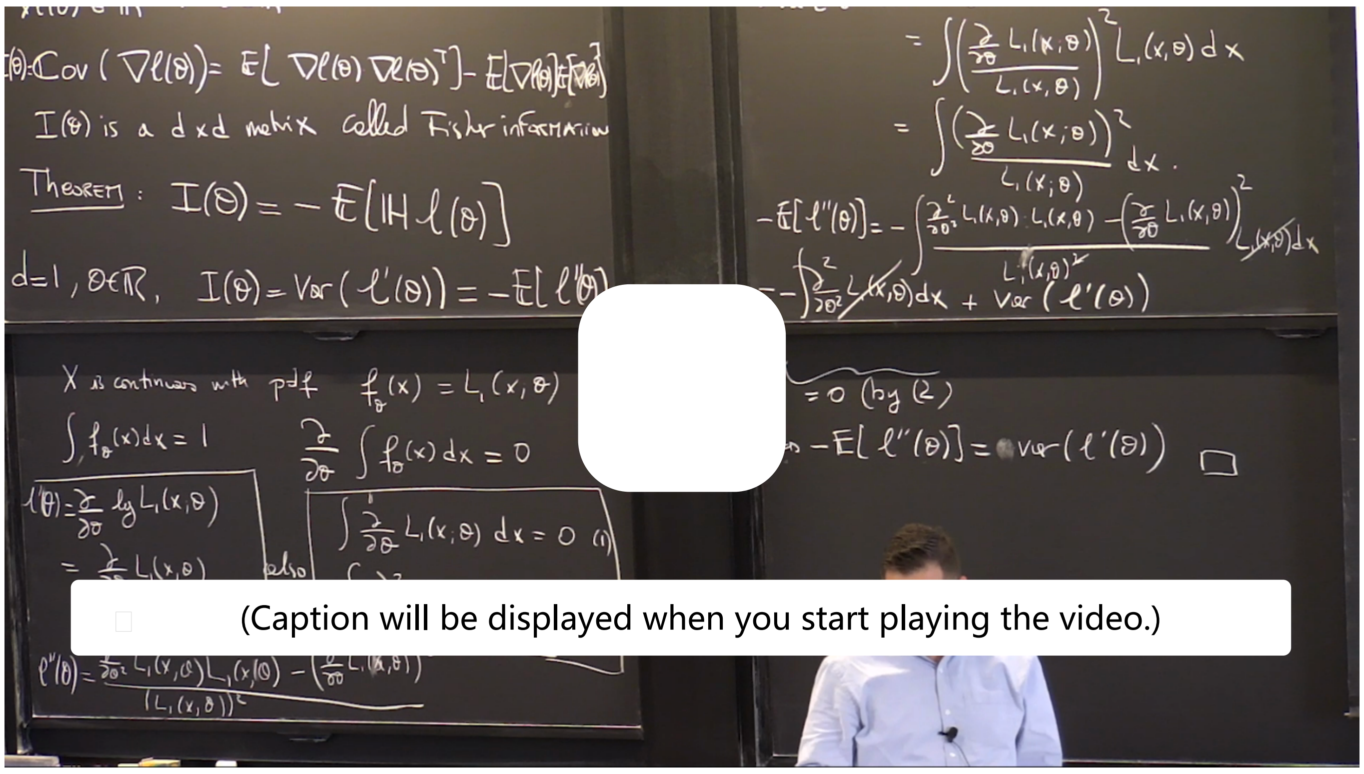
4. Examples of Fisher Information Computation

Fisher Information of the Bernoulli Random Variable

Start of transcript. Skip to the end.



All right.
So still I just defined a quantity.
I'm playing with it and I'm telling you
it's actually equal to some other quantity.
And before I even tell you where this is actually going in,
let's actually play around a little bit
and compute it in some specific examples.
So what you have now--

视频
下载视频文件

字幕
下载 SubRip (.srt) file
下载 Text (.txt) file

Fisher Information of the Binomial Random Variable

1/1 point (graded)
Let X be distributed according to the binomial distribution of n trials and parameter $p \in (0,1)$. Compute the Fisher information $\mathcal{I}(p)$.

Hint: Follow the methodology presented for the Bernoulli random variable in the above video.

$\mathcal{I}(p)$: n*p/p^2 - (n*p - n)/(1-p)^ **Answer:** n/(p*(1-p))

$\frac{n \cdot p}{p^2} - \frac{n \cdot p - n}{(1-p)^2}$

STANDARD NOTATION

Solution:

The logarithm of the pmf of a binomial random variable X, treated as a random function, can be written as

$$\ell(p) \triangleq \ln \binom{n}{X} + X \ln p + (n-X)\ln(1-p), \quad X \in \{0,1,\ldots,n\}.$$

The derivative of $\ell(p)$ with respect to p is

$$\ell'(p) = \frac{X}{p} - \frac{n-X}{1-p},$$

which means the second derivative is

$$\ell''(p) = -\frac{X}{p^2} - \frac{n-X}{(1-p)^2}.$$

The Fisher information $\mathcal{I}(p)$, therefore, is

$$\begin{aligned}\mathcal{I}(p) = -\mathbb{E}[\ell''(p)] &= \mathbb{E}\left[\frac{X}{p^2} + \frac{n-X}{(1-p)^2}\right] \\ &= \frac{np}{p^2} + \frac{n-np}{(1-p)^2} \\ &= \frac{n}{p(1-p)}.\end{aligned}$$

提交 你已经尝试了2次（总共可以尝试2次）

Answers are displayed within the problem

Fisher Information of a Bernoulli-Like Random Variable

1/1 point (graded)

Consider the following experiment: You take a coin that lands a head (H) with probability $0 < p < 1$ and you toss it twice. Define X as the following random variable:

$$X = \begin{cases} 1 & \text{if outcome is HH} \\ 0 & \text{otherwise} \end{cases}$$

Compute the Fisher information $\mathcal{I}(p)$.

