Kyle McGraw

ACM 157 Set 4

2.

$X_1, \dots, X_n \sim U[\alpha, \beta] \quad \alpha < \beta$

a)

$m_q(\alpha,\beta) = \hat{m}_q \qquad \hat{m}_1 = \frac{1}{n}\sum_{i=1}^n X_i \quad \hat{m}_2 = \frac{1}{n}\sum_{i=1}^n X_i^2$

$m_1(\alpha,\beta) = \mathbb{E}[x] = \int_\alpha^\beta x \frac{1}{\beta-\alpha}\,dx = \left(\frac{x^2}{2}\frac{1}{\beta-\alpha}\right)\Big|_\alpha^\beta = \frac{\beta^2-\alpha^2}{2(\beta-\alpha)} = \frac{\beta+\alpha}{2}$

$m_2(\alpha,\beta) = \mathbb{E}[x^2] = \int_\alpha^\beta x^2 \frac{1}{\beta-\alpha}\,dx = \left(\frac{x^3}{3}\frac{1}{\beta-\alpha}\right)\Big|_\alpha^\beta = \frac{\beta^3-\alpha^3}{3(\beta-\alpha)} = \frac{\alpha^2+\alpha\beta+\beta^2}{3}$

$\hat{m}_1 = \frac{\alpha+\beta}{2} \qquad \hat{m}_2 = \frac{\alpha^2+\alpha\beta+\beta^2}{3}$

$\alpha = \hat{m}_1 - \sqrt{3(\hat{m}_2 - \hat{m}_1^2)} \qquad \beta = \hat{m}_1 + \sqrt{3(\hat{m}_2 - \hat{m}_1^2)}$

b)

$\mathcal{L}(\alpha,\beta \mid X_1,\dots,X_n) = \prod_{i=1}^n \left(\frac{1}{\beta-\alpha}\right) = \left(\frac{1}{\beta-\alpha}\right)^n$

$\hat{\alpha}_{MLE} = \underset{\alpha \in [X_1,\dots,X_n]}{\operatorname{argmax}} \left(\frac{1}{\beta-\alpha}\right)^n = \min(X_1,\dots,X_n) = X_{(1)}$

$\hat{\beta}_{MLE} = \underset{\beta \in [X_1,\dots,X_n]}{\operatorname{argmax}} \left(\frac{1}{\beta-\alpha}\right)^n = \max(X_1,\dots,X_n) = X_{(n)}$

3.

$X_1, \dots, X_n \sim U[\alpha, \beta] \quad \mu = \int x\,dF(x)$

a)

$\mu = \frac{\alpha+\beta}{2}$

$\hat{\mu}_{MLE} = \mu(\hat{\alpha}_{MLE}, \hat{\beta}_{MLE}) = \frac{\hat{\alpha}_{MLE} + \hat{\beta}_{MLE}}{2} = \frac{X_{(1)} + X_{(n)}}{2}$

b)

$\hat{\mu}_n = \overline{X}_n \quad n = 10 \quad \alpha = 1 \quad \beta = 3$

$$\begin{aligned}
MSE[\hat{\mu}_n] &= \text{bias}[\hat{\mu}_n]^2 + se[\hat{\mu}_n]^2 \\
&= (\mathbb{E}[\hat{\mu}_n] - \mu)^2 + \mathbb{V}(\hat{\mu}_n)/n \\
&= (\mathbb{E}[\overline{X}_n] - \mu)^2 + \sigma^2/10 \\
&= (\mu-\mu)^2 + \tfrac{1}{10}(\mathbb{E}[x^2] - \mathbb{E}[x]^2) \\
&= 0 + \tfrac{1}{10}\left(\int_\alpha^\beta x^2 \tfrac{1}{\beta-\alpha}\,dx - \left(\int_\alpha^\beta x \tfrac{1}{\beta-\alpha}\,dx\right)^2\right) \\
&= \tfrac{1}{10}\left(\tfrac{x^3}{3}\tfrac{1}{3-1}\Big|_1^3 - \left(\tfrac{x^2}{2}\tfrac{1}{3-1}\Big|_1^3\right)^2\right) \\
&= \tfrac{1}{10}\left(\tfrac{27}{3}\tfrac{1}{2} - \tfrac{1}{3}\tfrac{1}{2} - \left(\tfrac{9}{2}\tfrac{1}{2} - \tfrac{1}{2}\tfrac{1}{2}\right)^2\right)
\end{aligned}$$

