4. Maximum Likelihood Estimation, Tests, and Confidence Intervals

Setup :

Let $X_1, \ldots, X_n \stackrel{iid}{\sim} X$ be distributed i.i.d. with probability density function

$$f_\theta(x) = (x/\theta^2)\exp\left(-x^2/2\theta^2\right)\mathbf{1}(x \geq 0), \theta > 0.$$

(a)

3.0/3 points (graded)

Let $l(\theta) = \ln L(X_1, \ldots, X_n, \theta)$ denote the log likelihood. Find the critical point of $l(\theta)$. (The critical point is unique because KL divergence is definite.)

(If applicable, enter **barX_n** for $\overline{X}_n$ and **bar(X_n^2)** for $\overline{X_n^2}$.)

Critical point of $l(\theta)$ is at $\theta =$ sqrt(bar(X_n^2)/(2))

Answer: sqrt(bar(X_n^2)/(2))

Find the second derivative $l'' = \dfrac{d^2 l}{d\theta^2}$ of $l(\theta)$. Your answer should be a function of θ and the data $X_1, \ldots, X_n$.

(Do **not** evaluate l'' at the critical point at this stage.)

(If applicable, enter **Sigma_i(X_i)** for $\sum_{i=1}^n X_i$ and and **Sigma_i(X_i^2)** for $\sum_{i=1}^n X_i^2$.)

$l'' = \dfrac{d^2 l}{d\theta^2} =$ 2*n/theta^2 - 3*(Sigma_i(X_i^2))/theta^4

Answer: 2*n/theta^2-3*Sigma_i(X_i^2)/theta^4

Using the second derivative test, is the critcal point you obtain above a global maximum, a global minimum, or neither of $l(\theta)$ in the domain $\theta > 0$?

- (•) global maximum
- ( ) global minimum
- ( ) neither

What can you conclude about the maximum likelihood estimator $\hat{\theta}$ for θ?
(There is no answer box for this question.)

STANDARD NOTATION

Solution:

Given

$$f_\theta(x) = (x/\theta^2)\exp(-x^2/2\theta^2)\mathbf{1}(x \geq 0), \theta > 0.$$

The log-likelihood is

$$l_n(\theta) = \ln \prod_i f_\theta(x_i) = \sum_{i=1}^n \ln x_i - 2n\ln\theta - \frac{1}{2\theta^2}\sum_{i=1}^n (x_i)^2$$

Now, find the critical point of $\ln L(x_1, \ldots, x_n, \theta)$ (there is a unique one because the KL divergence is definite):

$$\begin{aligned} \frac{dl_n}{d\theta} &= -\frac{2n}{\theta} + \frac{\sum_{i=1}^n (x_i)^2}{\theta^3} = 0 \\ \implies \theta &= \sqrt{\frac{\sum_{i=1}^n (x_i)^2}{2n}}. \end{aligned}$$

Check that the critical point is indeed a maximum of $l_n(\theta)$:

$$\begin{aligned} \frac{d^2 l_n}{d\theta^2} &= \frac{2n}{\theta^2} - 3\frac{\sum_{i=1}^n (x_i)^2}{\theta^4} = \frac{1}{\theta^2}\left(2n - 3\frac{\sum_{i=1}^n (x_i)^2}{\theta^2}\right) \\ \left.\frac{d^2 l_n}{d\theta^2}\right|_{\theta=\sqrt{\frac{\sum_{i=1}^n (x_i)^2}{2n}}} &= \frac{2n}{\sum_{i=1}^n (x_i)^2}(2n - 6n) \\ &= -8n^2\sum_{i=1}^n (x_i)^2 < 0. \end{aligned}$$

This means that the critical point we found is a local maximum.

Finally, check that the critial point is a global maximum. Since $l_n'(\theta)$ is defined for all $\theta > 0$, and there is only one critical point, it follows that this critical point is a global maximum in $\theta > 0$. (The function $l_n(\theta)$ is strictly increasing to the left of the critical point and strictly decreasing to the right of the critical point.) Hence, the MLE of θ is

$$\hat{\theta} = \sqrt{\frac{\overline{X_n^2}}{2}}.$$

提交 你已经尝试了2次（总共可以尝试3次）

Answers are displayed within the problem

(b)

1/1 point (graded)

What is the Fisher information $I(\theta)$ of the random variables X_i?

$I(\theta) =$ 4/theta^2 Answer: 4/theta^2

$\frac{4}{\theta^2}$

STANDARD NOTATION

Solution:

Setting $n = 1$ in the expression for $l_n''(\theta)$ computed in the question above, we get

$$l_1''(\theta) = \frac{2}{\theta^2} - \frac{3x^2}{\theta^4}.$$

This gives Fisher information $I(\theta)$:

$$I(\theta) = -\mathbb{E}\left(l_1''(\theta)\right) = -\frac{2}{\theta^2} + \frac{3}{\theta^4}\mathbb{E}\left[x^2\right].$$

