Lecture 30: Comparing estimators

Ciaran Evans

Course so far

- Maximum likelihood estimation
- Logistic regression
- Asymptotics
- Asymptotic properties of MLEs
- Hypothesis testing
- Confidence intervals

Common theme: Likelihoods and MLEs

Question: Why maximum likelihood estimation?

Today:

- Another approach to estimation: method of moments
- What makes a good estimator?

Example

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} Uniform(0, \theta)$. How could I estimate θ?

1. MLE: $\hat{\theta} = X_{(n)}$

2. $\mathbb{E}[X] = \frac{\theta}{2}$ $\qquad \bar{X} \approx \mathbb{E}[X]$

   $\Rightarrow \hat{\theta} = 2\bar{X}$

3. $\mathbb{E}[X^2] = Var(X) + (\mathbb{E}[X])^2$

   $= \frac{\theta^2}{12} + \frac{\theta^2}{4} = \frac{\theta^2}{3} \Rightarrow \hat{\theta} = \sqrt{\frac{3}{n}\sum_i X_i^2}$

4. $\hat{\theta} = 5$ (probably a terrible estimate)

Example

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} \mathit{Uniform}(a, b)$. How could I estimate a and b?

① MLE: $\hat{a} = X_{(1)}$, $\hat{b} = X_{(n)}$

② $\mathbb{E}[X] = \dfrac{a+b}{2} = \mu_1$ $\qquad \hat{\mu}_1 = \frac{1}{n}\sum_i X_i = \bar{X}$

$\mathbb{E}[X^2] = \frac{1}{3}(a^2 + ab + b^2) = \mu_2$ $\qquad \hat{\mu}_2 = \frac{1}{n}\sum_i X_i^2$

$$b = 2\mu_1 - a$$

$$\Rightarrow \mu_2 = \frac{1}{3}\left(a^2 + a(2\mu_1 - a) + (2\mu_1 - a)^2\right)$$

$$= \frac{1}{3}\left(a^2 - 2a\mu_1 + 4\mu_1^2\right)$$

⋮ (algebra)

$a = \mu_1 - \sqrt{3(\mu_2 - \mu_1)^2}$ $\qquad b = \mu_1 + \sqrt{3(\mu_2 - \mu_1)^2}$

$\hat{a} = \hat{\mu}_1 - \sqrt{3(\hat{\mu}_2 - \hat{\mu}_1)^2}$ $\qquad \hat{b} = \hat{\mu}_1 + \sqrt{3(\hat{\mu}_2 - \hat{\mu}_1)^2}$

Method of moments

Let $X_1, ..., X_n$ be a sample from a distribution with probability function $f(x|\theta_1, ..., \theta_k)$, with k parameters $\theta_1, ..., \theta_k$.

Let

$$\mu_1 = \mathbb{E}[X] = g_1(\theta_1, ..., \theta_k) \qquad \hat{\mu}_1 = \frac{1}{n}\sum_i X_i$$

$$\mu_2 = \mathbb{E}[X^2] = g_2(\theta_1, ..., \theta_k) \qquad \hat{\mu}_2 = \frac{1}{n}\sum_i X_i^2$$

$$\mu_3 = \mathbb{E}[X^3] = g_3(\theta_1, ..., \theta_k) \qquad \hat{\mu}_3 = \frac{1}{n}\sum_i X_i^3$$

$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$

$$\mu_k = \mathbb{E}[X^k] = g_k(\theta_1, ..., \theta_k) \qquad \hat{\mu}_k = \frac{1}{n}\sum_i X_i^k$$

The method of moments (MoM) approach estimates $\theta_1, ..., \theta_k$ by the solutions to

$$\hat{\mu}_1 = g_1(\hat{\theta}_1, ..., \hat{\theta}_k)$$

$$\vdots$$

$$\hat{\mu}_k = g_k(\hat{\theta}_1, ..., \hat{\theta}_k)$$

Example

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$.

