Estimation and Mean Squared Errors

Matthew Seguin

Importing Libraries

```
library(tidyverse)
library(latex2exp)
```

1.

$$\text{Let } X_1, X_2, ... \overset{\text{iid}}{\sim} Unif(0, \theta).$$

$$\text{Consider } \hat{\theta}_n = max\{X_1, ..., X_n\} \text{ and } \tilde{\theta}_n = 2\bar{X}_n.$$

$$\text{Recall, for } X \sim Unif(a, b) \text{ that:}$$

$$F_X(x) = \mathbb{P}[X \leq x] = \begin{cases} 0 & \text{for} \quad x < a \\ \frac{x-a}{b-a} & \text{for} \quad a \leq x \leq b \\ 1 & \text{for} \quad x > b \end{cases}$$

a.

We can easily find the CDF then use it to find the PDF:

Here $a = 0$ and $b = \theta$ so:

$$F_{X_1}(x) = \mathbb{P}[X_1 \leq x] = \begin{cases} 0 & \text{for} \quad x < 0 \\ \frac{x}{\theta} & \text{for} \quad 0 \leq x \leq \theta \\ 1 & \text{for} \quad x > \theta \end{cases}$$

$$F_{\hat{\theta}_n}(x) = \mathbb{P}[\hat{\theta}_n \leq x] = \mathbb{P}[max\{X_1, ..., X_n\} \leq x] = \mathbb{P}[X_1 \leq x, X_2 \leq x, ..., X_n \leq x]$$

$$= \mathbb{P}[X_1 \leq x]\, \mathbb{P}[X_2 \leq x] \ldots \mathbb{P}[X_n \leq x] \text{ by independence.}$$

Then since the X_i's are identically distributed

$$F_{\hat{\theta}_n}(x) = \mathbb{P}[\hat{\theta}_n \leq x] = \mathbb{P}[X_1 \leq x]\, \mathbb{P}[X_2 \leq x] \ldots \mathbb{P}[X_n \leq x] = \Big(\mathbb{P}[X_1 \leq x]\Big)^n = \Big(F_{X_1}(x)\Big)^n$$

$$= \begin{cases} 0 & \text{for} \quad x < 0 \\ \frac{x^n}{\theta^n} & \text{for} \quad 0 \leq x \leq \theta \\ 1 & \text{for} \quad x > \theta \end{cases}$$

First note that this piecewise function is differentiable in x since it is a polynomial of x for $x \in [0, \theta]$ and the boundary limits are equal to the function value from both sides, that is:

$$\lim_{x \downarrow 0} F_{\hat{\theta}_n}(x) = \frac{x^n}{\theta^n}\Big|_{x=0} = 0 = F_{\hat{\theta}_n}(0) = 0 = 0\Big|_{x=0} = \lim_{x \uparrow 0} F_{\hat{\theta}_n}(x)$$

and

$$\lim_{x \downarrow \theta} F_{\hat{\theta}_n}(x) = 1\Big|_{x=\theta} = 1 = F_{\hat{\theta}_n}(\theta) = 1 = \frac{x^n}{\theta^n}\Big|_{x=\theta} = \lim_{x \uparrow \theta} F_{\hat{\theta}_n}(x)$$

So $F_{\hat{\theta}_n}(x)$ is differentiable in x on $(-\infty, 0)$ and (θ, ∞) since it is constant there, differentiable in x on $(0, \theta)$ since it is a polynomial there, and still differentiable in x at $x = 0$ and $x = \theta$ from the results above.

Therefore $F_{\hat{\theta}_n}(x)$ is differentiable in x on $\mathbb{R}$.

That was the CDF of $\hat{\theta}_n$ so to find the PDF we can take the derivative with respect to x.

$$f_{\hat{\theta}_n}(x) = \frac{\partial}{\partial x} F_{\hat{\theta}_n}(x) = \begin{cases} \frac{\partial}{\partial x} 0 & \text{for} \quad x < 0 \\ \frac{\partial}{\partial x} \frac{x^n}{\theta^n} & \text{for} \quad 0 \leq x \leq \theta \\ \frac{\partial}{\partial x} 1 & \text{for} \quad x > \theta \end{cases} = \begin{cases} 0 & \text{for} \quad x < 0 \\ \frac{n x^{n-1}}{\theta^n} & \text{for} \quad 0 \leq x \leq \theta \\ 0 & \text{for} \quad x > \theta \end{cases}$$

b.

