{R}_+$ defined by $g(a,b) = (a/b, 1/b)$ is bijection. Hence, if the existence and uniqueness of of $\hat{\eta^*}$ of η^* hold, then the same hold for $\hat{\eta}$. Furthermore, if we find the asymptotic distribution of $\sqrt{n}(\hat{\eta^*} - \eta^*)$, we can also find that of $\sqrt{n}(\hat{\eta} - \eta)$ by $\Delta-$method.

(R0) Note that parameter space $\Theta = \mathbb{R} \times \mathbb{R}_+$. Suppose $\eta_1^* = (r_1, t_1), \eta_2^* = (r_2, t_2) \in \Theta$. Let $f(\cdot;\eta_i^*)$ be p.d.f. of Logistic($r_i/t_i, 1/t_i$) for $i = 1, 2$. Assume

$$f(x;\eta_1^*) = f(x;\eta_2^*) \tag{1}$$

for all $x \in \mathfrak{X} = \mathbb{R}$. Recall that if $Y \sim$ Logistic$(0,1)$, then $\mathbb{E}Y = 0$ and $\mathrm{Var}(Y) = \pi^2/3$. Hence if random variable Y_i follows the distribution with p.d.f. $f(\cdot;\eta_i^*)$ for $i = 1,2$, implies $\mathbb{E}Y_1 = \mathbb{E}Y_2$ and $\mathrm{Var}(Y_1) = \mathrm{Var}Y_2$ so that

$$r_1/t_1 = r_2/t_2, \tag{2}$$

$$\frac{\pi^2}{3}\frac{1}{t_1^2} = \frac{\pi^2}{3}\frac{1}{t_2^2}. \tag{3}$$

(3) implies $t_1 = t_2$ as $t_1, t_2 > 0$ and substituting $t_1 = t_2$ into (2) gives $r_1 = r_2$. Thus, $\eta_1^* = \eta_2^*$, which implies the identifiability of η^* in Θ.

(R1) The density $f(\cdot;\eta^*)$ has support $\mathfrak{X} = \mathbb{R}$, which does not depend on η^*.

(R2) $\Theta = \mathbb{R} \times \mathbb{R}_+$ is open in $\mathbb{R}^2$.

(R3) Log-likelihood of η^* is given as following:

$$\ell(\eta^*) = \sum_{i=1}^{n} \log f(X_i;\eta^*) = \sum_{i=1}^{n} \log t \frac{\exp(tX_i - r)}{(1+\exp(tX_i - r))^2}$$
$$= n\log t + \sum_{i=1}^{n}(tX_i - r) - 2\sum_{i=1}^{n}\log(1+\exp(tX_i - r)).$$

This gives following partial derivatives of $\ell(\eta^*)$:

$$\frac{\partial \ell}{\partial t} = \frac{n}{t} - \sum_{i=1}^{n} X_i - 2\sum_{i=1}^{n}\frac{\exp(tX_i - r)}{1+\exp(tX_i - r)}X_i, \tag{4}$$

$$\frac{\partial \ell}{\partial r} = -n + 2\sum_{i=1}^{n}\frac{\exp(tX_i - r)}{1+\exp(tX_i - r)}, \tag{5}$$

$$\frac{\partial^2 \ell}{\partial t^2} = -\frac{n}{t^2} - 2\sum_{i=1}^{n}\frac{\exp(tX_i - r)}{(1+\exp(tX_i - r))^2}X_i^2$$
$$= -\frac{n}{t^2} - 2\sum_{i=1}^{n}(\frac{1}{1+\exp(tX_i - r)} - \frac{1}{(1+\exp(tX_i - r))^2})X_i^2, \tag{6}$$

$$\frac{\partial^2 \ell}{\partial t \partial r} = \frac{\partial^2 \ell}{\partial r \partial t} = 2\sum_{i=1}^{n}\frac{\exp(tX_i - r)}{(1+\exp(tX_i - r))^2}X_i$$
$$= 2\sum_{i=1}^{n}(\frac{1}{1+\exp(tX_i - r)} - \frac{1}{(1+\exp(tX_i - r))^2})X_i, \tag{7}$$

$$\frac{\partial^2 \ell}{\partial r^2} = -2\sum_{i=1}^{n}\frac{\exp(tX_i - r)}{(1+\exp(tX_i - r))^2}$$
$$= -2\sum_{i=1}^{n}(\frac{1}{1+\exp(tX_i - r)} - \frac{1}{(1+\exp(tX_i - r))^2}). \tag{8}$$

One can clearly see that each derivative is continuous.