Want: $\hat{\mu}$ and $\hat{\sigma}^2$

$\mu_1 = \mathbb{E}[X] = \mu$

$\hat{\mu}_1 = \bar{X}$

$\mu_2 = \mu^2 + \sigma^2$

$\hat{\mu}_2 = \frac{1}{n} \sum_i X_i^2$

$\hat{\mu} = \bar{X}$ ✓

$\sigma^2 = \mu_2 - \mu^2 \quad \Rightarrow \quad \hat{\sigma}^2 = \hat{\mu}_2 - \hat{\mu}_1^2$

$= \frac{1}{n} \sum_i X_i^2 - (\bar{X})^2$

$= \frac{1}{n} \sum_i (X_i - \bar{X})^2$ ✓

$\sum_i (X_i - \bar{X})^2 = \sum_i (X_i - \bar{X})(X_i - \bar{X})$

$= \sum_i X_i (X_i - \bar{X}) - \bar{X} \cancel{\sum_i (X_i - \bar{X})}\ 0$

$= \sum_i X_i^2 - \bar{X} \sum_i X_i = \sum_i X_i^2 - \bar{X}\, n \bar{X} = \sum_i X_i^2 - n(\bar{X})^2$

What makes a good estimator?

Suppose $X_1, ..., X_n \overset{iid}{\sim} Uniform(0, \theta)$. Two possible estimates:

$$\text{MLE: } \widehat{\theta} = X_{(n)} \qquad \text{MoM: } \widehat{\theta} = 2\overline{X}$$

Question: How would I choose between these estimators?

Bias: $E[\hat{\theta}] - \theta$ (want bias close to 0)

MSE

Consistency: $\hat{\theta} \xrightarrow{p} \theta$ as $n \to \infty$

$\hat{\theta}$ is a function of a sufficient statistic?
(we will talk about this later)

Variance: $Var(\hat{\theta})$ (want variance close to 0)

Bias, variance, and MSE (mean squared error)

MSE: Let $\hat{\theta}$ be an estimator of θ

The MSE of $\hat{\theta}$ is $\mathbb{E}_\theta[(\hat{\theta}-\theta)^2]$

$$\mathbb{E}_\theta[(\hat{\theta}-\theta)^2] = \mathrm{Var}(\hat{\theta}) + \underbrace{(\mathbb{E}_\theta[\hat{\theta}]-\theta)^2}_{(\mathrm{Bias})^2}$$

$$= \mathrm{Var}(\hat{\theta}) + \mathrm{Bias}^2(\hat{\theta})$$

- One approach to choosing $\hat{\theta}$: try and minimize MSE

- Another approach: restrict ourselves to unbiased estimators, minimize variance
  MVUE: "minimum variance unbiased estimator"

Example

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} Uniform(0, \theta)$.

MoM: $\hat{\theta} = 2\bar{X}$

$\mathbb{E}[2\bar{X}] = 2\,\mathbb{E}[\bar{X}] = 2\left(\frac{\theta}{2}\right)$
$= \theta$

$\Rightarrow$ Bias $= 0$

$Var(2\bar{X}) = 4\,Var(\bar{X})$
$= 4\,\frac{Var(X)}{n}$
$= 4\left(\frac{\theta^2}{12n}\right) = \frac{\theta^2}{3n}$

$MSE(2\bar{X}) = \frac{\theta^2}{3n} \to 0$ as $n \to \infty$

MLE: $\hat{\theta} = X_{(n)}$

$\mathbb{E}[X_{(n)}] = \frac{n}{n+1}\theta$

$Bias(X_{(n)}) = -\frac{\theta}{n+1}$

(tend to underestimate)

$Var(X_{(n)}) = \frac{\theta^2 n}{(n+1)^2(n+2)}$

$MSE(X_{(n)}) =$

$\frac{\theta^2}{(n+1)^2} + \frac{\theta^2 n}{(n+1)^2(n+2)} = \frac{2\theta^2}{(n+1)(n+2)}$

$< \frac{\theta^2}{3n}$ $(MSE(\bar{X}))$

Example

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} Uniform(0, \theta)$.

Try unbiasing $X_{(n)}$:

$$\mathbb{E}[X_{(n)}] = \frac{n}{n+1}\theta$$

unbiasing: $\left(\frac{n+1}{n}\right) X_{(n)}$

$$\mathbb{E}\left[\left(\frac{n+1}{n}\right) X_{(n)}\right] = \theta \quad ✓ \qquad (\text{Bias} = 0)$$

$$Var\left(\left(\frac{n+1}{n}\right) X_{(n)}\right) = \left(\frac{n+1}{n}\right)^2 Var(X_{(n)})$$

$$= \frac{\theta^2}{n(n+2)} < \frac{2\theta^2}{(n+1)(n+2)} = MSE(X_{(n)})$$

$$\Rightarrow MSE\left(\left(\frac{n+1}{n}\right) X_{(n)}\right) < MSE(X_{(n)}) \quad ✓$$

Example

Suppose $X_1, ..., X_n \overset{iid}{\sim} N(\mu, \sigma^2)$. Consider two estimates of σ^2:

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 \qquad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$