Math 362: Mathematical Statistics II

Le Chen
le.chen@emory.edu

Emory University
Atlanta, GA

Last updated on April 13, 2021

2021 Spring

Chapter 5. Estimation

Plan

Chapter 5. Estimation

Definition. An estimator $\widehat{\theta}_n = h(W_1, \cdots, W_n)$ is said to be consistent if it converges to θ *in probability*, i.e., for any $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(|\widehat{\theta}_n - \theta| < \epsilon\right) = 1.$$

Comment: In the ϵ-δ language, the above convergence in probability says

$$\forall\epsilon > 0,\ \forall\delta > 0,\ \exists n(\epsilon,\delta) > 0,\ s.t.\ \forall n \geq n(\epsilon,\delta),$$

$$\mathbb{P}\left(|\widehat{\theta}_n - \theta| < \epsilon\right) > 1 - \delta.$$

Definition. An estimator $\widehat{\theta}_n = h(W_1, \cdots, W_n)$ is said to be consistent if it converges to θ *in probability*, i.e., for any $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(|\widehat{\theta}_n - \theta| < \epsilon\right) = 1.$$

Comment: In the ϵ-δ language, the above convergence in probability says

$$\forall\epsilon > 0,\ \forall\delta > 0,\ \exists n(\epsilon,\delta) > 0,\ s.t.\ \forall n \geq n(\epsilon,\delta),$$

$$\mathbb{P}\left(|\widehat{\theta}_n - \theta| < \epsilon\right) > 1 - \delta.$$

A useful tool to check convergence in probability is

Theorem. (Chebyshev's inequality) Let W be any r.v. with finite mean μ and variance σ^2. Then for any $\epsilon > 0$

$$\mathbb{P}\left(|W - \mu| < \epsilon\right) \geq 1 - \frac{\sigma^2}{\epsilon^2},$$

or, equivalently,

$$\mathbb{P}\left(|W - \mu| \geq \epsilon\right) \leq \frac{\sigma^2}{\epsilon^2}.$$

Proof. ... □

As a consequence of Chebyshev's inequality, we have

Proposition. The sample mean $\widehat{\mu}_n = \frac{1}{n}\sum_{i=1}^n W_i$ is consistent for $\mathbb{E}(W) = \mu$, provided that the population W has finite mean μ and variance σ^2.

Proof.

$$\mathbb{E}(\widehat{\mu}_n) = \mu \qquad \text{and} \qquad \mathrm{Var}(\widehat{\mu}_n) = \frac{\sigma^2}{n}.$$

$$\forall \epsilon > 0, \quad \mathbb{P}\left(|\widehat{\mu}_n - \mu| \leq \epsilon\right) \geq 1 - \frac{\sigma^2}{n\epsilon^2} \to 1.$$

□

E.g. 1. Let $Y_1, \cdots, Y_n$ be a random sample of size n from the uniform pdf $f_Y(y;\theta) = 1/\theta$, $y \in [0,\theta]$. Let $\widehat{\theta}_n = Y_{max}$. We know that Y_{max} is biased. Is it consistent?

Sol. The c.d.f. of Y is equal to $F_Y(y) = y/\theta$ for $y \in [0,\theta]$. Hence,

$$f_{Y_{max}}(y) = nF_Y(y)^{n-1} f_Y(y) = \frac{ny^{n-1}}{\theta^n}, \qquad y \in [0,\theta].$$

Therefore,

$$\begin{aligned}\mathbb{P}(|\widehat{\theta}_n - \theta| < \epsilon) &= \mathbb{P}(\theta - \epsilon < \widehat{\theta}_n < \theta + \epsilon) \\ &= \int_{\theta-\epsilon}^{\theta} \frac{ny^{n-1}}{\theta^n} \mathrm{d}y + \int_{\theta}^{\theta+\epsilon} 0 \mathrm{d}y \\ &= 1 - \left(\frac{\theta - \epsilon}{\theta}\right)^n \\ &\to 0 \quad \text{as } n \to \infty.\end{aligned}$$

□

E.g. 1. Let $Y_1, \cdots, Y_n$ be a random sample of size n from the uniform pdf $f_Y(y;\theta) = 1/\theta$, $y \in [0,\theta]$. Let $\widehat{\theta}_n = Y_{max}$. We know that Y_{max} is biased. Is it consistent?