$\mathcal{I}(p)$: 2+2*(p^2+1)/(1-p^2)

$2 + \frac{2\cdot(p^2+1)}{1-p^2}$

Answer: 4/(1-p^2)

STANDARD NOTATION

Solution:

Following the Bernoulli and binomial examples,

$$\ell(p) \triangleq 2X \ln p + (1-X)\ln(1-p^2), \quad X \in \{0,1\}.$$

The derivative of $\ell(p)$ with respect to p is

$$\ell'(p) = \frac{2X}{p} - 2p \cdot \frac{1-X}{1-p^2},$$

which means the second derivative is

$$\ell''(p) = -\frac{2X}{p^2} - 2\cdot\frac{(1-X)}{1-p^2} - 4p^2 \cdot \frac{1-X}{(1-p^2)^2}.$$

The Fisher information $\mathcal{I}(p)$, therefore, is

$$\begin{aligned}\mathcal{I}(p) = -\mathbb{E}[\ell''(p)] &= \mathbb{E}\left[\frac{2X}{p^2} + 2\cdot\frac{(1-X)}{1-p^2} + 4p^2\cdot\frac{1-X}{(1-p^2)^2}\right] \\ &= \frac{2p^2}{p^2} + \frac{2(1-p^2)}{(1-p^2)} + 4p^2\cdot\frac{1-p^2}{(1-p^2)^2}\end{aligned}$$

$$= 4 + \frac{4p^2}{1-p^2}$$

$$= \frac{4}{1-p^2}$$

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

Fisher Information of a Modified Gaussian Random Vector

3/4 points (graded)

Let $\mathbf{X}$ be a gaussian random vector with **independent** components $X^{(i)} \sim \mathcal{N}(\alpha + \beta t_i, 1)$ for $i = 1, \ldots, d$, where t_i are known constants and α and β are unknown parameters.

Compute the Fisher information matrix $\mathcal{I}(\theta)$ using the formula $\mathcal{I}(\theta) = -\mathbb{E}[\mathbf{H}\ell(\theta)]$.

Use **S_1** for $\sum_{i=1}^d t_i$ and **S_2** for $\sum_{i=1}^d t_i^2$.

$\mathcal{I}(\theta)_{1,1} =$ d

d

Answer: d + 0*S_1 + 0*S_2

$\mathcal{I}(\theta)_{1,2} =$ S_1

S_1

Answer: 0*d + S_1 + 0*S_2

$\mathcal{I}(\theta)_{2,1} =$ S_1

S_1

Answer: 0*d + S_1 + 0*S_2

$\mathcal{I}(\theta)_{2,2} =$ S_2-d*S_1

$S_2 - d \cdot S_1$

Answer: 0*d + 0*S_1 + S_2

算错了，对b求偏导了以后，忘记把b去掉了

Hint: Let $\theta = [\alpha \ \ \beta]^T$ denote the paramaters of the statistical model. $\ell(\theta)$ is a real-valued function of θ as given by the joint pdf at any fixed $\mathbf{x}$.

Solution:

Let $\theta = [\alpha \ \ \beta]^T$ denote the paramaters of the statistical model. The Gaussian random vector $\mathbf{X}$ has the pdf

$$f_\theta(\mathbf{x}) = (2\pi)^{-\frac{d}{2}} e^{-\frac{1}{2}\sum_{i=1}^d \left(x^{(i)} - \alpha - \beta t_i\right)^2}, \quad \mathbf{x} = \left[x^{(1)} \ \ x^{(2)} \ \ \cdots \ \ x^{(d)}\right]^T \in \mathbb{R}^d,$$

as the variance of each individual component is equal to 1 and the components are independent.

Taking $\ln$ of the pdf yields (written as a random function)

$$\ell(\theta) = -\frac{d}{2}\ln(2\pi) - \frac{1}{2}\left[\sum_{i=1}^d \left(\left(X^{(i)} - \beta t_i\right)^2 - 2\alpha\left(X^{(i)} - \beta t_i\right) + \alpha^2\right)\right]$$

Therefore,

$$\nabla \ell(\theta) = \begin{bmatrix} \sum_{i=1}^d \left(X^{(i)} - \beta t_i - \alpha\right) \\ \sum_{i=1}^d \left(t_i X^{(i)} - \beta t_i^2 - \alpha t_i\right) \end{bmatrix},$$

from which we can obtain the hessian

$$\mathbf{H}\ell(\theta) = \begin{bmatrix} \sum_{i=1}^d (-1) & \sum_{i=1}^d (-t_i) \\ \sum_{i=1}^d (-t_i) & \sum_{i=1}^d (-t_i^2) \end{bmatrix}.$$

Therefore,

$$\mathcal{I}(\theta) = -\mathbb{E}[\mathbf{H}\ell(\theta)] = \begin{bmatrix} d & \sum_{i=1}^d t_i \\ \sum_{i=1}^d t_i & \sum_{i=1}^d t_i^2 \end{bmatrix},$$

where the expectation is taken with respect to the pdf of the random vector $\mathbf{X}$. Since none of the entries of the hessian contained any $X^{(i)}$, the expectation was simply the hessian matrix itself.

提交 你已经尝试了4次（总共可以尝试4次）

Answers are displayed within the problem

讨论

显示讨论

主题： Unit 3 Methods of Estimation:Lecture 11: Fisher Information, Asymptotic Normality of MLE; Method of Moments / 4. Examples of Fisher Information Computation