Homework 9

STAT 984

Emily Robinson

November 21, 2019

Exercise 7.3

Suppose that $X_1, ..., X_n$ are independent and identically distributed with density $f_\theta(x)$, where $\theta \in (0, \infty)$. For each of the following forms of $f_\theta(x)$, prove that the likelihood equation has a unique solution and that this solution maximizes the likelihood function.

(a) *Weibull:* For some constant $a > 0$,

$$f_\theta(x) = a\theta^a x^{a-1} \exp\{-(\theta x)^a\} I\{x > 0\}$$

Consider

$$\begin{aligned}
& f_\theta(x) = a\theta^a x^{a-1} \exp\{-(\theta x)^a\} I\{x > 0\} \\
\implies \quad & L(\theta) = \prod_{i=1}^{n} a\theta^a x_i^{a-1} \exp\{-(\theta x_i)^a\} I\{x_i > 0\} \\
& \quad = a^n \theta^{an} \prod_{i=1}^{n} x + i^{a-1} e^{-\theta^a \sum_{i=1}^n (x_i^a)} \\
\implies \quad & \ell(\theta) = n\log(a) + an\log(\theta) + (a-1)\sum_{i=1}^{n}\log(x_i) - \theta^a \sum_{i=1}^{n}(x_i^a) \\
\implies \quad & \ell'(\theta) = \frac{an}{\theta} - a\theta^{(a-1)}\sum_{i=1}^{n}(x_i^a).
\end{aligned}$$

Then setting $\ell'(\theta) = 0$, implies

$$\begin{aligned}
& \frac{an}{\theta} - a\theta^{(a-1)}\sum_{i=1}^{n}(x_i^a) = 0 \\
\implies \quad & n - \theta^a \sum_{i=1}^{n}(x_i^a) = 0 \\
\implies \quad & \theta^a \sum_{i=1}^{n}(x_i^a) = n \\
\implies \quad & \theta^a = \frac{n}{\sum_{i=1}^n (x_i^a)} \\
\implies \quad & \hat{\theta}_{\text{MLE}} = \left(\frac{n}{\sum_{i=1}^n (x_i^a)}\right)^{1/a}.
\end{aligned}$$

Then consider

$$\ell''(\theta) = -\frac{an}{\theta^2} - a(a-1)\theta^{(a-2)}\sum_{i=1}^{n}(x_i^a) = -an - a(a-1)\theta^a\sum_{i=1}^{n}(x_i^a) \leq 0.$$

Therefore, $\hat{\theta}_{\text{MLE}} = \left(\frac{n}{\sum_{i=1}^{n}(x_i^a)}\right)^{1/a}$ is unique maximizes the likelihood function.

(b) *Cauchy:*

$$f_\theta(x) = \frac{\theta}{\pi}\frac{1}{x^2+\theta^2}$$

Consider

$$\begin{aligned}
& & f_\theta(x) &= \frac{\theta}{\pi}\frac{1}{x^2+\theta^2} \\
&\implies & L(\theta) &= \prod_{i=1}^{n}\frac{\theta}{\pi}\frac{1}{x_i^2+\theta^2} \\
& & &= \frac{\theta^n}{\pi^n}\frac{1}{\prod_{i=1}^{n}(x_i^2-\theta^2)} \\
&\implies & \ell(\theta) &= n\log(\theta) - n\log(\pi) - \sum_{i=1}^{n}\log(x_i^2+\theta^2) \\
&\implies & \ell'(\theta) &= \frac{n}{\theta} - 2\theta\sum_{i=1}^{n}\frac{1}{x_i^2-\theta^2}.
\end{aligned}$$

Then setting $\ell'(\theta)=0$, implies

$$\begin{aligned}
& & \frac{n}{\theta} - 2\theta\sum_{i=1}^{n}\frac{1}{x_i^2-\theta^2} &= 0 \\
&\implies & \sum_{i=1}^{n}\frac{\theta^2}{x_i^2-\theta^2} &= \frac{n}{2}.
\end{aligned}$$

Then consider

$$\ell''(\theta) = -\frac{n}{\theta^2} + \frac{2(\theta^2-x^2)}{(x^2+\theta^2)^2} \leq 0.$$

Therefore, $\sum_{i=1}^{n}\frac{\hat{\theta}^2}{x_i^2-\hat{\theta}^2} = \frac{n}{2}$ has a unique solution that maximizes the likelihood function.

