

# 10: ASYMPTOTIC PROPERTIES

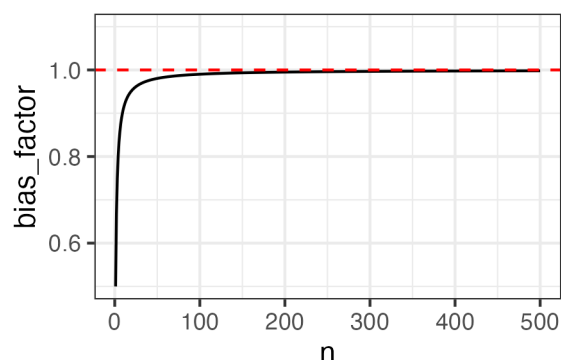
Stat250 S25

Prof Amanda Luby

When we've considered bias and variance of estimators, we've assumed that our data has a fixed sample size. This makes sense in the context of historical statistics: data was time-consuming and expensive to gather, and so experiments were very rigorously designed with a lot of consideration for sample sizes. As data has become easier and cheaper to gather, the *asymptotic* behavior of estimators has also become an important consideration. We may find, for example, that an estimator has a desired behavior *in the limit* that it fails to have for any fixed sample size.

## Asymptotic bias

**Example:** Recall that for  $X_i \sim \text{Unif}(0, \theta)$ ,  $\hat{\theta}_{MLE} = X_{\max}$ . We also showed that the MLE is *biased*:  $E(\hat{\theta}_{MLE}) = \frac{n}{n+1}\theta$ . Does this bias matter?



## Asymptotically unbiased

Let  $Y_1, \dots, Y_n \sim f_y(y|\theta)$ . If  $E(\hat{\theta}_n) \rightarrow \theta$  as  $n \rightarrow \infty$ , then  $\hat{\theta}_n$  is *asymptotically unbiased*

## Consistency

**Example:** Which estimator do you prefer?

Estimator A, n = 10



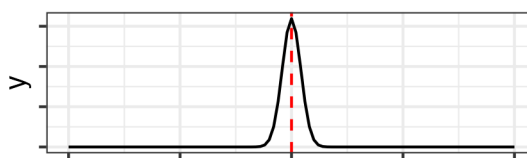
Estimator B, n = 10



Estimator A, n = 100



Estimator B, n = 100



### Consistent estimator

An estimator  $\hat{\theta}_n$  is *consistent* if it converges in probability to  $\theta$ :

**Example:** If an estimator is asymptotically unbiased, does that mean it is also consistent?

**Code Example:** Cauchy vs Normal

### A detour into tail probability inequalities

#### Markov's Inequality

For any random variable  $W$  and any constant  $a$ ,

#### Chebyshev's inequality

Let  $W$  be any random variable with mean  $\mu$  and variance  $\sigma^2$ . For any  $a > 0$ ,

#### Chernoff's inequality

Let  $W$  be any random variable and constants  $a$  and  $t$ ,

**Example:** Let  $X_1, \dots, X_n$  an iid sample of discrete variables from  $p_x(x|\theta)$  where  $E(X_i) = \mu$  and  $\text{Var}(X_i) = \sigma^2 < \infty$ . Let  $\hat{\mu}_n = \frac{1}{n} \sum X_i$ . Is  $\hat{\mu}$  a consistent estimator for  $\mu$ ?

Note:

#### An alternative method to check for consistency

Let  $\{\hat{\theta}_n\}$  be a sequence of estimators for  $\theta$ . If

$$\lim_{n \rightarrow \infty} E[\hat{\theta}_n] \rightarrow \theta \text{ and } \lim_{n \rightarrow \infty} \text{Var}[\hat{\theta}_n] \rightarrow 0$$

**Exercise:** Let  $X_1, \dots, X_n \sim \text{Unif}(0, \theta)$ . Recall that  $\hat{\theta}_{MoM} = 2\bar{X}$  and  $E(\hat{\theta}_{MoM}) = \theta$  and  $\text{Var}(\hat{\theta}_{MoM}) = \frac{\theta^2}{3n}$ . Is  $\hat{\theta}_{MoM}$  consistent for  $\theta$ ? Check using both definitions of consistency.

## Invariance

#### Transformation invariance

An estimation procedure is *transformation invariant* if it yields equivalent results for transformations of parameters. That is, if  $\zeta = h(\theta)$  then  $\hat{\zeta} = h(\hat{\theta})$

**Example:** In class, we showed that the MLEs for  $\mu$  and  $\sigma^2$  in a  $N(\mu, \sigma^2)$  distribution were  $\hat{\mu} = \bar{X}$  and  $\hat{\sigma}^2 = \frac{1}{n} \sum (X_i - \bar{X})^2$ . If I changed my mind and instead want the MLE for the standard error  $\sigma$ , do I have to re-derive the MLE solution?

Note:

## Asymptotic Properties of the MLE

I've mentioned a couple of times in class that the MLE is nice/optimal in some ways. Here's one of the major reasons: under smoothness conditions of  $f_x$ , the sampling distribution of the MLE is *approximately normal*

### Sampling distribution of the MLE

Let  $\hat{\theta}_{MLE} = h(Y)$  be the MLE for  $\theta$ , where  $Y \sim f_y(y|\theta)$ . Then

$$\hat{\theta} \approx N(\theta, \frac{1}{nI(\theta)})$$

**Example:** Use the sampling distribution of the MLE to show that:

1. MLE Estimators are *asymptotically unbiased*
2. Under appropriate smoothness conditions of  $f_x$ , the MLE is *consistent*
3. MLE estimators are *asymptotically efficient*: for large  $n$ , other estimators do not have smaller variance