Econ0190 Section 6

Oct. 4, 2024

Yiwei Sun

4. [Hong] Suppose $\{X_1, X_2 \dots X_n\}$ is iid $N(0, \sigma^2)$. Consider the following estimator for σ^2:

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$$

Find:

(a) the sampling distribution of $n\hat{\sigma}^2/\sigma^2$.

(b) $\mathbb{E}\hat{\sigma}^2$.

(c) $\mathrm{var}(\hat{\sigma}^2)$.

(d) $\mathrm{MSE}(\hat{\sigma}^2)$.

(a) $$\frac{n\hat{\sigma}^2}{\sigma^2} = \frac{\cancel{n}\cdot\frac{1}{\cancel{n}}\sum_{i=1}^{n} x_i^2}{\sigma^2} = \sum_{i=1}^{n}\frac{x_i^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{x_i}{\sigma}\right)^2$$

Since $x_i \sim \mathcal{N}(0,\sigma^2)$, $\frac{x_i}{\sigma} \sim \mathcal{N}(0,1)$.

Since x_i are iid, $\left(\frac{x_i}{\sigma}\right)$ are independent standard normals.

$$\Rightarrow \frac{n\hat{\sigma}^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{x_i}{\sigma}\right)^2 \sim \chi^2_n.$$

R.V. X, a, b constant

$E[aX+b] = aE[X]+b$

$E\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n}E\left[\sum_{i=1}^{n} x_i\right]$

$= \frac{1}{n}\sum_{i=1}^{n} E[x_i]$

(b) $$E[\hat{\sigma}^2] = E\left[\frac{1}{n}\sum_{i=1}^{n} x_i^2\right]$$

$E[\cdot]$ linear operator $= \frac{1}{n}\sum_{i=1}^{n} \underbrace{E[x_i^2]}$

b/c iid $E[x_i^2] = E[x^2], \forall i$.

$$= \frac{1}{\cancel{n}}\cancel{n}\cdot E[x^2]$$

$$= E[x^2]$$

3 ways to go from here:

1) integrate out the pdf

$$E[x^2] = \int x^2 f(x)\,dx$$

$$= \int x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}x^2}\,dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int \underbrace{x}_{u}\,\underbrace{\left(xe^{-\frac{1}{2\sigma^2}x^2}\right)dx}_{dv}$$

$\int u\,dv = uv - \int v\,du$

we want $\frac{dv}{dx} = xe^{-\frac{x^2}{2\sigma^2}} \Rightarrow v = -\sigma^2 e^{-\frac{x^2}{2\sigma^2}}$

$$= \frac{1}{\sigma\sqrt{2\pi}}\left(\underbrace{-\sigma^2 x e^{-\frac{x^2}{2\sigma^2}}\Big|_{-\infty}^{\infty}}_{\substack{=0\text{, b/c L'Hopital's} \\ \text{rule. See Section 3}}} - \underbrace{\int -\sigma^2 e^{-\frac{x^2}{2\sigma^2}}\,dx}_{=+\sigma^2\int e^{-\left(\frac{x}{\sqrt{2\sigma^2}}\right)^2}dx}\right)$$

Recall Gaussian integral:

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$$

Let $u = \frac{x}{\sqrt{2\sigma^2}} \Rightarrow \frac{du}{dx} = \frac{1}{\sqrt{2\sigma^2}}$

$\Rightarrow dx = \sqrt{2\sigma^2}\,du$.

$$= \sigma^2\left(\sqrt{2\sigma^2}\underbrace{\int e^{-u^2}du}_{\sqrt{\pi}}\right)$$

$$= \sigma^2 \cdot \sqrt{2\sigma^2\pi}$$

$$= \frac{1}{\cancel{\sqrt{2\pi\sigma^2}}} \cdot \sigma^2 \cdot \cancel{\sqrt{2\sigma^2\pi}}$$

$$= \sigma^2 \qquad \blacksquare$$

2) Notice: $\underbrace{var(x)}_{\sigma^2} = \underbrace{E[x^2]}_{?} - \underbrace{(\underbrace{E[x]}_{0})^2}_{0}$

$$\Rightarrow E[x^2] = \sigma^2 \qquad \blacksquare$$

In general: $E[x^2] = var(x) + (E[x])^2$.