From the result of the previous problem we know $\hat{\theta}_n$ has density $f_{\hat{\theta}_n}(x) = \dfrac{nx^{n-1}}{\theta^n}$ when $x \in [0, \theta]$.

Finding bias:

$$\mathbb{E}[\hat{\theta}_n] = \int_{-\infty}^{\infty} x f_{\hat{\theta}_n}(x)\, dx = \int_0^\theta x \frac{nx^{n-1}}{\theta^n}\, dx = \frac{n}{\theta^n}\int_0^\theta x^n\, dx = \frac{n}{\theta^n}\left(\frac{x^{n+1}}{n+1}\Big|_0^\theta\right) = \frac{n}{\theta^n}\Big(\frac{\theta^{n+1}}{n+1}\Big) = \frac{n}{n+1}\theta$$

$$\text{Therefore } \mathbf{BIAS}[\hat{\theta}_n] = \mathbb{E}[\hat{\theta}_n] - \theta = \frac{n}{n+1}\theta - \theta = -\frac{\theta}{n+1} \;\square$$

Finding standard error:

$$\mathbb{E}[(\hat{\theta}_n)^2] = \int_{-\infty}^{\infty} x^2 f_{\hat{\theta}_n}(x)\, dx = \int_0^\theta x^2 \frac{nx^{n-1}}{\theta^n}\, dx = \frac{n}{\theta^n}\int_0^\theta x^{n+1}\, dx = \frac{n}{\theta^n}\left(\frac{x^{n+2}}{n+2}\Big|_0^\theta\right) = \frac{n}{\theta^n}\Big(\frac{\theta^{n+2}}{n+2}\Big) = \frac{n}{n+2}\theta^2$$

Then we know

$$\mathbb{V}[\hat{\theta}_n] = \mathbb{E}[(\hat{\theta}_n)^2] - \Big(\mathbb{E}[\hat{\theta}_n]\Big)^2 = \frac{n}{n+2}\theta^2 - \frac{n^2}{(n+1)^2}\theta^2 = \theta^2\Big(\frac{n(n+1)^2}{(n+2)(n+1)^2} - \frac{n^2(n+2)}{(n+2)(n+1)^2}\Big)$$

$$= \theta^2\Big(\frac{n^3 + 2n^2 + n - n^3 - 2n^2}{(n+2)(n+1)^2}\Big) = \theta^2\frac{n}{(n+2)(n+1)^2}$$

$$\text{Therefore } \mathbf{SE}[\hat{\theta}_n] = \sqrt{\mathbb{V}[\hat{\theta}_n]} = \sqrt{\theta^2\tfrac{n}{(n+2)(n+1)^2}} = \frac{\theta}{n+1}\sqrt{\frac{n}{(n+2)}} \;\square$$

Finding mean squared error:

Note for any estimator $\hat{X}$ of a parameter x that

$$\mathbf{MSE}[\hat{X}] = \mathbb{E}[(\hat{X} - x)^2] = \mathbb{E}[(\hat{X} - \mathbb{E}[\hat{X}] + \mathbb{E}[\hat{X}] - x)^2]$$
$$= \mathbb{E}[(\hat{X} - \mathbb{E}[\hat{X}])^2 + 2(\hat{X} - \mathbb{E}[\hat{X}])(\mathbb{E}[\hat{X}] - x) + (\mathbb{E}[\hat{X}] - x)^2]$$
$$= \mathbb{E}[(\hat{X} - \mathbb{E}[\hat{X}])^2] + 2\mathbb{E}[(\hat{X} - \mathbb{E}[\hat{X}])(\mathbb{E}[\hat{X}] - x)] + \mathbb{E}[(\mathbb{E}[\hat{X}] - x)^2]$$
$$= \mathbb{V}[\hat{X}] + 2(\mathbb{E}[\hat{X}] - x)(\mathbb{E}[\hat{X} - \mathbb{E}[\hat{X}]]) + (\mathbb{E}[\hat{X}] - x)^2$$
$$= \mathbb{V}[\hat{X}] + 2(\mathbb{E}[\hat{X}] - x)(\mathbb{E}[\hat{X}] - \mathbb{E}[\mathbb{E}[\hat{X}]]) + \big(\mathbf{BIAS}[\hat{X}]\big)^2$$
$$= \mathbb{V}[\hat{X}] + 2(\mathbb{E}[\hat{X}] - x)(\mathbb{E}[\hat{X}] - \mathbb{E}[\hat{X}]) + \big(\mathbf{BIAS}[\hat{X}]\big)^2 = \mathbb{V}[\hat{X}] + \big(\mathbf{BIAS}[\hat{X}]\big)^2$$

$$\mathbf{MSE}[\hat{\theta}_n] = \mathbb{V}[\hat{\theta}_n] + \big(\mathbf{BIAS}[\hat{\theta}_n]\big)^2 = \theta^2\frac{n}{(n+2)(n+1)^2} + \Big(-\frac{\theta}{n+1}\Big)^2 = \theta^2\frac{n}{(n+2)(n+1)^2} + \theta^2\frac{1}{(n+1)^2}$$

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

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

c.

Finding bias:

$$\mathbb{E}[X_1] = \int_{-\infty}^{\infty} x f_{X_1}(x)\,dx = \int_0^{\theta} \frac{x}{\theta}\,dx = \frac{1}{\theta}\Big(\frac{x^2}{2}\Big|_0^{\theta}\Big) = \Big(\frac{1}{\theta}\Big)\Big(\frac{\theta^2}{2}\Big) = \frac{\theta}{2}$$

Recall the linearity of expectation and that $X_1, X_2, ...$ are iid.

$$\mathbb{E}[\tilde{\theta}_n] = \mathbb{E}[2\bar{X}_n] = 2\mathbb{E}[\frac{X_1 + ... + X_n}{n}] = \frac{2}{n}\mathbb{E}[X_1 + ... + X_n] = \frac{2}{n}\Big(\mathbb{E}[X_1] + ... + \mathbb{E}[X_n]\Big) = \frac{2}{n}\Big(n\mathbb{E}[X_1]\Big)$$

$$= 2\mathbb{E}[X_1] = 2\frac{\theta}{2} = \theta$$

$$\text{Therefore } \mathbf{BIAS}[\tilde{\theta}_n] = \mathbb{E}[\tilde{\theta}_n] - \theta = \theta - \theta = 0 \ \square$$

Finding standard error:

$$\mathbb{E}[(X_1)^2] = \int_{-\infty}^{\infty} x^2 f_{X_1}(x)\,dx = \int_0^{\theta} \frac{x^2}{\theta}\,dx = \frac{1}{\theta}\Big(\frac{x^3}{3}\Big|_0^{\theta}\Big) = \Big(\frac{1}{\theta}\Big)\Big(\frac{\theta^3}{3}\Big) = \frac{\theta^2}{3}$$

Recall that for independent variables P and Q that $Var(P+Q) = Var(P) + Var(Q)$ and that $X_1, X_2, ...$ are iid.

$$\mathbb{V}[\tilde{\theta}_n] = \mathbb{V}[2\bar{X}_n] = 4\mathbb{V}[\frac{X_1 + ... + X_n}{n}] = \frac{4}{n^2}\mathbb{V}[X_1 + ... + X_n] = \frac{4}{n^2}\Big(\mathbb{V}[X_1] + ... + \mathbb{V}[X_n]\Big) = \frac{4}{n^2}\Big(n\mathbb{V}[X_1]\Big)$$

$$= \frac{4}{n}\mathbb{V}[X_1] = \frac{4}{n}\Big(\mathbb{E}[(X_1)^2] - (\mathbb{E}[X_1])^2\Big) = \frac{4}{n}\Big(\frac{\theta^2}{3} - \frac{\theta^2}{4}\Big) = \Big(\frac{4}{n}\Big)\Big(\frac{\theta^2}{12}\Big) = \frac{\theta^2}{3n}$$

$$\text{Therefore } \mathbf{SE}[\tilde{\theta}_n] = \sqrt{\mathbb{V}[\tilde{\theta}_n]} = \sqrt{\tfrac{\theta^2}{3n}} = \frac{\theta}{\sqrt{3n}} \ \square$$