Sol. The c.d.f. of Y is equal to $F_Y(y) = y/\theta$ for $y \in [0,\theta]$. Hence,

$$f_{Y_{max}}(y) = nF_Y(y)^{n-1} f_Y(y) = \frac{ny^{n-1}}{\theta^n}, \qquad y \in [0,\theta].$$

Therefore,

$$\begin{aligned}
\mathbb{P}(|\widehat{\theta}_n - \theta| < \epsilon) &= \mathbb{P}(\theta - \epsilon < \widehat{\theta}_n < \theta + \epsilon) \\
&= \int_{\theta-\epsilon}^{\theta} \frac{ny^{n-1}}{\theta^n} \mathrm{d}y + \int_{\theta}^{\theta+\epsilon} 0 \mathrm{d}y \\
&= 1 - \left(\frac{\theta - \epsilon}{\theta}\right)^n \\
&\to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

□

E.g. 1. Let $Y_1, \cdots, Y_n$ be a random sample of size n from the uniform pdf $f_Y(y;\theta) = 1/\theta$, $y \in [0,\theta]$. Let $\widehat{\theta}_n = Y_{max}$. We know that Y_{max} is biased. Is it consistent?

Sol. The c.d.f. of Y is equal to $F_Y(y) = y/\theta$ for $y \in [0,\theta]$. Hence,

$$f_{Y_{max}}(y) = nF_Y(y)^{n-1} f_Y(y) = \frac{ny^{n-1}}{\theta^n}, \qquad y \in [0,\theta].$$

Therefore,

$$\begin{aligned}
\mathbb{P}(|\widehat{\theta}_n - \theta| < \epsilon) &= \mathbb{P}(\theta - \epsilon < \widehat{\theta}_n < \theta + \epsilon) \\
&= \int_{\theta-\epsilon}^{\theta} \frac{ny^{n-1}}{\theta^n} \mathrm{d}y + \int_{\theta}^{\theta+\epsilon} 0 \mathrm{d}y \\
&= 1 - \left(\frac{\theta - \epsilon}{\theta}\right)^n \\
&\to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

□

E.g. 1. Let $Y_1, \cdots, Y_n$ be a random sample of size n from the uniform pdf $f_Y(y;\theta) = 1/\theta$, $y \in [0,\theta]$. Let $\widehat{\theta}_n = Y_{max}$. We know that Y_{max} is biased. Is it consistent?

Sol. The c.d.f. of Y is equal to $F_Y(y) = y/\theta$ for $y \in [0,\theta]$. Hence,

$$f_{Y_{max}}(y) = nF_Y(y)^{n-1} f_Y(y) = \frac{ny^{n-1}}{\theta^n}, \qquad y \in [0,\theta].$$

Therefore,

$$\begin{aligned}
\mathbb{P}(|\widehat{\theta}_n - \theta| < \epsilon) &= \mathbb{P}(\theta - \epsilon < \widehat{\theta}_n < \theta + \epsilon) \\
&= \int_{\theta-\epsilon}^{\theta} \frac{ny^{n-1}}{\theta^n} \mathrm{d}y + \int_{\theta}^{\theta+\epsilon} 0 \mathrm{d}y \\
&= 1 - \left(\frac{\theta-\epsilon}{\theta}\right)^n \\
&\to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

□

E.g. 1. Let $Y_1, \cdots, Y_n$ be a random sample of size n from the uniform pdf $f_Y(y;\theta) = 1/\theta$, $y \in [0,\theta]$. Let $\widehat{\theta}_n = Y_{max}$. We know that Y_{max} is biased. Is it consistent?

