MIT OpenCourseWare
http://ocw.mit.edu

1.010 Uncertainty in Engineering
Fall 2008

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.

Example Application 19
(Parameter estimation)

COMPARISON OF ESTIMATORS FOR THE UPPER LIMIT OF THE UNIFORM DISTRIBUTION

Consider a random variable X with uniform distribution between 0 and b. The PDF and CDF of X are

$$
\begin{aligned}
f_X(x) &= \begin{cases} 1/b, & \text{for } 0 \le x \le b \\ 0, & \text{elsewhere} \end{cases} \\
F_X(x) &= \begin{cases} 0, & \text{for } x \le 0 \\ x/b, & \text{for } 0 \le x \le b \\ 1, & \text{for } x \ge b \end{cases}
\end{aligned}
\tag{1}
$$

The mean value and variance of X are m = b/2 and $\sigma^2 = b^2/12$.

Suppose that the upper limit b is unknown and needs to be estimated from a random sample $\{X_1, X_2, ..., X_n\}$ from the population of X. Often used methods for problems of this type are the method of moments (MOM), maximum likelihood (ML), and Bayesian estimation. Here we compare these parameter estimation procedures in the context of the present problem.

1. Method of Moments (MOM)

Since the mean value of X is m = b/2, the method of moments produces the following estimator of b:

$$\hat{b}_{MOM} = 2\overline{X} = \frac{2}{n}\sum_{i=1}^{n} X_i \tag{2}$$

This estimator is unbiased, since $E[\hat{b}_{MOM}] = 2m = b$, and has variance $Var[\hat{b}_{MOM}] = \left(\frac{2}{n}\right)^2 n\sigma^2 = \frac{b^2}{3n}$, where we have used $\sigma^2 = b^2/12$.

A problem with the method of moments is that $\hat{b}_{MOM}$ may be smaller than the largest value in the sample, $X_{max} = \max\{X_1, X_2, ..., X_n\}$. In this case it makes no sense to use $\hat{b}_{MOM}$.

2. Maximum Likelihood (ML)

The previous problem can be avoided by using the method of maximum likelihood (ML). The likelihood function has the form

$$\ell(b \mid X_1, X_2, ..., X_n) \propto \prod_{i=1}^{n} f_X(X_i \mid b) \propto \begin{matrix} \frac{1}{b^n}, & \text{for} \quad b \geq X_{max} \\ 0, & \text{for} \quad b < X_{max} \end{matrix} \tag{3}$$

The function in Eq. 3 is maximum for $\hat{b}_{ML} = X_{max}$. Since X_{max} is the maximum of n iid variables with the distribution in Eq. 1, the distribution of this estimator has the following CDF and PDF:

$$\begin{aligned} F_{\hat{b}_{ML}}(x) &= \begin{cases} 0, & \text{for } x \leq 0 \\ (x/b)^n, & \text{for } 0 \leq x \leq b \\ 1, & \text{for } x \geq b \end{cases} \\ f_{\hat{b}_{ML}}(x) &= \begin{cases} \frac{n}{b^n} x^{n-1}, & \text{for } 0 \leq x \leq b \\ 0, & \text{elsewhere} \end{cases} \end{aligned} \tag{4}$$

The mean value and variance of $\hat{b}_{ML}$ are

$$
\begin{aligned}
E[\hat{b}_{ML}] &= \int_0^{\infty} x f_{\hat{b}_{ML}}(x)dx = \int_0^{b} \frac{n}{b^n} x^n dx = \frac{n}{n+1} b \\
Var[\hat{b}_{ML}] &= E[\hat{b}_{ML}^2] - (E[\hat{b}_{ML}])^2 = \int_0^{b} \frac{n}{b^n} x^{n+1} dx - \left(\frac{n}{n+1} b\right)^2 \\
&= \left[\frac{n}{n+2} - \frac{n^2}{(n+1)^2}\right] b^2 = \frac{n}{(n+2)(n+1)^2} b^2
\end{aligned}
\tag{5}
$$

A comparison of the variance prefactors, $\frac{1}{3n}$ for $\hat{b}_{MOM}$ and $\frac{n}{(n+2)(n+1)^2}$ for $\hat{b}_{ML}$, is shown in Figure 1. As one can see, the variance of $\hat{b}_{ML}$ is smaller than that of $\hat{b}_{MOM}$, especially for large n. In this regard, notice that the variance of $\hat{b}_{MOM}$ depends on n like 1/n, whereas (for large n) the variance of $\hat{b}_{ML}$ has an unusual $1/n^2$ behavior.