$$= \frac{1}{10}\left(\frac{13}{3} - (2)^2\right)$$

$$= \frac{1}{10} \cdot \frac{1}{3} = \frac{1}{30}$$

4. $X_1, \dots, X_n \sim N(\theta, 1)$ $\quad Y_i = \begin{cases} 1, & \text{if } X_i > 0 \\ 0, & \text{if } X_i \leq 0 \end{cases}$ $\quad \psi = \mathbb{E}[Y_1]$

a)

$$\psi = \mathbb{E}[Y_1] = \int_{-\infty}^{\infty} Y\, N(\theta,1)\, dY = \int_{-\infty}^{0} 0 \cdot N(\theta,1)\, dY + \int_0^{\infty} 1 \cdot N(\theta,1)\, dY$$

$$= \int_0^{\infty} N(\theta,1)\, dx = \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\theta)^2}{2}} dx$$

$\hat{\theta}_{MLE} = \overline{X_n}$ from Maximum Likelihood notes eq 12 with $\mu = \theta$

$$\hat{\psi}_{MLE} = \psi(\hat{\theta}_{MLE}) = \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\hat{\theta}_{MLE})^2}{2}} dx = \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\overline{X}_n)^2}{2}} dx$$

b)

$$I_\theta = \hat{\theta}_n \pm z_{1-\alpha/2}\, \hat{se} = \overline{X}_n \pm z_{0.975} \frac{\sigma}{\sqrt{n}} = \overline{X}_n \pm \frac{1.96}{\sqrt{n}}$$

$$I_\psi = \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-I_\theta)^2}{2}} dx = \int_0^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{\left(x-\left(\overline{x}_n \pm \frac{1.96}{\sqrt{n}}\right)\right)^2}{2}} dx$$

5.

$X_1, \dots, X_n \sim U[0,\theta]$ $\quad \hat{\theta}_{MLE} = X_{(n)}$

a)

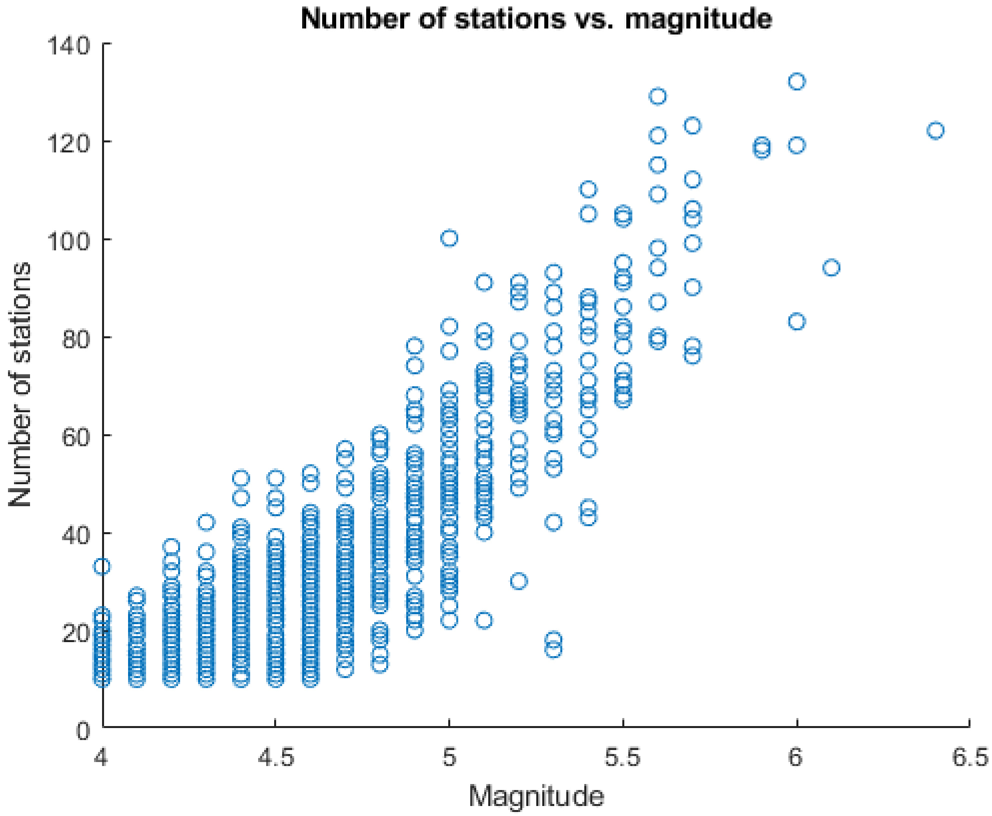
$$CDF = P(X_{(n)} \leq x) = P(X_1 \leq x) \cdot \ldots \cdot P(x_n \leq x)$$