Sol. The c.d.f. of Y is equal to $F_Y(y) = y/\theta$ for $y \in [0,\theta]$. Hence,

$$f_{Y_{max}}(y) = nF_Y(y)^{n-1} f_Y(y) = \frac{ny^{n-1}}{\theta^n}, \qquad y \in [0,\theta].$$

Therefore,

$$\begin{aligned}
\mathbb{P}(|\widehat{\theta}_n - \theta| < \epsilon) &= \mathbb{P}(\theta - \epsilon < \widehat{\theta}_n < \theta + \epsilon) \\
&= \int_{\theta-\epsilon}^{\theta} \frac{ny^{n-1}}{\theta^n} \mathrm{d}y + \int_{\theta}^{\theta+\epsilon} 0 \mathrm{d}y \\
&= 1 - \left(\frac{\theta - \epsilon}{\theta}\right)^n \\
&\to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

□

E.g. 1. Let $Y_1, \cdots, Y_n$ be a random sample of size n from the uniform pdf $f_Y(y;\theta) = 1/\theta$, $y \in [0,\theta]$. Let $\widehat{\theta}_n = Y_{max}$. We know that Y_{max} is biased. Is it consistent?

Sol. The c.d.f. of Y is equal to $F_Y(y) = y/\theta$ for $y \in [0,\theta]$. Hence,

$$f_{Y_{max}}(y) = nF_Y(y)^{n-1} f_Y(y) = \frac{ny^{n-1}}{\theta^n}, \qquad y \in [0,\theta].$$

Therefore,

$$\begin{aligned}\mathbb{P}(|\widehat{\theta}_n - \theta| < \epsilon) &= \mathbb{P}(\theta - \epsilon < \widehat{\theta}_n < \theta + \epsilon) \\ &= \int_{\theta-\epsilon}^{\theta} \frac{ny^{n-1}}{\theta^n} \mathrm{d}y + \int_{\theta}^{\theta+\epsilon} 0 \mathrm{d}y \\ &= 1 - \left(\frac{\theta-\epsilon}{\theta}\right)^n \\ &\to 0 \quad \text{as } n \to \infty.\end{aligned}$$

□

E.g. 1. Let $Y_1, \cdots, Y_n$ be a random sample of size n from the uniform pdf $f_Y(y;\theta) = 1/\theta$, $y \in [0,\theta]$. Let $\widehat{\theta}_n = Y_{max}$. We know that Y_{max} is biased. Is it consistent?

Sol. The c.d.f. of Y is equal to $F_Y(y) = y/\theta$ for $y \in [0,\theta]$. Hence,

$$f_{Y_{max}}(y) = nF_Y(y)^{n-1} f_Y(y) = \frac{ny^{n-1}}{\theta^n}, \qquad y \in [0,\theta].$$

Therefore,

$$\begin{aligned}
\mathbb{P}(|\widehat{\theta}_n - \theta| < \epsilon) &= \mathbb{P}(\theta - \epsilon < \widehat{\theta}_n < \theta + \epsilon) \\
&= \int_{\theta-\epsilon}^{\theta} \frac{ny^{n-1}}{\theta^n} \mathrm{d}y + \int_{\theta}^{\theta+\epsilon} 0 \mathrm{d}y \\
&= 1 - \left(\frac{\theta-\epsilon}{\theta}\right)^n \\
&\to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

□

E.g. 1. Let $Y_1, \cdots, Y_n$ be a random sample of size n from the uniform pdf $f_Y(y;\theta) = 1/\theta$, $y \in [0,\theta]$. Let $\widehat{\theta}_n = Y_{max}$. We know that Y_{max} is biased. Is it consistent?

Sol. The c.d.f. of Y is equal to $F_Y(y) = y/\theta$ for $y \in [0,\theta]$. Hence,

$$f_{Y_{max}}(y) = nF_Y(y)^{n-1} f_Y(y) = \frac{ny^{n-1}}{\theta^n}, \qquad y \in [0,\theta].$$

Therefore,

$$\begin{aligned}
\mathbb{P}(|\widehat{\theta}_n - \theta| < \epsilon) &= \mathbb{P}(\theta - \epsilon < \widehat{\theta}_n < \theta + \epsilon) \\
&= \int_{\theta-\epsilon}^{\theta} \frac{ny^{n-1}}{\theta^n} \mathrm{d}y + \int_{\theta}^{\theta+\epsilon} 0 \mathrm{d}y \\
&= 1 - \left(\frac{\theta - \epsilon}{\theta}\right)^n \\
&\to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

□

E.g. 2. Suppose $Y_1, Y_2, \cdots, Y_n$ is a random sample from the exponential pdf, $f_Y(y;\lambda) = \lambda e^{-\lambda y}$, $y > 0$. Show that $\widehat{\lambda}_n = Y_1$ is not consistent for λ.