(c)

$$f_\theta(x) = \frac{3\theta^2\sqrt{3}}{2\pi(x^3+\theta^3)}I\{x>0\}$$

Consider

$$\begin{aligned}
f_\theta(x) &= \frac{3\theta^2\sqrt{3}}{2\pi(x^3+\theta^3)} I\{x>0\} \\
\implies \quad L(\theta) &= \prod_{i=1}^{n} \frac{3\theta^2\sqrt{3}}{2\pi(x_i^3+\theta^3)} I\{x_i>0\} \\
&= \frac{3^n\theta^{2n}3^{n/2}}{2^n\pi^n\prod_{i=1} n(x_i^3+\theta^3)} \\
\implies \quad \ell(\theta) &= n\log(3)+2n\log(\theta)+\frac{n}{2}\log(3)-n\log(2)-n\log(\pi)-\sum_{i=1}^{n}\log(x_i^3+\theta^3) \\
\implies \quad \ell'(\theta) &= \frac{2n}{\theta}-3\theta^2\sum_{i=1}^{n}\frac{1}{x_i^3-\theta^3}.
\end{aligned}$$

Then setting $\ell'(\theta)=0$, implies

$$\begin{aligned}
\frac{2n}{\theta}-3\theta^2\sum_{i=1}^{n}\frac{1}{x_i^3-\theta^3} &= 0 \\
\implies \quad \sum_{i=1}^{n}\frac{\theta^3}{x_i^3-\theta^3} &= \frac{2n}{3}.
\end{aligned}$$

Then consider

$$\ell''(\theta) = -\frac{2n}{\theta^2}+\frac{3(\theta^4-2x^3\theta)}{(x^3+\theta^3)^2} \le 0.$$

Therefore, $\sum_{i=1}^{n}\frac{\hat{\theta}^3}{x_i^3-\hat{\theta}^3}=\frac{2n}{3}$ has a unique solution that maximizes the likelihood function.

Exercise 7.8

Prove Theorem 7.9

Hint: Start with $\sqrt{n}(\delta_n-\theta_0)=\sqrt{n}(\delta_n-\tilde{\theta}_n)+\sqrt{n}(\tilde{\theta}_n-\theta_0)$, then expand $\ell'(\tilde{\theta}_n)$ in a Taylor series about θ_0 and substitute the result into Equation (7.15). After simplifying. use the result of Exercise 2.2 along with arguments similar to those leading up to Theorem 7.8.

Proof. Consider $\tilde{\theta}_n \xrightarrow{P} \theta_0$. Then by Taylor's expansion,

$$\ell'(\tilde{\theta}_n)=\ell'(\theta_0)+(\tilde{\theta}_n-\theta_0)[\ell''(\theta_0)+o_p(1)]$$

and

$$\ell''(\tilde{\theta}_n)=\theta_0)+o_p(1).$$

Then substituing in,

$$\begin{aligned}
\sqrt{n}(\delta_n - \theta_0) &= \sqrt{n}(\delta_n - \tilde{\theta}_n) + \sqrt{n}(\tilde{\theta}_n - \theta_0) \\
&= -\frac{\sqrt{n}\ell'(\tilde{\theta}_n)}{\ell''(\tilde{\theta}_n)} + \sqrt{n}(\tilde{\theta}_n - \theta_0) \\
&= -\sqrt{n}\left(\frac{\ell'(\theta_0) + (\tilde{\theta}_n - \theta_0)[\ell''(\theta_0) + o_p(1)]}{\ell''(\theta_0) + o_p(1)}\right) + \sqrt{n}(\tilde{\theta}_n - \theta_0) \\
&= -\frac{\sqrt{n}\ell'(\theta_0)}{\ell''(\theta_0) + o_p(1)} + \sqrt{n}(\tilde{\theta}_n - \theta_0)\left[1 - \frac{\ell''(\theta_0) + o_p(1)}{\ell''(\theta_0) + o_p(1)}\right] \\
&\xrightarrow{p} -\frac{\sqrt{n}\ell'(\theta_0)}{\ell''(\theta_0)} + \sqrt{n}(\tilde{\theta}_n - \theta_0) \times 0 \\
&= -\frac{\sqrt{n}\ell'(\theta_0)}{\ell''(\theta_0)}.
\end{aligned}$$