Finding mean squared error:

Note for any estimator $\hat{X}$ of a parameter x that

$$\mathbf{MSE}[\hat{X}] = \mathbb{E}[(\hat{X} - x)^2] = \mathbb{E}[(\hat{X} - \mathbb{E}[\hat{X}] + \mathbb{E}[\hat{X}] - x)^2]$$

$$= \mathbb{E}[(\hat{X} - \mathbb{E}[\hat{X}])^2 + 2(\hat{X} - \mathbb{E}[\hat{X}])(\mathbb{E}[\hat{X}] - x) + (\mathbb{E}[\hat{X}] - x)^2]$$

$$= \mathbb{E}[(\hat{X} - \mathbb{E}[\hat{X}])^2] + 2\mathbb{E}[(\hat{X} - \mathbb{E}[\hat{X}])(\mathbb{E}[\hat{X}] - x)] + \mathbb{E}[(\mathbb{E}[\hat{X}] - x)^2]$$

$$= \mathbb{V}[\hat{X}] + 2(\mathbb{E}[\hat{X}] - x)(\mathbb{E}[\hat{X} - \mathbb{E}[\hat{X}]]) + (\mathbb{E}[\hat{X}] - x)^2$$

$$= \mathbb{V}[\hat{X}] + 2(\mathbb{E}[\hat{X}] - x)(\mathbb{E}[\hat{X}] - \mathbb{E}[\mathbb{E}[\hat{X}]]) + \Big(\mathbf{BIAS}[\hat{X}]\Big)^2$$

$$= \mathbb{V}[\hat{X}] + 2(\mathbb{E}[\hat{X}] - x)(\mathbb{E}[\hat{X}] - \mathbb{E}[\hat{X}]) + \Big(\mathbf{BIAS}[\hat{X}]\Big)^2 = \mathbb{V}[\hat{X}] + \Big(\mathbf{BIAS}[\hat{X}]\Big)^2$$

$$\mathbf{MSE}[\tilde{\theta}_n] = \mathbb{V}[\tilde{\theta}_n] + \Big(\mathbf{BIAS}[\tilde{\theta}_n]\Big)^2 = \frac{\theta^2}{3n} + 0^2 = \frac{\theta^2}{3n} \ \square$$

d.

