Kyle McGraw

ACM 157 Set 3

1. $X_1, \dots, X_n \sim F \in \mathcal{F}$ $\quad \hat{F}_n \quad \hat{F}_n(x) \quad \hat{F}_n(y) \quad x \neq y$

$$\mathrm{Cov}(\hat{F}_n(x), \hat{F}_n(y)) = \mathbb{E}[\hat{F}_n(x) \cdot \hat{F}_n(y)] - \mathbb{E}[\hat{F}_n(x)] \cdot \mathbb{E}[\hat{F}_n(y)]$$

$$= \mathbb{E}[\hat{F}_n(x) \cdot \hat{F}_n(y)] - F(x) \cdot F(y)$$

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} H(x - X_i) \qquad \hat{F}_n(y) = \frac{1}{n} \sum_{j=1}^{n} H(y - X_j)$$

$$\mathbb{E}[\hat{F}_n(x) \cdot \hat{F}_n(y)] = \mathbb{E}\left[\frac{1}{n^2} \sum_{i=1}^{n} H(x - X_i) \sum_{j=1}^{n} H(y - X_j)\right]$$

If $i \neq j$ they are independent so we get $F(x) \cdot F(y)$
If $i = j$ we get whichever of x or y is smaller
because the heaviside function is 0 for values larger

$$\mathbb{E}[\hat{F}_n(x) \cdot \hat{F}_n(y)] = \begin{cases} \frac{1}{n^2}(n(n-1)F(x)F(y) + nF(y)) & x > y \\ \frac{1}{n^2}(n(n-1)F(x)F(y) + nF(x)) & x \leq y \end{cases}$$

$$\mathrm{Cov}(\hat{F}_n(x), \hat{F}_n(y)) = \begin{cases} \frac{1}{n^2}(n(n-1)F(x)F(y) + nF(y)) - F(x)F(y) & x > y \\ \frac{1}{n^2}(n(n-1)F(x)F(y) + nF(x)) - F(x)F(y) & x \leq y \end{cases}$$

$$= \begin{cases} \frac{n-1}{n}F(x)F(y) + \frac{1}{n}F(y) - F(x)F(y) & x > y \\ \frac{n-1}{n}F(x)F(y) + \frac{1}{n}F(x) - F(x)F(y) & x \leq y \end{cases}$$

$$= \begin{cases} \frac{F(y) - F(x)F(y)}{n} & x > y \\ \frac{F(x) - F(x)F(y)}{n} & x \leq y \end{cases}$$

2. $$\kappa_F = \frac{\int (x - \mu_F)^3 \, dF(x)}{\left(\int (x - \mu_F)^2 \, dF(x)\right)^{3/2}} \qquad F(x) \to \hat{F}_n(x) \quad \hat{\mu}_n \to \bar{X}_n$$

$$= \frac{\int (x - \bar{X}_F)^3 \, d\hat{F}_n(x)}{\left(\int (x - \bar{X}_F)^2 \, d\hat{F}_n(x)\right)^{3/2}}$$

$$= \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X}_n)^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X}_n)^2\right)^{3/2}} \qquad \hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X}_n)^2$$

$$= \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X}_n)^3}{\hat{\sigma}_n^3}$$

4. $X_1, \dots, X_n \quad U[0, \theta] \qquad \mathbb{E}[X_{(k)}] = \frac{k\theta}{n+1}$

a) $\theta = \min\{x : F(x) = 1\}$

$$\hat{\theta}_n = \min\{x : \hat{F}_n(x) = 1\}$$

$$= \min\{x : \frac{1}{n}\sum_{i=1}^{n} H(x - X_i) = 1\}$$

since $\frac{1}{n}\sum_{i=1}^{n} H(x-X_i) = 1$ only if $x \geq X_i \; \forall X_i$

$$= \min\{x : x \geq X_{(n)}\}$$

$$= X_{(n)}$$

b)

$$\text{bias}[\hat{\theta}_n] = \mathbb{E}[\hat{\theta}_n] - \theta = \mathbb{E}[X_{(n)}] - \theta = \frac{n\theta}{n+1} - \theta = -\frac{\theta}{n+1}$$

c)

$$\hat{\theta}_n^J = n\hat{\theta}_n - (n-1)\bar{\theta}_n^J$$

$\hat{\theta}_n^{(-i)} = X_{(n)}$ except for the replication omitting $X_{(n)}$ where it is $X_{(n-1)}$

$$\bar{\theta}_n^J = \frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_n^{(-i)} = \frac{1}{n}\left((n-1)X_{(n)} + X_{(n-1)}\right)$$

$$\hat{\theta}_n^J = nX_{(n)} - (n-1)\frac{1}{n}\left((n-1)X_{(n)} + X_{(n-1)}\right)$$

$$= nX_{(n)} - \frac{(n-1)^2}{n}X_{(n)} - \frac{n-1}{n}X_{(n-1)}$$

$$= \frac{2n-1}{n}X_{(n)} - \frac{n-1}{n}X_{(n-1)}$$

d)

$$\mathbb{B}[\hat{\theta}_n^J] = \mathbb{E}[\hat{\theta}_n] - \mathbb{E}[\hat{\mathbb{B}}_J[\hat{\theta}_n]] - \theta$$