3) From (a), $\frac{n\hat{\sigma}^2}{\sigma^2} \sim \chi^2_n$.

<u>Property of χ^2:</u> If $X \sim \chi^2_n$, then $E[X] = n$.

$var(X) = 2n$.

$$\Rightarrow E\left[\frac{n\hat{\sigma}^2}{\sigma^2}\right] = n$$

constants

$$\Rightarrow \frac{n}{\sigma^2} E[\hat{\sigma}^2] = n$$

$$\Rightarrow E[\hat{\sigma}^2] = \cancel{n} \cdot \frac{\sigma^2}{\cancel{n}} = \sigma^2 \qquad \blacksquare$$

(c) $var(\hat{\sigma}^2)$?

$var(\hat{\sigma}^2) = E[(\hat{\sigma} - \overbrace{E[\hat{\sigma}^2]}^{\sigma^2})^2]$ → complicated, but can worked out.

Alternatively: from (a)

$$\mathrm{var}\left(\frac{n\hat{\sigma}^2}{\sigma^2}\right) = 2n$$

$$\Rightarrow \frac{n^2}{\sigma^4}\mathrm{var}(\hat{\sigma}^2) = 2n$$

$$\Rightarrow \mathrm{var}(\hat{\sigma}^2) = 2n \cdot \frac{\sigma^4}{n^2} = \frac{2\sigma^4}{n}$$

(d) $$\mathrm{MSE}(\hat{\sigma}^2) = \underbrace{\mathrm{var}(\hat{\sigma}^2)}_{\frac{2\sigma^4}{n}} + \big(\underbrace{\mathrm{bias}(\hat{\sigma}^2)}_{\underbrace{\mathbb{E}[\hat{\sigma}^2]}_{\sigma^2} - \sigma^2 = 0}\big)^2$$

$$= \frac{2\sigma^4}{n}.$$

5. Let $\{X_1, \ldots, X_n\}$ be a random sample from a Poisson distribution with parameter λ
iid

$$P\{X_i = j\} = \frac{e^{-\lambda}\lambda^j}{j!}, j = 0, 1 \ldots$$

(a) Find a minimal sufficient statistic for λ, say T.

(b) Suppose we are interested in estimating probability of a count of zero $\theta = P\{X = 0\} = e^{-\lambda}$. Find an unbiased estimator for θ, say $\hat{\theta}_1$. (Note $P\{X = 0\} = \mathbb{E}[\mathbf{1}\{X = 0\}]$.)

(c) Is the estimator in (b) a function of the minimal sufficient statistics T?

(d) Use the definition of a sufficient statistic and an unbiased estimator, show that the estimator $\hat{\theta}_2 = \mathbb{E}[\hat{\theta}_1 | T]$ is also unbiased and $\mathrm{MSE}(\hat{\theta}_2) \le \mathrm{MSE}(\hat{\theta}_1)$.

(e) Based on (d), find an analytic form of $\hat{\theta}_2$.

(a) $$f(\underbrace{\mathbf{x}}_{\text{DATA}} | \lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!}$$

$$= \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}, \text{ for } x_i \in \{0, 1, \ldots\}$$

Let $g(T(\mathbf{x})|\lambda) = e^{-n\lambda}\lambda^{\sum_{i=1}^n x_i}$, and $h(\mathbf{x}) = \dfrac{1}{\prod_{i=1}^n x_i!}$

By factorization theorem, $T(\mathbf{x}) = \sum_{i=1}^n x_i$ is a S.S.

$T(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^n x_i$ also work.
with constants adjusted properly.

$\Rightarrow$ Still need to show it's a minimal S.S.