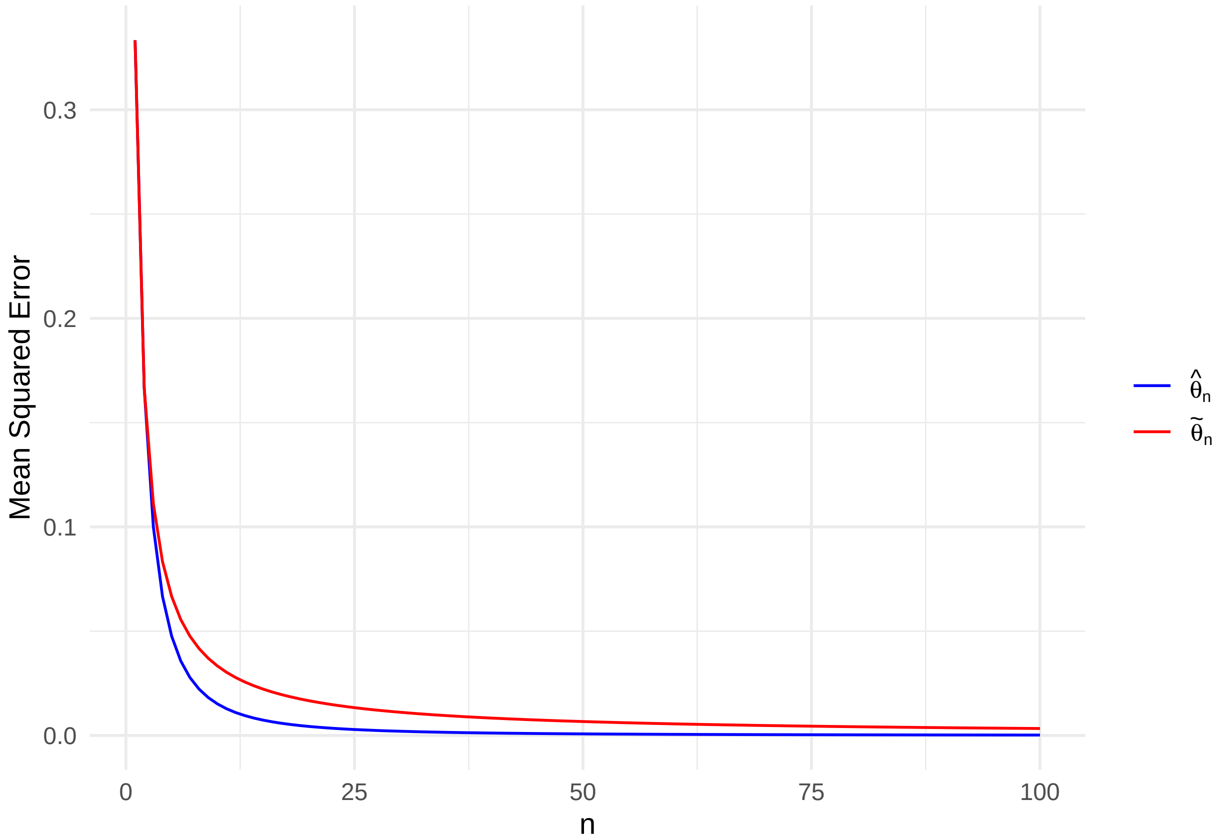
Here we will plot the mean squared error of both $\hat{\theta}_n$ and $\widetilde{\theta}_n$ after fixing $\theta = 1$:

```r
mse_theta_hat <- function(theta, n){
  return((2*theta^2)/((n+2)*(n+1)))
}
mse_theta_tilde <- function(theta, n){
  return((theta^2)/(3*n))
}

data <- data.frame(n = 1:100)

data <- data %>%
  mutate(mse_hat = mse_theta_hat(1, n),
         mse_tilde = mse_theta_tilde(1, n)
         ) %>%
  gather()

graph_df <- data.frame(n = c(filter(data, key == "n")$value,
                             filter(data, key == "n")$value),
                       mse = filter(data, key != "n")$value,
                       group = filter(data, key != "n")$key
                       )
graph_df %>%
  ggplot(aes(x = n,
             y = mse,
             col = group)) +
    geom_line(aes(group = group),
              linewidth = 0.5
              ) +
    labs(x = "n",
         y = "Mean Squared Error",
         col = ""
         ) +
    scale_color_manual(values = c(mse_hat = "blue",
                                  mse_tilde = "red"
                                  ),
                       labels = c(mse_hat = TeX("$\\hat{\\theta}_n$"),
                                  mse_tilde = TeX("$\\tilde{\\theta}_n$")
                                  )
                       ) +
    theme_minimal()
```



We can clearly see that $\hat{\theta}_n$ has lower mean squared error over essentially all values of n. Although $\tilde{\theta}_n$ is unbiased and $\hat{\theta}_n$ is not, we would still prefer $\hat{\theta}_n$ over $\tilde{\theta}_n$ due to the lower mean squared error it provides.

2.

Recall that for disjoint events A and B that $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B]$.

We know $(F \cap G) \cup (F \cap G)^C = \Omega$. Clearly $F \cap G$ and $(F \cap G)^C$ are disjoint.