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

$$= -\frac{\theta}{n+1} - \mathbb{E}\left[(n-1)\left(\frac{1}{n}\left((n-1)X_{(n)} + X_{(n-1)}\right) - X_{(n)}\right)\right]$$

$$= -\frac{\theta}{n+1} - \mathbb{E}\left[\frac{(n-1)^2}{n}X_{(n)}\right] - \mathbb{E}\left[\frac{n-1}{n}X_{(n-1)}\right] + \mathbb{E}[(n-1)X_{(n)}]$$

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

$$= \frac{1}{n+1}\left(-\theta - (n-1)^2\theta - \frac{(n-1)^2}{n}\theta + (n-1)n\theta\right)$$

$$= \frac{1}{n+1}\left(-\theta - n^2\theta + 2n\theta - \theta - n\theta + 2\theta - \frac{1}{n}\theta + n^2\theta - n\theta\right)$$

$$= \frac{1}{n+1}\left(-\frac{1}{n}\theta\right)$$

$$= -\frac{\theta}{n(n+1)}$$

5. $X_1, \dots, X_n \sim N(\mu, \sigma^2)$ $\sigma^2 = 1$ $\theta = e^{\mu}$ $\hat{\theta}_n = e^{\bar{X}_n}$

a)

$$\mathbb{B}[\hat{\theta}_n] = \mathbb{E}[\hat{\theta}_n] - \theta = \mathbb{E}[e^{\bar{X}_n}] - e^{\mu}$$

$$\mathbb{E}[e^{\bar{X}_n}] = \mathbb{E}[e^{\frac{1}{n}\sum_{i=1}^{n} x_i}] = \mathbb{E}[e^{\frac{1}{n}X_1} \cdot \ldots \cdot e^{\frac{1}{n}X_n}] = \mathbb{E}[e^{\frac{1}{n}X_1}] \cdot \ldots \cdot \mathbb{E}[e^{\frac{1}{n}X_n}]$$

$$= e^{\frac{\mu}{n} + \frac{\sigma^2}{2n^2}} \cdot \ldots \cdot e^{\frac{\mu}{n} + \frac{\sigma^2}{2n^2}} = e^{\mu + \frac{\sigma^2}{2n}} = e^{\mu + \frac{1}{2n}} = e^{\mu} e^{\frac{1}{2n}}$$

$$\mathbb{B}[\hat{\theta}_n] = e^{\mu}\left(e^{\frac{1}{2n}} - 1\right)$$

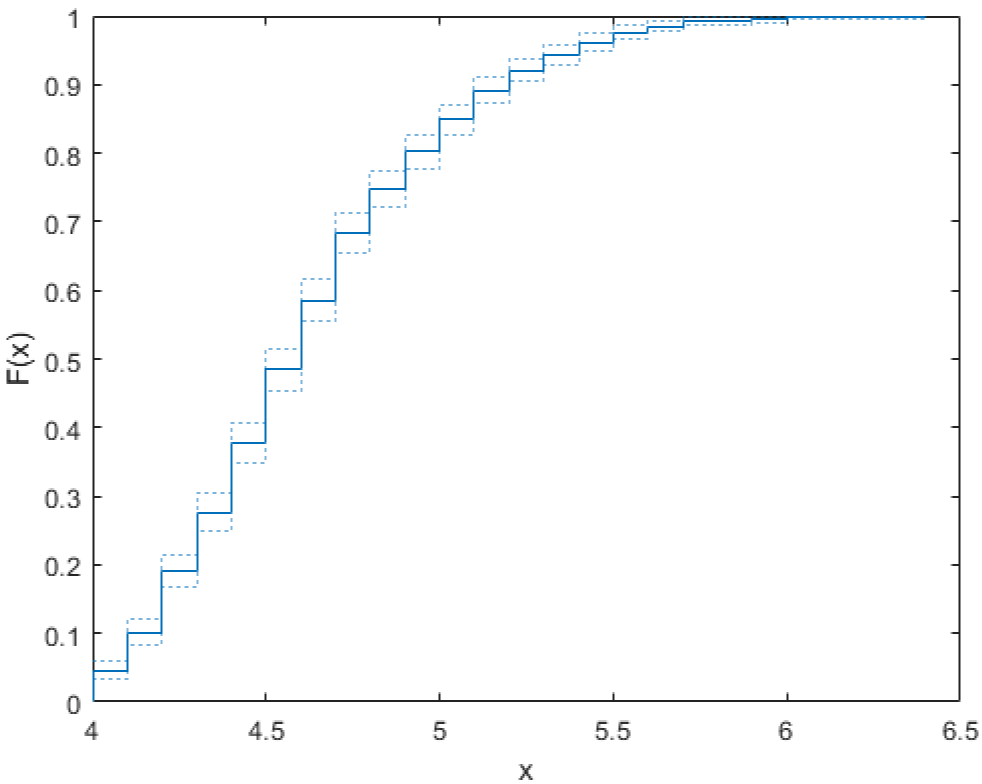
$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