Consider 2 sample realizations, x, y

$$\frac{f(x|\lambda)}{f(y|\lambda)} = \frac{\cancel{e^{-n\lambda}} \lambda^{\sum_{i=1}^{n} x_i} / \prod_{i=1}^{n}(x_i!)}{\cancel{e^{-n\lambda}} \lambda^{\sum_{i=1}^{n} y_i} / \prod_{i=1}^{n}(y_i!)}$$

$$= \underbrace{\left(\lambda^{\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} y_i}\right)}_{\text{doesn't depend on } \lambda \text{ iff } \underbrace{\sum_{i=1}^{n} x_i}_{T(x)} = \underbrace{\sum_{i=1}^{n} y_i}_{T(y)}} \underbrace{\left(\frac{\prod_{i=1}^{n} y_i}{\prod_{i=1}^{n} x_i!}\right)}_{\text{doesn't depend on } \lambda}$$

$\Rightarrow$ We have shown that $\frac{f(x|\lambda)}{f(y|\lambda)}$ doesn't depend on λ iff $T(x) = T(y)$.

$\Rightarrow$ By theorem from class, $T(x) = \sum_{i=1}^{n} x_i$ is a minimal s.s.

(b) unbiased estimator $\hat{\theta}_1$ for $\theta = P(X=0)$

Hint: $P(X=0) = E[\mathbb{1}\{x=0\}]$ .... ①

A natural estimator is the sample analog of ①

$$\Rightarrow \hat{\theta}_1 = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{x_i = 0\}$$

$$= \frac{\sum_{i=1}^{n} \mathbb{1}\{x_i = 0\}}{n}$$

→ counting # 0 in the data

$\{0, 2, 3, 0, \ldots\}$

$1 \quad 0 \quad 0 \quad 1 \quad \ldots$

Need to check if $\hat{\theta}_1$ is unbiased:

$$E[\hat{\theta}_1] = E\left[\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{x_i=0\}\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n} \underbrace{E[\mathbb{1}\{x_i=0\}]}_{P(X=0)}$$

b/c iid

$P(x_1=0)$

$P(X_2=0)$

$\vdots$

$$= P(X=0) = \theta \qquad \blacksquare$$

(c) NO!

samples (x_1, x_2, x_3)	$T(x) = \sum_{i=1}^{n} x_i$	$\hat{\theta}_1 = \frac{1}{n}\sum_{i=1}^{n}\mathbb{1}\{x_i=0\}$
$(0, 0, 5)$	5	$2/3$
$(0, 2, 3)$	5	$1/3$

One value of $T(x)$ can map to multiple values of $\hat{\theta}_1$

$\Rightarrow$ $\hat{\theta}_1$ is a function of T.

(d) $\hat{\theta}_2 = E[\hat{\theta}_1 \mid T]$

i) <u>unbiasedness</u>:

$$E[\hat{\theta}_2] = E[\, E[\hat{\theta}_1 \mid T] \,] \overset{LIE}{=} E[\hat{\theta}_1] \overset{\text{by (b)}}{=} \theta$$

ii) WTS: $MSE(\hat{\theta}_2) \leq MSE(\hat{\theta}_1)$

$$\begin{aligned} MSE(\hat{\theta}_1) &\overset{def}{=} E[(\hat{\theta}_1 - \theta)^2] \\ &= E[(\hat{\theta}_1 - \hat{\theta}_2 + \hat{\theta}_2 - \theta)^2] \\ &= E[(\hat{\theta}_1 - \hat{\theta}_2)^2] + \underbrace{E[(\hat{\theta}_2 - \theta)^2]}_{MSE(\hat{\theta}_2)} + 2\underbrace{E[(\hat{\theta}_1 - \hat{\theta}_2)(\hat{\theta}_2 - \theta)]}_{} \end{aligned}$$

$$\overset{LIE}{=} E[E[(\hat{\theta}_1 - \hat{\theta}_2)(\hat{\theta}_2 - \theta) \mid T]]$$

$E[\hat{\theta}_1 \mid T]$ para.

don't depend on T. can be treated as constant

$$= E[(\underbrace{E[\hat{\theta}_1 \mid T]}_{:= \theta_2} - \hat{\theta}_2)(\hat{\theta}_2 - \theta)]$$

0

$= 0$

$$= \underbrace{E[(\hat{\theta}_1 - \hat{\theta}_2)^2]}_{\geq 0.} + MSE(\hat{\theta}_2)$$

$$\geq MSE(\hat{\theta}_2)$$

$\Rightarrow$ Rao-Blackwell theorem.