Sol. To prove $\widehat{\lambda}_n$ is not consistent for λ, we need only to find out one $\epsilon > 0$ such that the following limit does not hold:

$$\lim_{n\to\infty} \mathbb{P}\left(|\widehat{\lambda}_n - \lambda| < \epsilon\right) = 1. \tag{3}$$

We can choose $\epsilon = \lambda/m$ for any $m \geq 1$. Then

$$|\widehat{\lambda}_n - \lambda| \leq \frac{\lambda}{m} \quad \Longleftrightarrow \quad \left(1 - \frac{1}{m}\right)\lambda \leq \widehat{\lambda}_n \leq \left(1 + \frac{1}{m}\right)\lambda$$

$$\Longrightarrow \quad \widehat{\lambda}_n \geq \left(1 - \frac{1}{m}\right)\lambda.$$

E.g. 2. Suppose $Y_1, Y_2, \cdots, Y_n$ is a random sample from the exponential pdf, $f_Y(y;\lambda) = \lambda e^{-\lambda y}$, $y > 0$. Show that $\widehat{\lambda}_n = Y_1$ is not consistent for λ.

Sol. To prove $\widehat{\lambda}_n$ is not consistent for λ, we need only to find out one $\epsilon > 0$ such that the following limit does not hold:

$$\lim_{n\to\infty} \mathbb{P}\left(|\widehat{\lambda}_n - \lambda| < \epsilon\right) = 1. \tag{3}$$

We can choose $\epsilon = \lambda/m$ for any $m \geq 1$. Then

$$|\widehat{\lambda}_n - \lambda| \leq \frac{\lambda}{m} \iff \left(1 - \frac{1}{m}\right)\lambda \leq \widehat{\lambda}_n \leq \left(1 + \frac{1}{m}\right)\lambda$$
$$\implies \widehat{\lambda}_n \geq \left(1 - \frac{1}{m}\right)\lambda.$$

E.g. 2. Suppose $Y_1, Y_2, \cdots, Y_n$ is a random sample from the exponential pdf, $f_Y(y;\lambda) = \lambda e^{-\lambda y}$, $y > 0$. Show that $\widehat{\lambda}_n = Y_1$ is not consistent for λ.

Sol. To prove $\widehat{\lambda}_n$ is not consistent for λ, we need only to find out one $\epsilon > 0$ such that the following limit does not hold:

$$\lim_{n\to\infty} \mathbb{P}\left(|\widehat{\lambda}_n - \lambda| < \epsilon\right) = 1. \tag{3}$$

We can choose $\epsilon = \lambda/m$ for any $m \geq 1$. Then

$$|\widehat{\lambda}_n - \lambda| \leq \frac{\lambda}{m} \quad \Longleftrightarrow \quad \left(1 - \frac{1}{m}\right)\lambda \leq \widehat{\lambda}_n \leq \left(1 + \frac{1}{m}\right)\lambda$$

$$\Longrightarrow \quad \widehat{\lambda}_n \geq \left(1 - \frac{1}{m}\right)\lambda.$$

E.g. 2. Suppose $Y_1, Y_2, \cdots, Y_n$ is a random sample from the exponential pdf, $f_Y(y;\lambda) = \lambda e^{-\lambda y}$, $y > 0$. Show that $\widehat{\lambda}_n = Y_1$ is not consistent for λ.