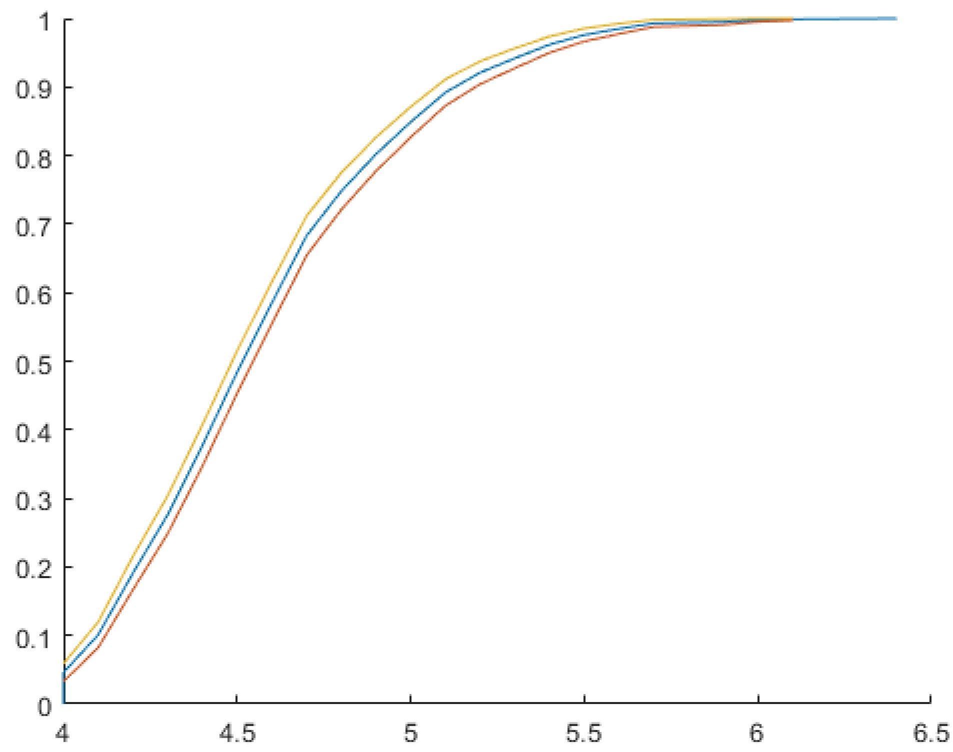
$$= e^{\mu}\left(1 + \frac{1}{2n} + \frac{1}{8n^2} + O\left(\frac{1}{n^3}\right) - 1\right)$$

$$= \frac{e^{\mu}/2}{n} + \frac{e^{\mu}/8}{n^2} + O\left(\frac{1}{n^3}\right)$$

```
A = importdata("fiji.txt");
mags = A(:,5);
n = length(mags);
a = 1 - 0.95;
figure
ecdf(mags,'Alpha',a,'Bounds','on')

[f,x,flo,fup] = ecdf(mags,'Alpha',a);
figure
hold on
plot(x,f)
plot(x,flo)
plot(x,fup)
hold off
```





Published with MATLAB® R2021a

```
mu = 5;
n = 100;
theta = exp(mu);
data = randn([n 1])+mu;

total = 0;
for i=1:n
    dataS = cat(1,data(1:i-1),data(i+1:n));
    total = total+exp(mean(dataS));
end

thetaJ = total/n;
thetaH = exp(mean(data));
bias = (n-1)*(thetaJ-thetaH);
realBias = exp(mu+(1/(2*n))) - theta;
disp([bias realBias])
% (b) The jackknife bias is close to the real bias of 0.7439
% Jackknife bias was 0.8624 in the last run

r = 10^4;
totalN = 0;
totalJ = 0;

for i=1:r
    data = randn([n 1])+mu;
    total = 0;
    for j=1:n
        dataS = cat(1,data(1:j-1),data(j+1:n));
        total = total + exp(mean(dataS));
    end
    thetaB = total/n;
    thetaN = exp(mean(data));
    thetaJ = n*thetaN - (n-1)*thetaB;
    totalN = totalN + thetaN;
    totalJ = totalJ + thetaJ;
end

B1 = totalN/r - theta;
B2 = totalJ/r - theta;
disp([B1 B2])
% (c) B1 is close to what we see for the bias but B2 is consistently
 lower
% than B1 and the bias.
% In the last run B1 was 0.7376 and B2 was -0.0090
%
```

```
    0.8316    0.7439

    0.8349    0.0883
```

Published with MATLAB® R2021a