Therefore, by Slutsky's Theroem and proof of Theorem 7.8 done in class, we know

$$-\frac{\sqrt{n}\ell'(\theta_0)}{\ell''(\theta_0)} \xrightarrow{d} N\left(0, \frac{1}{I(\theta_0)}\right).$$

$\square$

Exercise 7.9

Suppose that the following is a random sample from a logistic density with distribution function $F_\theta(x) = (1 + \exp\{\theta - x\})^{-1}$(I'll cheat and tell you that I used $\theta = 2$.)

1.0944	6.4723	3.118	3.8318	4.1262

1.2853	1.0439	1.7472	4.9483	1.7001

1.0422	0.169	3.6111	0.997	2.9438

(a) Evaluate the unique root of the likelihood equation numerically. Then, taking the sample median as our known $\sqrt{n}$-consistent estimator $\tilde{\theta}_n$ of θ, evaluate the estimator δ_n in Equation (7.15) numerically.

Consider

$$
\begin{aligned}
f_\theta(x) &= \frac{d}{d\theta}F_\theta(x) \\
&= \frac{e^{\theta-x}}{(1+e^{\theta-x})^2} \\
\implies \qquad L(\theta) &= \prod_{i=1}^{n} \frac{e^{\theta-x}}{(1+e^{\theta-x})^2} \\
&= \frac{e^{n\theta-\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n}(1+e^{\theta-x_i})^2} \\
\implies \qquad \ell(\theta) &= n\theta - \sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n}\log(1+e^{\theta-x_i}) \\
\implies \qquad \ell'(\theta) &= n - 2\sum_{i=1}^{n}\frac{e^{\theta}}{(e^{x_i}+e^{\theta})^2} \\
\implies \qquad \ell''(\theta) &= -2\sum_{i=1}^{n}\frac{e^{\theta+x_i}}{(e^{x_i}+e^{\theta})^2} \\
\implies \qquad I(\theta) &= -E[\ell''(\theta)] \\
&= \frac{1}{3}.
\end{aligned}
$$

Then the code below solves for $\hat{\theta}_{\text{MLE}} = 2.39173$ and $\delta_n = 2.385235$.

```
logLogistic <- function(theta = theta, der = 0, x = x) {
    n = length(x)

    value = theta * n - sum(x) - 2 * sum(log(1 + exp(theta -
        x)))
    if (der == 0)
        return(value)

    der1 = n - 2 * sum(exp(theta)/(exp(x) + exp(theta)))
    if (der == 1)
        return(list(value = value, der1 = der1))

    der2 = -2 * sum(exp(theta + x)/(exp(x) + exp(theta))^2)
    return(list(value = value, der1 = der1, der2 = der2))
}

newtonUni = function(f, xInit, maxIt = 20, relConvCrit = 1e-10,
    ...) {

    results = matrix(NA, maxIt, 5)
    colnames(results) = c("value", "x", "Conv", "slope",
        "Hess")
```

```r
    xCurrent = xInit
    for (t in 1:maxIt) {
        evalF = f(xCurrent, der = 2, ...)
        results[t, "value"] = evalF$value
        results[t, "x"] = xCurrent
        results[t, "slope"] = evalF$der1
        results[t, "Hess"] = evalF$der2
        xNext = xCurrent - evalF$der1/evalF$der2
        Conv = abs(xNext - xCurrent)/(abs(xCurrent) +
            relConvCrit)
        results[t, "Conv"] = Conv
        if (Conv < relConvCrit | t > maxIt)
            break
        xCurrent = xNext
    }
    return(list(theta = xNext, value = f(xNext, der = 0,
        ...), convergence = (Conv < relConvCrit), t = t))
}

thetaMLE <- newtonUni(logLogistic, xInit = median(x),
    x = x)$theta
delta <- newtonUni(logLogistic, xInit = median(x),
    x = x, maxIt = 1)$theta

kable(cbind(thetaMLE, delta))
```