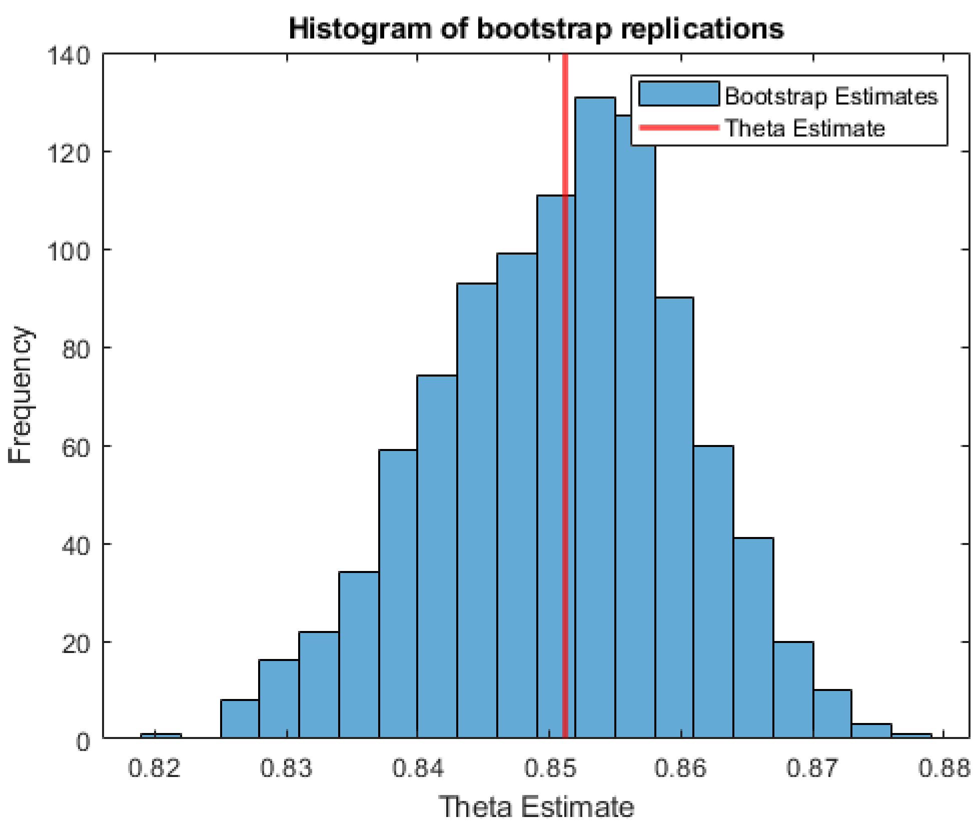
$$= \frac{x}{\theta} \cdot \ldots \cdot \frac{x}{\theta} = \left(\frac{x}{\theta}\right)^n$$