Therefore $\mathbb{P}[(F \cap G) \cup (F \cap G)^C] = \mathbb{P}[F \cap G] + \mathbb{P}[(F \cap G)^C] = \mathbb{P}[F \cap G] + \mathbb{P}[F^C \cup G^C] = \mathbb{P}[\Omega] = 1$

Therefore $\mathbb{P}[F \cap G] = 1 - \mathbb{P}[(F \cap G)^C] = 1 - \mathbb{P}[F^C \cup G^C]$

Let $X_1, X_2, ... \overset{\text{iid}}{\sim} Unif(0, \theta)$.

Consider $\hat{\theta}_n = max\{X_1, ..., X_n\}$ and the confidence interval for θ given by $C_n = [a\hat{\theta}_n, b\hat{\theta}_n]$.

Recall that:

$$F_{\hat{\theta}_n}(x) = \mathbb{P}[\hat{\theta}_n \leq x] = \begin{cases} 0 & \text{for} \quad x < 0 \\ \frac{x^n}{\theta^n} & \text{for} \quad 0 \leq x \leq \theta \\ 1 & \text{for} \quad x > \theta \end{cases}$$

Then:

$$\mathbb{P}[\theta \in C_n] = \mathbb{P}[\theta \in [a\hat{\theta}_n, b\hat{\theta}_n]] = \mathbb{P}[a\hat{\theta}_n \leq \theta \leq b\hat{\theta}_n] = \mathbb{P}[a\hat{\theta}_n \leq \theta,\ \theta \leq b\hat{\theta}_n]$$

$$= 1 - \mathbb{P}[a\hat{\theta}_n > \theta \text{ or } \theta > b\hat{\theta}_n] = 1 - \mathbb{P}[\hat{\theta}_n < \theta/b \text{ or } \hat{\theta}_n > \theta/a]$$

Since $a < b$ (taking $a > 0$) we know $\dfrac{1}{a} > \dfrac{1}{b}$ so $\dfrac{\theta}{b} < \dfrac{\theta}{a}$. Hence the events $\hat{\theta}_n < \dfrac{\theta}{b}$ and $\hat{\theta}_n > \dfrac{\theta}{a}$ are disjoint as shown:

If $\hat{\theta}_n < \dfrac{\theta}{b} < \dfrac{\theta}{a}$ then $\hat{\theta}_n$ can not be greater than $\dfrac{\theta}{a}$ as well.

If $\hat{\theta}_n > \dfrac{\theta}{a} > \dfrac{\theta}{b}$ then $\hat{\theta}_n$ can not be less than $\dfrac{\theta}{b}$ as well.

Therefore we know:

$$\mathbb{P}[\theta \in C_n] = 1 - \mathbb{P}[\hat{\theta}_n < \theta/b \text{ or } \hat{\theta}_n > \theta/a] = 1 - \mathbb{P}[\hat{\theta}_n < \theta/b] - \mathbb{P}[\hat{\theta}_n > \theta/a] = 1 - \mathbb{P}[\hat{\theta}_n \leq \theta/b] - \mathbb{P}[\hat{\theta}_n > \theta/a]$$

$$= 1 - \mathbb{P}[\hat{\theta}_n \leq \theta/b] - (1 - \mathbb{P}[\hat{\theta}_n \leq \theta/a]) = \mathbb{P}[\hat{\theta}_n \leq \theta/a] - \mathbb{P}[\hat{\theta}_n \leq \theta/b] = F_{\hat{\theta}_n}\left(\frac{\theta}{a}\right) - F_{\hat{\theta}_n}\left(\frac{\theta}{b}\right)$$

$$= \frac{(\theta/a)^n}{\theta^n} - \frac{(\theta/b)^n}{\theta^n} = \frac{1}{a^n} - \frac{1}{b^n} \ \square$$

First note the coverage above depends only on a, b, and n as desired.

If $a = 1$ and we want $\mathbb{P}[\theta \in C_n] = 0.95$ we need:

$0.95 = \mathbb{P}[\theta \in C_n] = \dfrac{1}{a^n} - \dfrac{1}{b^n} = 1 - \dfrac{1}{b^n}$. Which is equivalently $0.05 = \dfrac{1}{b^n}$ and $b^n = \frac{1}{0.05} = 20$ and finally $b = \sqrt[n]{20}$ $\square$