thetaMLE	delta
2.39173	2.385235

(b) Find the asymptotic distributions of $\sqrt{n})(\tilde{\theta}_n - 2)$ and $\sqrt{n}(\delta_n - 2)$. Then, simulate 200 samples of size $n = 15$ from the logistic distribution with $\theta = 2$. Find the sample variances of the resulting sample medians and δ_n-estimators. How well does the asymptotic theory match reality?

Then by Theorem 6.7, since $p = 1/2$ and $f(\theta) = 1/4$, we know

$$\sqrt{n}(\tilde{\theta}_n - 2) \xrightarrow{d} N(0, 4).$$

Then by Theorem 7.9, since $I(\theta) = 1/3$, we know

$$\sqrt{n}(\delta_n - 2) \xrightarrow{d} N(0, 3).$$

Our empirical results are consistent with the theoretical results above, δ_n is more efficent than $\tilde{\theta}_n$.

```r
empiricalLogistic <- function(samps = 200, n = 15,
    theta = 2) {
    median = rep(0, samps)
    delta = rep(0, samps)
    for (i in 1:samps) {
        x = rlogis(n, location = theta)
        median[i] = median(x)
        delta[i] = newtonUni(logLogistic, xInit = median[i],
            x = x, maxIt = 1)$theta
    }
    estVar <- n * c(var(median), var(delta))
    estVar
}
empiricalLogistic(samps = 200, n = 15, theta = 2)
```

```
## [1] 4.591794 3.366153
```

Exercise 7.11

If $f_\theta(x)$ forms a location family, so that $f_\theta(x) = f(x-\theta)$ for some density $f(x)$, then the Fisher information $I(\theta)$ is a constant (you may assume this fact without proof).

(a) Verify that for the Cauchy location family,

$$f_\theta(x) = \frac{1}{\pi\{1+(x-\theta)^2\}},$$

we have $I(\theta) = \frac{1}{2}$.

Consider

$$\begin{aligned}
& & f_\theta(x) &= \frac{1}{\pi\{1+(x-\theta)^2\}} \\
&\implies & L(\theta) &= \prod_{i=1}^{n} \frac{1}{\pi\{1+(x-\theta)^2\}} \\
& & &= \pi^{-n} \prod_{i=1}^{n} \frac{1}{1+(x_i-\theta)^2} \\
&\implies & \ell(\theta) &= -n\log(\pi) - \sum_{i=1}^{n} \log(1+(x_i-\theta)^2) \\
&\implies & \ell'(\theta) &= \sum_{i=1}^{n} \frac{2(x_i-\theta)}{1+(x_i-\theta)^2} \\
&\implies & \ell''(\theta) &= \sum_{i=1}^{n} \frac{2(x_i-\theta)^2-2}{(1+(x_i-\theta)^2)^2} \\
&\implies & I(\theta) &= -E[\ell''(\theta)] \\
& & &= \frac{1}{2}.
\end{aligned}$$

(b) For 500 samples of size $n = 51$ from a standard Cauchy distribution, calculate the sample median $\tilde{\theta}_n$ and the efficient estimator δ_n^* of Equation (7.19). Compare the variances of $\tilde{\theta}_n$ and δ_n^* with their theoretical asymptotic limits.