(R4) This holds because $\prod_{i=1}^{n} f(x_i;\eta^*)$ is very smooth, i.e., has derivative of all order and each derivative is continuous. Thus, by Leibniz rule, **(R4)** holds.

(R5) Write

$$I(\eta^*) = \begin{pmatrix} I_{11} & I_{12} \\ I_{21} & I_{22} \end{pmatrix}.$$

We calculate I_{ij}, where $1 \le i, j \le 2$. In the calculation, we may use the fact that

$$\int_{\mathbb{R}} x^2 \frac{\exp(2x)}{(1+\exp(x))^4}dx = \frac{\pi^2}{18} - \frac{1}{3}.$$

Set $n = 1$ in (6)-(8). Then,

$$
\begin{aligned}
I_{11} = -\mathbb{E}_{\eta^*}\frac{\partial^2 \ell}{\partial r^2} &= 2\int_{\mathbb{R}} t\left(\frac{\exp(tx-r)}{(1+\exp(tx-r))^2}\right)^2 dx \\
&= 2\int_{\mathbb{R}}\left(\frac{\exp(y)}{(1+\exp(y))^2}\right)^2 dy \\
&= 2\int_1^{\infty}\frac{t-1}{t^4}dt = 2\left[-\frac{1}{2t^2}+\frac{1}{3t^3}\right]_1^{\infty} = \frac{1}{3}, \\
I_{12} = I_{21} = -\mathbb{E}_{\eta^*}\frac{\partial^2 \ell}{\partial r \partial t} &= -2\int_{\mathbb{R}} tx\left(\frac{\exp(tx-r)}{(1+\exp(tx-r))^2}\right)^2 dx \\
&= -2\int_{\mathbb{R}}\frac{y+r}{t}\frac{\exp(2y)}{(1+\exp(y))^4}dy \\
&= -\frac{2}{t}\int_{\mathbb{R}} y\frac{\exp(2y)}{(1+\exp(y))^4}dy - 2\frac{r}{t}\int_{\mathbb{R}}\frac{\exp(2y)}{(1+\exp(y))^4}dy \\
&= -\frac{r}{t}\cdot\frac{1}{3} = -\frac{r}{3t} \\
I_{22} = -\mathbb{E}_{\eta^*}\frac{\partial^2 \ell}{\partial t^2} &= \frac{1}{t^2} + 2\int_{\mathbb{R}} tx^2\left(\frac{\exp(tx-r)}{(1+\exp(tx-r))^2}\right)^2 dx \\
&= \frac{1}{t^2} + 2\int_{\mathbb{R}}\left(\frac{y+r}{t}\right)^2\frac{\exp(2y)}{(1+\exp(y))^4}dy \\
&= \frac{1}{t^2} + \frac{2}{t^2}\left(\int_{\mathbb{R}} y^2\frac{\exp(2y)}{(1+\exp(y))^4}dy + 2r\int_{\mathbb{R}} y\frac{\exp(2y)}{(1+\exp(y))^4}dy + r^2\int_{\mathbb{R}}\frac{\exp(2y)}{(1+\exp(y))^4}dy\right) \\
&= \frac{1}{t^2} + \frac{2}{t^2}\left(\frac{\pi^2}{18} - \frac{1}{3} + \frac{1}{6}r^2\right).
\end{aligned}
$$

In the calculation, we used the fact that $f(y) = y\frac{exp(2y)}{(1+\exp(y))^4}$ is odd function of $\mathbb{R}$. Indeed, for all $y \in \mathbb{R}$,

$$
f(-y) = -y\frac{\exp(-2y)}{(1+\exp(-y))^4} = -y\frac{\exp(-2y)\exp(4y)}{(1+\exp(-y))^4\exp(4y)} = -y\frac{\exp(2y)}{(1+\exp(y))^4} = -f(y).
$$

Consequently,

$$
I(\eta^*) = \begin{pmatrix} \frac{1}{3} & -\frac{r}{3t} \\ -\frac{r}{3t} & \frac{1}{t^2} + \frac{2}{t^2}(\frac{\pi^2}{18} - \frac{1}{3} + \frac{1}{6}r^2) \end{pmatrix}. \tag{9}
$$