Sol. To prove $\widehat{\lambda}_n$ is not consistent for λ, we need only to find out one $\epsilon > 0$ such that the following limit does not hold:

$$\lim_{n\to\infty} \mathbb{P}\left(|\widehat{\lambda}_n - \lambda| < \epsilon\right) = 1. \tag{3}$$

We can choose $\epsilon = \lambda/m$ for any $m \geq 1$. Then

$$\begin{aligned} |\widehat{\lambda}_n - \lambda| \leq \frac{\lambda}{m} \quad &\Longleftrightarrow \quad \left(1 - \frac{1}{m}\right)\lambda \leq \widehat{\lambda}_n \leq \left(1 + \frac{1}{m}\right)\lambda \\ &\Longrightarrow \quad \widehat{\lambda}_n \geq \left(1 - \frac{1}{m}\right)\lambda. \end{aligned}$$

Hence,

$$\mathbb{P}\left(|\widehat{\lambda}_n - \lambda| < \frac{\lambda}{m}\right) \leq \mathbb{P}\left(\widehat{\lambda}_n \geq \left(1 - \frac{1}{m}\right)\lambda\right)$$
$$= \mathbb{P}\left(Y_1 \geq \left(1 - \frac{1}{m}\right)\lambda\right)$$
$$= \int_{\left(1-\frac{1}{m}\right)\lambda}^{\infty} \lambda e^{-\lambda y} \mathrm{d}y$$
$$= e^{-\left(1-\frac{1}{m}\right)\lambda^2} < 1.$$

Therefore, the limit in (3) cannot hold. □

Hence,

$$\begin{aligned}
\mathbb{P}\left(|\widehat{\lambda}_n - \lambda| < \frac{\lambda}{m}\right) &\le \mathbb{P}\left(\widehat{\lambda}_n \ge \left(1 - \frac{1}{m}\right)\lambda\right) \\
&= \mathbb{P}\left(Y_1 \ge \left(1 - \frac{1}{m}\right)\lambda\right) \\
&= \int_{\left(1-\frac{1}{m}\right)\lambda}^{\infty} \lambda e^{-\lambda y}\mathrm{d}y \\
&= e^{-\left(1-\frac{1}{m}\right)\lambda^2} < 1.
\end{aligned}$$

Therefore, the limit in (3) cannot hold. □

Hence,

$$\begin{aligned}\mathbb{P}\left(|\widehat{\lambda}_n - \lambda| < \frac{\lambda}{m}\right) &\leq \mathbb{P}\left(\widehat{\lambda}_n \geq \left(1 - \frac{1}{m}\right)\lambda\right)\\ &= \mathbb{P}\left(Y_1 \geq \left(1 - \frac{1}{m}\right)\lambda\right)\\ &= \int_{\left(1-\frac{1}{m}\right)\lambda}^{\infty} \lambda e^{-\lambda y}\mathrm{d}y\\ &= e^{-\left(1-\frac{1}{m}\right)\lambda^2} < 1.\end{aligned}$$

Therefore, the limit in (3) cannot hold. □

Hence,

$$\begin{aligned}\mathbb{P}\left(|\widehat{\lambda}_n - \lambda| < \frac{\lambda}{m}\right) &\leq \mathbb{P}\left(\widehat{\lambda}_n \geq \left(1 - \frac{1}{m}\right)\lambda\right) \\ &= \mathbb{P}\left(Y_1 \geq \left(1 - \frac{1}{m}\right)\lambda\right) \\ &= \int_{\left(1-\frac{1}{m}\right)\lambda}^{\infty} \lambda e^{-\lambda y}\mathrm{d}y \\ &= e^{-\left(1-\frac{1}{m}\right)\lambda^2} < 1.\end{aligned}$$

Therefore, the limit in (3) cannot hold. □

Hence,

$$\begin{aligned}\mathbb{P}\left(|\widehat{\lambda}_n - \lambda| < \frac{\lambda}{m}\right) &\leq \mathbb{P}\left(\widehat{\lambda}_n \geq \left(1 - \frac{1}{m}\right)\lambda\right) \\ &= \mathbb{P}\left(Y_1 \geq \left(1 - \frac{1}{m}\right)\lambda\right) \\ &= \int_{\left(1-\frac{1}{m}\right)\lambda}^{\infty} \lambda e^{-\lambda y} \mathrm{d}y \\ &= e^{-\left(1-\frac{1}{m}\right)\lambda^2} < 1.\end{aligned}$$

Therefore, the limit in (3) cannot hold. □