$$PDF = \frac{n x^{n-1}}{\theta^n}$$

```matlab
A = importdata("fiji.txt");
X = A(:,5);
Y = A(:,6);
scatter(X,Y);
title('Number of stations vs. magnitude');
xlabel('Magnitude');
ylabel('Number of stations');

theta = corr(X,Y);
disp(['Theta estimate ', num2str(theta)]);

n = length(X);
B = 10^3;
boot = zeros(B,1);
for i=1:B
    xy = datasample([X Y],n);
    boot(i) = corr(xy(:,1),xy(:,2));
end
bootSE = sqrt(var(boot));
disp(['Bootstrap estimate ', num2str(bootSE)]);

figure;
histogram(boot);
title('Histogram of bootstrap replications');
ylabel('Frequency');
xlabel('Theta Estimate');
hold on;
xline(theta, 'LineWidth', 2, 'Color', 'r');
legend('Bootstrap Estimates', 'Theta Estimate');

a = 0.025;
norm = norminv([a 1-a], theta, bootSE);
pivot = bootci(n, @corr, X, Y);
disp(['Normal interval ', num2str(norm(1)), ' ', num2str(norm(2))]);
disp(['Pivotal interval ', num2str(pivot(1)), ' ',
 num2str(pivot(2))]);
```

```
Theta estimate 0.85118
Bootstrap estimate 0.0094947
Normal interval 0.83257 0.86979
Pivotal interval 0.82966 0.86808
```


Number of stations vs. magnitude
Number of stations
Magnitude
140
120
100
80
60
40
20
0
4
4.5
5
5.5
6
6.5


Histogram of bootstrap replications
Bootstrap Estimates
Theta Estimate
Frequency
Theta Estimate
140
120
100
80
60
40
20
0
0.82
0.83
0.84
0.85
0.86
0.87
0.88