Then by Theorem 6.7, since $p = 1/2$ and $f(\theta) = 1/\pi$, we know

$$\sqrt{n}(\tilde{\theta}_n - \theta) \xrightarrow{d} N\left(0, \frac{\pi^2}{4}\right).$$

Then by Equation (7.19), since $I(\theta) = 1/2$, we know

$$\sqrt{n}(\delta_n^* - 2) \xrightarrow{d} N(0, 2).$$

Our empirical results are consistent with the theoretical results above, δ_n^* is more efficent than $\tilde{\theta}_n$.

```
logCauchy <- function(theta = theta, der = 0, x = x) {
    n = length(x)

    value = -n * log(pi) - sum(log(1 + (x - theta)^2))
    if (der == 0)
        return(value)

    der1 = sum((2 * (x - theta))/(1 + (x - theta)^2))
    if (der == 1)
        return(list(value = value, der1 = der1))

    der2 = sum((2 * (x - theta)^2 - 2)/(1 + (x - theta)^2)^2)
    return(list(value = value, der1 = der1, der2 = der2))
}

empiricalCauchy <- function(samps = 500, n = 51, theta = 2) {
    median = rep(0, samps)
    delta = rep(0, samps)
    for (i in 1:samps) {
        x = rcauchy(n, location = theta)
        median[i] = median(x)
        # delta[i] = newtonUni(logCauchy, xInit =
        # median[i], x = x, maxIt = 1)$theta
        delta[i] = median[i] + (2 * logCauchy(theta = median[i],
            der = 1, x = x)$der1)/n
    }
    n * c(var(median), var(delta))
}
empiricalCauchy(samps = 500, n = 51, theta = 0)
```

```
## [1] 2.612552 2.151267
```

Exercise 7.15

Suppose that $\boldsymbol{\theta} \in \mathbb{R}x\mathbb{R}_+$ (that is $\theta_1 \in \mathbb{R}$ and $\theta_2 \in (0, \infty)$) and

$$f_\theta(x) = \frac{1}{\theta_2} f\left(\frac{x-\theta_1}{\theta_2}\right)$$

for some continuous, differentiable density $f(x)$ that is symmetric about the origin. Find $I(\boldsymbol{\theta})$.

Let $\boldsymbol{\theta} = (\theta_1, \theta_2)$ and suppose

$$f_\theta(x) = \frac{1}{\theta_2} f\left(\frac{x-\theta_1}{\theta_2}\right).$$

Then

$$\log f_\theta(x) = -\log(\theta_2) + \log\left(f\left(\frac{x-\theta_1}{\theta_2}\right)\right)$$

so

$$\frac{\partial}{\partial\theta_1} \log f_\theta(x) = -\frac{1}{\theta_2} \frac{f'\left(\frac{x-\theta_1}{\theta_2}\right)}{f\left(\frac{x-\theta_1}{\theta_2}\right)} \text{ and } f_\theta(x) = -\frac{1}{\theta_2}\left(\frac{f\left(\frac{x-\theta_1}{\theta_2}\right) + \left(\frac{x-\theta_1}{\theta_2}\right) f'\left(\frac{x-\theta_1}{\theta_2}\right)}{f\left(\frac{x-\theta_1}{\theta_2}\right)}\right).$$

Consider $u = \frac{x-\theta_1}{\theta_2}$, therefore, $du = \frac{dx}{\theta_2}$ implies $\theta_2 du = dx$. Thus, the entries in the information matrix are as follows:

$$\begin{aligned}
I_{11}(\boldsymbol{\theta}) &= E_\theta\left[\left(-\frac{1}{\theta_2}\frac{f'\left(\frac{x-\theta_1}{\theta_2}\right)}{f\left(\frac{x-\theta_1}{\theta_2}\right)}\right)^2\right] \\
&= \frac{1}{\theta_2^2}\int \frac{\left[f'\left(\frac{x-\theta_1}{\theta_2}\right)\right]^2}{f\left(\frac{x-\theta_1}{\theta_2}\right)^2} \frac{1}{\theta_2} f\left(\frac{x-\theta_1}{\theta_2}\right) dx \\
&= \frac{1}{\theta_2^3}\int \frac{\left[f'\left(\frac{x-\theta_1}{\theta_2}\right)\right]^2}{f\left(\frac{x-\theta_1}{\theta_2}\right)} dx \\
&= \frac{1}{\theta_2^2}\int \frac{[f'(u)]^2}{f(u)} du
\end{aligned}$$