(R6) As one can see in **(R4)**, ℓ is twice continuously differentiable on Θ. Also,

$$
\lim_{r\to\infty}\ell(\eta^*) = -\infty, \tag{10}
$$

$$
\lim_{r\to-\infty}\ell(\eta^*) = -\infty, \tag{11}
$$

$$
\lim_{t\to\infty}\ell(\eta^*) = -\infty, \tag{12}
$$

$$
\lim_{t\to 0+}\ell(\eta^*) = -\infty. \tag{13}
$$

Thus,

$$
\lim_{\eta^*\to\partial\Theta}\ell(\eta^*) = -\infty.
$$

Now we show the negative definiteness of $\ddot{\ell}$. Take $c = (c_1, c_2)^t \in \mathbb{R}^2 \setminus \{0\}$. Let $A_i = 2\frac{\exp(tX_i - r)}{(1+\exp(tX_i-r))^2}$. Clearly, $A_i > 0$. Since

$$
\ddot{\ell}(\eta^*) = \begin{pmatrix} \frac{\partial^2 \ell}{\partial r^2} & \frac{\partial^2 \ell}{\partial r \partial t} \\ \frac{\partial^2 \ell}{\partial r \partial t} & \frac{\partial^2 \ell}{\partial t^2} \end{pmatrix},
$$

by (6)-(8),

$$\begin{aligned}
c^t\ddot{\ell}(\eta^*)c &= \frac{\partial^2\ell}{\partial r^2}c_1^2 + 2\frac{\partial^2\ell}{\partial r\partial t}c_1c_2 + \frac{\partial^2\ell}{\partial t^2}c_2^2 \\
&= -\sum_{i=1}^n A_ic_1^2 + 2\sum_{i=1}^n A_iX_ic_1c_2 - \frac{n}{t^2}c_2^2 - \sum_{i=1}^n A_iX_ic_2^2 \\
&= -\frac{n}{t^2}c_2^2 - \sum_{i=1}^n A_i(c_1 - X_ic_2)^2.
\end{aligned}$$

Because $\frac{n}{t^2}c_2^2$ and $\sum_{i=1}^n A_i(c_1 - X_ic_2)^2$ are both nonnegative as $A_i > 0$, $c^t\ddot{\ell}(\eta^*)c \le 0$ and the equality holds if and only if both two terms are 0. But this holds if and only if $c_2^2 = 0$ and $(c_1 - X_ic_2^2) = 0$ for $i = 1, 2, \ldots, n$ and this can be held only when $c_1, c_2 = 0$ which implies $c = 0$. As we chose c to be nonzero, we see that $c^t\ddot{\ell}(\eta^*)c < 0$ whenever $c \in \mathbb{R}^2 \setminus \{0\}$ and so $\ddot{\ell}$ is negative definite.

Put $n = 1$ and let $\eta_0^* = (r_0, t_0)$ be true parameter of η^*. Then,

$$\begin{aligned}
\mathbb{E}_{\eta_0^*}\log f(X_1;\eta^*) &= \log t - r + t\mathbb{E}_{\eta_0^*}X_i - 2\mathbb{E}_{\eta_0^*}\log(1+\exp(tX_1 - r)) \\
&= \log t - r + t\frac{r_0}{t_0} - 2\mathbb{E}_{\eta_0^*}\log(1+\exp(tX_1 - r)).
\end{aligned}$$

Therefore, $\mathbb{E}_{\eta_0^*}\log f(X_1;\eta^*)$ is continuous function with respect to η^*. Consequently, by Theorem 6.4.2 in the textbook, the MLE $\hat{\eta^*}$ of η^* exists and unique and is consistent.

(R7) Put $n = 1$. From (6)-(8), one can obtain followings:

$$\frac{\partial^3\ell}{\partial t^3} = \frac{2}{t^3} + 2\left(\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^2} - 2\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^3}\right)X_1^3, \tag{14}$$

$$\frac{\partial^3\ell}{\partial t\partial r^2} = \frac{\partial^3\ell}{\partial r^2\partial t} = 2\left(\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^2} - 2\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^3}\right)X_1, \tag{15}$$

$$\frac{\partial^3\ell}{\partial t^2\partial r} = \frac{\partial^3\ell}{\partial r\partial t^2} = 2\left(-\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^2} + 2\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^3})\right)X_1^2, \tag{16}$$

$$\frac{\partial^3\ell}{\partial r^3} = 2\left(-\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^2} + 2\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^3}\right). \tag{17}$$

