# 3. Bias of Estimators; Jensen's Inequality

## Bias Estimators and an application of Jensen's Inequality

(Caption will be displayed when you start playing the video.)

Start of transcript. Skip to the end.

Once I have several properties of an estimator--

I know that maybe it's asymptotically normal,

and maybe it's consistent--

I'm going to want to talk about it in terms

of how far it is from theta and how much variability it has.

How close is it to theta?

**视频**

下载视频文件

**字幕**

下载 SubRip (.srt) file

下载 Text (.txt) file

## The Expectation of the Average

1/1 point (graded)

Let $X_1, \ldots, X_n \stackrel{iid}{\sim} \mathcal{U}([a, a+1])$ where $a$ is an unknown parameter. Let $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ denote the sample mean. In terms of $a$, what is $\mathbb{E}[\overline{X}_n]$?

$\mathbb{E}\left[\overline{X}_n\right] =$ a+0.5 ✔ **Answer:** a+1/2

$a + 0.5$

**Solution:**

Note that since the $X_i$'s are identically distributed, by linearity of expectation,

$$\mathbb{E}[\overline{X}_n] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}[X_i] = \mathbb{E}[X_1] = a + \frac{1}{2}.$$

提交 你已经尝试了2次（总共可以尝试3次）

ℹ Answers are displayed within the problem

## Computing Bias

1/1 point (graded)

**Recall:** Let $\hat{\theta}_n$ denote an estimator for a true parameter $\theta$. Here $n$ specifies the sample size. The **bias** of $\hat{\theta}_n$ is defined to be

$$\mathbb{E}\left[\hat{\theta}_n\right] - \theta.$$

Let $X_1, \ldots, X_n$ be defined as in the previous question. Compute the bias of the estimator $\overline{X}_n$ with respect to the parameter $a$.

0.5

✔ Answer: .5

**Solution:**

The bias is given by $\mathbb{E}\left[\overline{X}_n\right] - a = 1/2$, where we applied the previous part. Note that this implies that $\overline{X}_n - \frac{1}{2}$ is an unbiased estimator.

提交 你已经尝试了2次（总共可以尝试3次）

Answers are displayed within the problem

**(Optional) Jensen's Inequality**

A function $g : \mathbb{R} \to \mathbb{R}$ is **convex** if for all pairs of real numbers $x_1 < x_2$

$$g\left(tx_1 + (1-t)\,x_2\right) \leq tg\left(x_1\right) + (1-t)\,g\left(x_2\right) \qquad \text{for all } 0 \leq t \leq 1.$$

Geometrically, this means that for $x_1 \geq x \geq x_2$, the graph of $g$ is below the secant line connecting the two points $\left(x_1, g\left(x_1\right)\right)$ and $\left(x_2, g\left(x_2\right)\right)$.

At $x = x_2 - t\left(x_2 - x_1\right) = tx_1 + (1-t)\,x_2$, the $y$-value of the graph of $g$ is $g\left(x\right) = g\left(tx_1 + (1-t)\,x_2\right)$, while the $y$-value of the secant line is $tg\left(x_1\right) + (1-t)\,g\left(x_2\right)$.

Note that for $x_1 = 0, x_2 = 1$, the inequality above can be reinterpreted as follows. Let $X \sim \text{Ber}\left(t\right)$ for some parameter $0 \leq t \geq 1$, then the left and right hand sides of inequality above can be rewritten respectively as:

$$\begin{aligned} g\left(t\left(0\right) + (1-t)\left(1\right)\right) &= g\left(1-t\right) = g\left(\mathbb{E}\left[X\right]\right) \\ tg\left(x_1\right) + (1-t)\,g\left(x_2\right) &= \mathbb{E}\left[g\left(X\right)\right], \end{aligned}$$

and hence the inequality defining convexity of $g$ implies

$$g\left(\mathbb{E}\left[X\right]\right) \leq \mathbb{E}\left[g\left(X\right)\right] \qquad (\text{for any Bernoulli random variable} X.$$

**Jensen's Inequality** generalizes this statement to other random variables. It states that for any random variable $X$, and any convex function $g$,

$$g\left(\mathbb{E}\left[X\right]\right) \leq \mathbb{E}\left[g\left(X\right)\right].$$

Jensen's Inequality is also true for random vectors and convex functions on $\mathbb{R}^n$.

**Memory aid:** To remember which way the inequality goes, remember the special case of the Bernoulli random variable above: the secant line, which is the graph of $\mathbb{E}\left[g\left(X\right)\right]$, is above the graph of $g$, which is the graph of $g\left(\mathbb{E}\left[X\right]\right)$.

(For a proof of Jensen's inequality when $g$ is differential, you may refer to *Additional Theoretical Material* in *Unit 8* inthe course 6.431x Probability–the Science of Uncertainty and Data.)

Hide

## (Optional) Expectation of nonlinear functions and Jensen's Inequality

0 points possible (ungraded)

Let $X$ be a positive random variable with expectation $\lambda$. How does $\mu = \mathbb{E}\left[\frac{1}{X}\right]$ compare to $\frac{1}{\lambda}$?

- In general, $\mu$ and $\lambda$ are not comparable
- $\mu \geq \frac{1}{\lambda}$ ✔
- $\mu \leq \frac{1}{\lambda}$

**Solution:**

Note that the function $x \mapsto \frac{1}{x}$ is a convex function on $(0, \infty)$, hence we can use Jensen's inequality that implies

$$\mathbb{E}\left[f(X)\right] \geq f\left(\mathbb{E}\left[X\right]\right)$$

for all convex functions $f$ to conclude

$$\mu = \mathbb{E}\left[\frac{1}{X}\right] \geq \frac{1}{\mathbb{E}\left[X\right]} = \frac{1}{\lambda}.$$

提交 你已经尝试了1次（总共可以尝试1次）

Answers are displayed within the problem

## 讨论

**主题:** Unit 2 Foundation of Inference:Lecture 4: Parametric Estimation and Confidence Intervals / 3. Bias of Estimators; Jensen's Inequality

显示讨论

认证证书是什么?