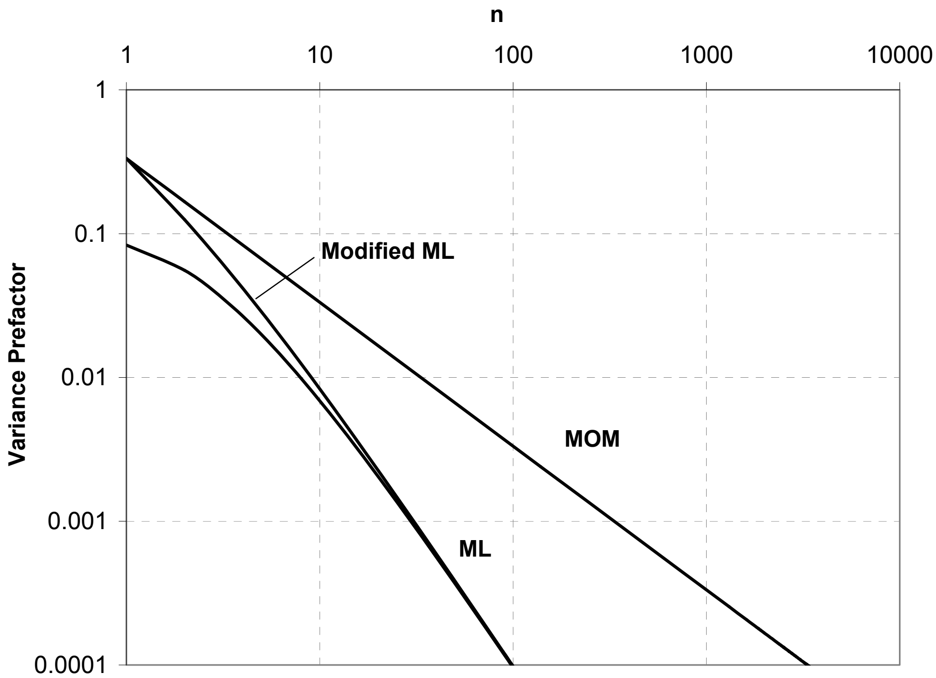

Figure 1. Comparison of the variance prefactors of $\hat{b}_{MOM}$, $\hat{b}_{ML}$, and $\hat{b}'_{ML}$

While the above feature makes $\hat{b}_{ML}$ a more attractive estimator than $\hat{b}_{MOM}$, a drawback of $\hat{b}_{ML}$ is that it is biased. To eliminate the bias, one may use the modified ML estimator

$$\hat{b}'_{ML} = \frac{n+1}{n}\hat{b}_{ML} = \frac{n+1}{n}X_{max} \tag{6}$$

whose mean value and variance are

$$\begin{aligned} &E[\hat{b}'_{ML}] = b \\ &Var[\hat{b}'_{ML}] = \frac{(n+1)^2}{n^2}\frac{n}{(n+2)(n+1)^2}b^2 = \frac{1}{n(n+2)}b^2 \end{aligned} \tag{7}$$

Also the variance prefactor $\frac{1}{n(n+2)}$ in Eq. 7 is shown in Figure 1. As one can see, this prefactor is intermediate between those of $\hat{b}_{MOM}$ and $\hat{b}_{ML}$.

Estimators that are not necessarily unbiased (here $\hat{b}_{MOM}$ and $\hat{b}'_{ML}$ are unbiased, but $\hat{b}_{ML}$ is biased) are often ranked based on the mean square error (MSE), which is the second initial moment of the error $e = \hat{b} - b$. Hence $MSE = E[e^2] = \sigma^2_{\hat{b}} + (E[\hat{b} - b])^2$. For the above estimators, MSE is given by:

$$\begin{aligned} &MSE[\hat{b}_{MOM}] = Var[\hat{b}_{MOM}] = \frac{1}{3n}b^2 \\ &MSE[\hat{b}'_{ML}] = Var[\hat{b}'_{ML}] = \frac{1}{n(n+2)}b^2 \\ &MSE[\hat{b}_{ML}] = Var[\hat{b}_{ML}] + \left(E[\hat{b}_{ML}] - b\right)^2 \\ &\qquad = \left[\frac{n}{(n+2)(n+1)^2} + \frac{1}{(n+1)^2}\right]b^2 = \frac{2}{(n+1)(n+2)}b^2 \end{aligned} \tag{8}$$

The prefactors of b^2 in these MSE expressions are compared in Figure 2. Notice that, except for n = 1, the modified ML estimator $\hat{b}'_{ML}$ has the best performance.