Note that $0 < \frac{\exp(tX_1-r)}{(1+\exp(tX_1-r))^2}, \frac{\exp(tX_1-r)}{(1+\exp(tX_1-r))^3} < 1$. Thus, by (9)-(12),

$$\begin{aligned}
\left|\frac{\partial^3\ell}{\partial t^3}\right| &\le \frac{2}{t^3} + \left|2\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^2}X_1^3\right| + \left|4\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^3}X_1^3\right| \\
&\le \frac{2}{t^3} + 2|X_1|^3 + 4|X_1|^3 = \frac{2}{t^3} + 6|X_1|^3 = M_1(X_1),
\end{aligned} \tag{18}$$

$$\begin{aligned}
\left|\frac{\partial^3\ell}{\partial t\partial r^2}\right| = \left|\frac{\partial^3\ell}{\partial r^2\partial t}\right| &= \left|2\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^2}X_1\right| + \left|4\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^3}X_1\right| \\
&\le 2|X_1| + 4|X_1| = 6|X_1| = M_2(X_1),
\end{aligned} \tag{19}$$

$$\begin{aligned}
\frac{\partial^3\ell}{\partial t^2\partial r} = \left|\frac{\partial^3\ell}{\partial r\partial t^2}\right| &\le \left|2\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^2}X_1^2\right| + \left|4\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^3}X_1^2\right| \\
&\le 2|X_1|^2 + 4|X_1|^2 = 6|X_1|^2 = M_3(X_1),
\end{aligned} \tag{20}$$

$$\left|\frac{\partial^3\ell}{\partial r^3}\right| \le \left|2\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^2}\right| + \left|4\frac{\exp(tX_1 - r)}{(1+\exp(tX_1 - r))^3}\right| \le 2 + 4 = 6 = M_4(X_1). \tag{21}$$

It suffices to show $\mathbb{E}_\eta^* M_i(X_1) < \infty$ for $i = 1, 2, 3$, and 4. But by Lyapunov's inequality, this can be achieved by showing $\mathbb{E}_\eta^*|X_1|^3 < \infty$. This can be held if $\int_0^\infty \frac{x^3}{\exp(tx)}dx < \infty$ because

$$
\begin{aligned}
\mathbb{E}_\eta^* &= \int_{\mathbb{R}} |x|^3 t \frac{\exp(tx-r)}{(1+\exp(tx-r))^2} dx \\
&= t\int_0^\infty x^3 \frac{\exp(tx-r)}{(1+\exp(tx-r))^2} dx + t\int_{-\infty}^0 (-x)^3 \frac{\exp(tx-r)}{(1+\exp(tx-r))^2} dx \\
&= t\int_0^\infty x^3 \frac{\exp(tx-r)}{(1+\exp(tx-r))^2} dx + t\int_{-\infty}^0 x^3 \frac{\exp(tx+r)}{(1+\exp(tx+r))^2} dx \\
&\le t\exp(-r)\int_0^\infty x^3 \frac{\exp(tx)}{(\exp(tx-r))^2} dx + t\exp(r)\int_0^\infty x^3 \frac{\exp(tx)}{(\exp(tx+r))^2} dx \\
&=\le t\exp(r)\int_0^\infty \frac{x^3}{\exp(tx)} dx + t\exp(-r)\int_0^\infty \frac{x^3}{\exp(tx)} dx
\end{aligned}
$$

and $t>0$. But if $t, x > 0$, we see that $\exp(tx) = \sum_{i=0}^n (tx)^i/i! \ge (tx)^5/5! = C(tx)^5$ for some constant $C > 0$. This implies that

$$
\begin{aligned}
\int_0^\infty \frac{x^3}{\exp(tx)} dx &= \int_0^1 \frac{x^3}{\exp(tx)} dx + \int_1^\infty \frac{x^3}{\exp(tx)} dx \\
&\le \int_0^1 x^3 dx + \frac{1}{Ct^5}\int_1^\infty \frac{1}{x^2} dx \\
&= \frac{1}{4} + \frac{1}{Ct^5} < \infty,
\end{aligned}
$$

which establishes **(R7)**.