It remains to compute the second moment $\mathbb{E}\left[x^2\right]$:

$$\begin{aligned}
\mathbb{E}\left[x^2\right] &= \int_0^\infty \frac{x^3}{\theta^2} e^{-\frac{x^2}{2\theta^2}}\, dx \\
&= \int_0^\infty \left(x^2\right)\left(\frac{x}{\theta^2} e^{-\frac{x^2}{2\theta^2}}\right)\, dx \\
&= Cx^2 e^{-\frac{x^2}{2\theta^2}}\Big|_0^\infty + \int_0^\infty (2x)\left(e^{-\frac{x^2}{2\theta^2}}\right)\, dx \qquad \text{(Integration by part} \\
&= 0 + 2\theta^2\left[-e^{-\frac{x^2}{2\theta^2}}\right]_0^\infty \\
&= 2\theta^2.
\end{aligned}$$

Plugging this back into the expression for the Fisher information, we get

$$I(\theta) = \frac{4}{\theta^2}.$$

提交 你已经尝试了2次（总共可以尝试3次）

Answers are displayed within the problem

(c)

0/2 points (graded)

Use the theorem for the MLE to write down the asymptotic distribution of the MLE $\hat{\theta}$.

Give an asymptotic 95% confidence interval $\mathcal{I}_{\text{plug-in}}$ for θ using the plug-in method. (You may use I in the answer box below to denote $I(\hat{\theta})$, the Fisher Information, which you found in the previous part, evaluated at $\hat{\theta}$.)

(If applicable, enter **I** for $I(\hat{\theta})$, **hattheta** for $\hat{\theta}$, and **q(alpha)** for q_α for any numerical value α. Recall q_α denotes the value such that $\mathbf{P}(Z \geq q_\alpha) = \alpha$ for $Z \sim \mathcal{N}(0,1)$.)

(Do not worry if the parser does not render properly; the graders will work independently. To render properly, add parentheses around **q(alpha)**, i.e. enter **(q(alpha))**.)

$\mathcal{I}_{\text{plug-in}} = [A, B]$ where

$A =$ (hattheta - q(0.05/2))/sqrt(I*n) / sqrt(pi/2) **Answer:** hattheta-q(0.025)/sqrt(n*I)

$B =$ (hattheta + q(0.05/2))/sqrt(I*n) / sqrt(pi/2) **Answer:** hattheta+q(0.025)/sqrt(n*I)

STANDARD NOTATION

Solution:

Since the asymptotic variance is given by $I^{-1}(\theta) = \frac{4}{\theta^2}$, a plug-in confidence interval at confidence level 95% is

$$\begin{aligned}
\mathcal{I} &= \left[\hat{\theta} - \frac{q_{0.025}}{\sqrt{nI}}, \hat{\theta} + \frac{q_{0.025}}{\sqrt{nI}}\right] \\
&= \left[\hat{\theta} - \frac{q_{0.025}}{\sqrt{n}}\frac{\theta}{2}, \hat{\theta} + \frac{q_{0.025}}{\sqrt{n}}\frac{\theta}{2}\right]
\end{aligned}$$

提交 你已经尝试了3次（总共可以尝试3次）

☐ Answers are displayed within the problem

(d)

1.0/1 point (graded)
Use the results from the previous parts to give a test with asymptotic level α for testing

$$H_0 : \theta = 1 \quad \text{v.s.} \quad H_1 : \theta \neq 1.$$

Suppose $n = 100$ and the data gives $\overline{X}_n = 1.5$ and $\overline{X_n^2} = 4.0$. Find the p-value associated to this data for this hypothesis test.

(If applicable, enter **Phi(z)** for the cdf $\Phi(z)$ of a normal variable Z, **q(alpha)** for q_α for any numerical value α.)

p-value: 1 - Phi(10*sqrt(2)*(2.5-sqrt(pi)))   ☐ **Answer:** 2*(1-Phi(10*sqrt(2)*(sqrt(2)-1)))

Correction Note: An earlier version gave the data $\bar{X}_n = 2.5$ instead, which led to the variance being negative, i.e. impossible data! The grader has no issue.

STANDARD NOTATION

Solution:

The desired test is

$$\Psi = \mathbf{1}\left(|T_n| > q_{\alpha/2}\right) \quad \text{where } T_n = \sqrt{nI(\hat{\theta})}\,(\hat{\theta} - 1)$$

With $n = 100$, $\overline{X}_n = 1.5$ and $\overline{X_n^2} = 4.0$, $\hat{\theta} = \sqrt{\frac{\overline{X_n^2}}{2}} = \sqrt{2}$, and $I(\hat{\theta}) = \frac{4}{\hat{\theta}^2} = 2$. This gives the associated p-value is

$$\begin{aligned} 2\left(1 - \Phi(T_n)\right) &= 2\left(1 - \Phi\left(\sqrt{nI(\hat{\theta})}\left(\hat{\theta} - 1\right)\right)\right) \\ &= 2\left(1 - \Phi\left(\sqrt{(100)(2)}\left(\sqrt{2} - 1\right)\right)\right) \approx 0.0006. \end{aligned}$$

(Hence, for any test with level $\alpha > 0.0006$, the test will reject the null hypothesis.)

提交 你已经尝试了1次（总共可以尝试3次）

☐ Answers are displayed within the problem

讨论

显示讨论

主题： Unit 3 Methods of Estimation:Homework 6 Maximum Likelihood Estimation and Method of Moments / 4. Maximum Likelihood Estimation, Tests, and Confidence Intervals