$$
\begin{aligned}
I_{22}(\boldsymbol{\theta}) &= E_\theta\left[\left(-\frac{1}{\theta_2}\left(\frac{f\left(\frac{x-\theta_1}{\theta_2}\right)+\left(\frac{x-\theta_1}{\theta_2}\right)f'\left(\frac{x-\theta_1}{\theta_2}\right)}{f\left(\frac{x-\theta_1}{\theta_2}\right)}\right)\right)^2\right] \\
&= \frac{1}{\theta_2^2}\int \frac{\left[f\left(\frac{x-\theta_1}{\theta_2}\right)+\left(\frac{x-\theta_1}{\theta_2}\right)f'\left(\frac{x-\theta_1}{\theta_2}\right)\right]^2}{f\left(\frac{x-\theta_1}{\theta_2}\right)^2}\frac{1}{\theta_2}f\left(\frac{x-\theta_1}{\theta_2}\right)dx \\
&= \frac{1}{\theta_2^3}\int \frac{\left[f\left(\frac{x-\theta_1}{\theta_2}\right)+\left(\frac{x-\theta_1}{\theta_2}\right)f'\left(\frac{x-\theta_1}{\theta_2}\right)\right]^2}{f\left(\frac{x-\theta_1}{\theta_2}\right)}dx \\
&= \frac{1}{\theta_2^2}\int \frac{[f(u)+uf'(u)]^2}{f(u)}du
\end{aligned}
$$

$$
\begin{aligned}
I_{12}(\boldsymbol{\theta}) = I_{21}(\boldsymbol{\theta}) &= E_\theta\left[\left(-\frac{1}{\theta_2}\frac{f'\left(\frac{x-\theta_1}{\theta_2}\right)}{f\left(\frac{x-\theta_1}{\theta_2}\right)}\right)\left(-\frac{1}{\theta_2}\left(\frac{f\left(\frac{x-\theta_1}{\theta_2}\right)+\frac{xf'\left(\frac{x-\theta_1}{\theta_2}\right)}{\theta_2^2}}{f\left(\frac{x-\theta_1}{\theta_2}\right)}\right)\right)\right] \\
&= \frac{1}{\theta_2^2}\int \frac{f'\left(\frac{x-\theta_1}{\theta_2}\right)\left[f\left(\frac{x-\theta_1}{\theta_2}\right)+\left(\frac{x-\theta_1}{\theta_2}\right)f'\left(\frac{x-\theta_1}{\theta_2}\right)\right]}{f\left(\frac{x-\theta_1}{\theta_2}\right)^2}\frac{1}{\theta_2}f\left(\frac{x-\theta_1}{\theta_2}\right)dx \\
&= \frac{1}{\theta_2^3}\int \frac{f'\left(\frac{x-\theta_1}{\theta_2}\right)\left[f\left(\frac{x-\theta_1}{\theta_2}\right)+\left(\frac{x-\theta_1}{\theta_2}\right)f'\left(\frac{x-\theta_1}{\theta_2}\right)\right]}{f\left(\frac{x-\theta_1}{\theta_2}\right)}dx \\
&= \frac{1}{\theta_2^2}\int \frac{f'(u)[f(u)+uf'(u)]}{f(u)}du
\end{aligned}
$$

Thus,

$$
I(\boldsymbol{\theta}) = \frac{1}{\theta_2^2}\begin{pmatrix} \int \frac{[f'(u)]^2}{f(u)}du & \int \frac{f'(u)[f(u)+uf'(u)]}{f(u)}du \\ \int \frac{f'(u)[f(u)+uf'(u)]}{f(u)}du & \int \frac{[f(u)+uf'(u)]^2}{f(u)}du \end{pmatrix}.
$$