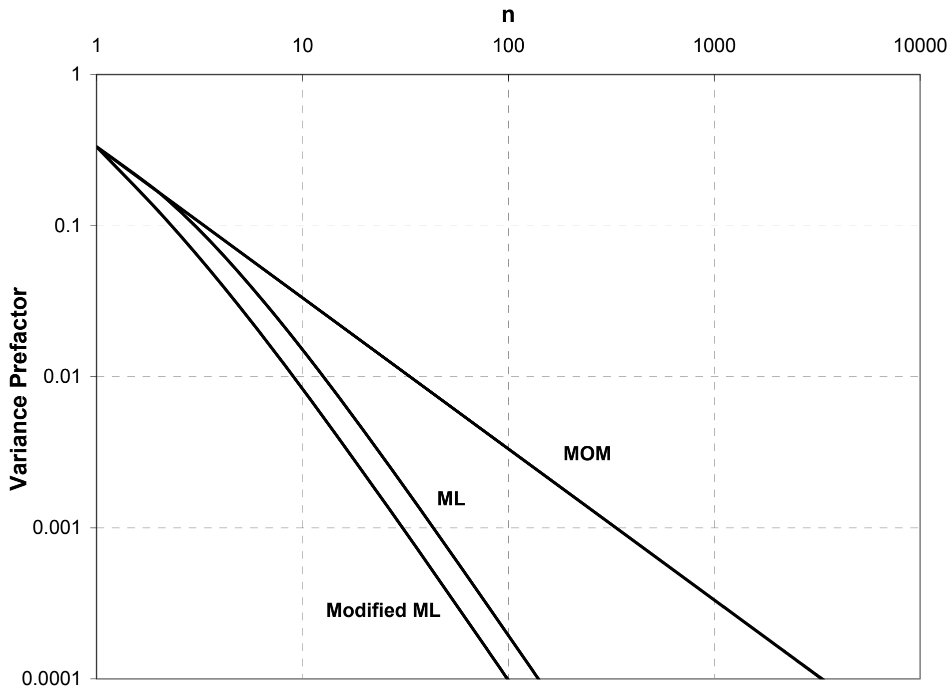

Figure 2. Comparison of the variance prefactors of $\hat{b}_{MOM}$, $\hat{b}_{ML}$, and $\hat{b}'_{ML}$ in the MSE expressions

3. Bayesian Estimation

Bayesian analysis requires specification of a prior distribution for b, say in the form of a prior density $f'(b)$, combination of the prior density with the likelihood function in Eq. 3 to obtain the posterior density $f''(b) \propto f'(b) \cdot l(b)$, and finally the choice of an estimator of b. The estimator may be based on $f''(b)$ (for example one might use the a-posteriori mode or the a-posteriori mean) or calculated by minimizing an expected loss function.

Consider for example a prior density of the gamma type,

$$f'(b) \propto b^{\alpha-1} e^{-b/\beta}, \qquad b \geq 0 \tag{9}$$

where α, β > 0 are parameters. Then the posterior density of b is

$$f''(b) \propto b^{(\alpha-n)-1} e^{-b/\beta}, \qquad \text{for } b \geq X_{max} \tag{10}$$

Ignoring for the moment the constraint $b \geq X_{max}$, the function in Eq. 10 is maximum for $b = \max\{0, \beta(\alpha - n - 1)\}$. We conclude that, if one estimates b as the value that maximizes $f''(b)$, the Bayesian estimator $\hat{b}_{Bayes}$ is given by

$$\hat{b}_{Bayes} = \max\{X_{max}, \beta(\alpha - n - 1)\} \tag{11}$$

It is interesting that $\hat{b}_{Bayes}$ is the same as $\hat{b}_{ML}$ when $\beta(\alpha - n - 1) \leq X_{max}$, which happens for any given α and β if n is sufficiently large.

Problem 19.1

Implement the previous estimators for the following samples:

Sample 1: {5.2, 0.8, 2.7, 3.1, 6.4, 3.8, 1.7, 6.3}

Sample 2: {5.2, 0.8, 2.7, 3.1, 6.4, 3.8, 1.7, 16.3}

For the Bayesian estimator, use α = 20 *and* β = 1 *in the prior. Comment on the results. Do you find anything suspicious in the second sample?*