By **(R0)-(R7)**,

$$\sqrt{n}(\hat{\eta^*} - \eta^*) \xrightarrow{\text{d}} N(0, I(\eta^*)^{-1}).$$

Recall the map g defined in the page 1 and the existence and existence of $\hat{\eta^*}$ can be held due to this map. Consistency also holds because g is continuous on Θ so the consistency is direct from continuous mapping theorem. Let $\dot{g}$ be Jacobian matrix of g. Then,

$$\dot{g}(a,b) = \begin{pmatrix} \frac{1}{b} & -\frac{a}{b^2} \\ 0 & -\frac{1}{b^2} \end{pmatrix}.$$

By $\Delta-$method,

$$\sqrt{n}(g(\hat{\eta^*}) - g(\eta^*)) = \sqrt{n}(\hat{\eta} - \eta) \xrightarrow{\text{d}} N(0, \dot{g}(\eta^*)I(\eta^*)^{-1}\dot{g}(\eta^*)^t).$$

Let $V = \dot{g}(\eta^*)I(\eta^*)^{-1}\dot{g}(\eta^*)^t$. It remains to calculate to V.

$$
\begin{aligned}
V &= \begin{pmatrix} \frac{1}{t} & -\frac{r}{t^2} \\ 0 & -\frac{1}{t^2} \end{pmatrix} \begin{pmatrix} \frac{1}{3} & -\frac{r}{3t} \\ -\frac{r}{3t} & \frac{1}{t^2} + \frac{2}{t^2}(\frac{\pi^2}{18} - \frac{1}{3} + \frac{1}{6}r^2) \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{t} & 0 \\ -\frac{r}{t^2} & -\frac{1}{t^2} \end{pmatrix}^t \\
&= \begin{pmatrix} \sigma & -\theta\sigma \\ 0 & -\sigma^2 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & -\frac{1}{3}\theta \\ -\frac{1}{3}\theta & \sigma^2 + 2\sigma^2(\frac{\pi^2}{18} - \frac{1}{3}) + \frac{1}{3}\theta^2 \end{pmatrix}^{-1} \begin{pmatrix} \sigma & 0 \\ -\theta\sigma & -\sigma^2 \end{pmatrix} \\
&= \frac{1}{1/3(\sigma^2 + 2\sigma^2(\frac{\pi^2}{18} - \frac{1}{3}) + \frac{1}{3}\theta^2) - \frac{1}{9}\theta^2} \begin{pmatrix} \sigma & -\theta\sigma \\ 0 & -\sigma^2 \end{pmatrix} \\
&\quad \begin{pmatrix} \sigma^2 + 2\sigma^2(\frac{\pi^2}{18} - \frac{1}{3}) + \frac{1}{3}\theta^2 & \frac{1}{3}\theta \\ \frac{1}{3}\theta & \frac{1}{3} \end{pmatrix} \begin{pmatrix} \sigma & 0 \\ -\theta\sigma & -\sigma^2 \end{pmatrix} \\
&= \frac{1}{\sigma^2(\frac{\pi^2}{27} + \frac{1}{9})} \begin{pmatrix} \sigma & -\theta\sigma \\ 0 & -\sigma^2 \end{pmatrix} \begin{pmatrix} \sigma^2\frac{\pi^2}{9} + \frac{1}{3}\sigma^2 + \frac{1}{3}\theta & \frac{1}{3}\theta \\ \frac{1}{3}\theta & \frac{1}{3} \end{pmatrix} \begin{pmatrix} \sigma & 0 \\ -\theta\sigma & -\sigma^2 \end{pmatrix} \\
&= \frac{1}{\sigma^2(\frac{\pi^2}{27} + \frac{1}{9})} \begin{pmatrix} \sigma^3\frac{\pi^2}{9} + \sigma^3\frac{1}{3} & 0 \\ -\frac{1}{3}\theta\sigma^2 & -\frac{1}{3}\sigma^2 \end{pmatrix} \begin{pmatrix} \sigma & 0 \\ -\theta\sigma & -\sigma^2 \end{pmatrix} = \frac{1}{\sigma^2(\frac{\pi^2}{27} + \frac{1}{9})} \begin{pmatrix} \sigma^4(\frac{\pi^2}{9} + \frac{1}{3}) & 0 \\ 0 & \frac{1}{3}\sigma^4 \end{pmatrix} \\
&= \begin{pmatrix} 3\sigma^2 & 0 \\ 0 & \frac{9}{\pi^2+3}\sigma^2 \end{pmatrix}.
\end{aligned}
$$