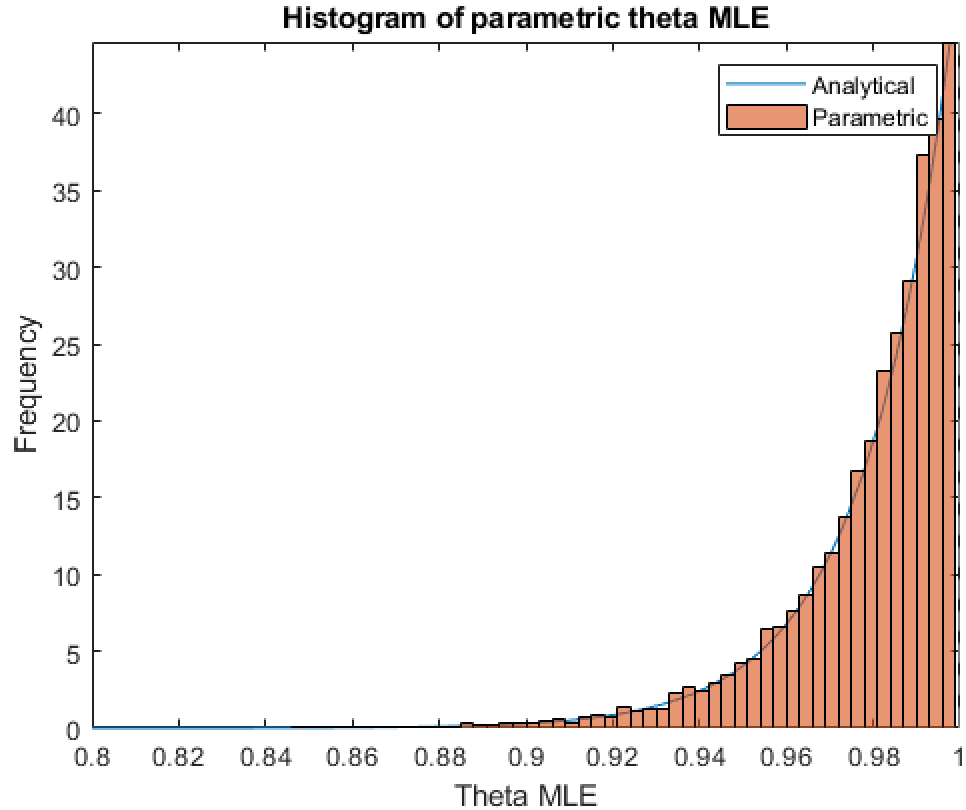
Published with MATLAB® R2021a

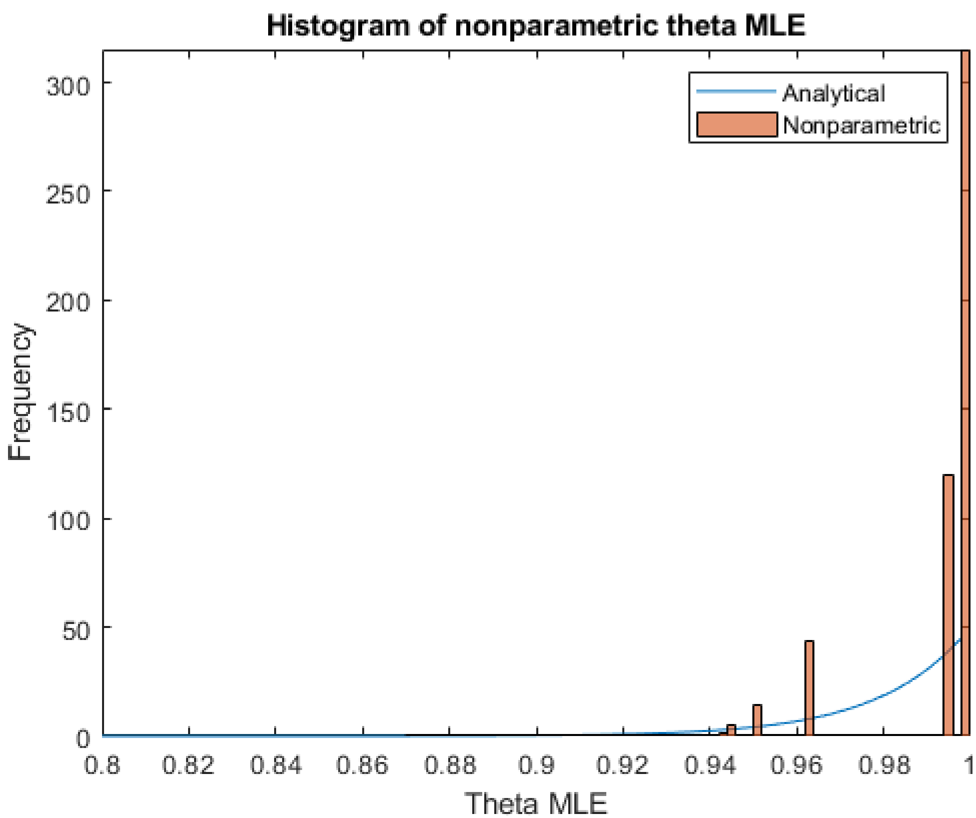
```
n = 10;
a = 1;
b = 3;
mu = (a+b)/2;

sum = 0;
B = 10^4;
for i=1:B
    data = unifrnd(a, b, [n 1]);
    muS = (min(data)+max(data))/2;
    sum = sum+(muS-mu)^2;
end

MSE = sum/B;
disp(['MSE of mu MLE = ', num2str(MSE)]);

% The simulation MSE of mu MLE is 0.015273, less than that of the
% analytical value of 1/30=0.0333
```

```
MSE of mu MLE = 0.015249
```

Published with MATLAB® R2021a

```
n = 50;
theta = 1;

data = unifrnd(0, theta, [n 1]);
thetaMLE = max(data);

B = 10^4;
para = zeros(B,1);
nonpara = zeros(B,1);

for i=1:B
    dataS1 = unifrnd(0, thetaMLE, [n 1]);
    dataS2 = datasample(data, n);
    para(i) = max(dataS1);
    nonpara(i) = max(dataS2);
end

figure
fplot(@(x) (n.*x.^(n-1)),[0.8,1])
hold on
histogram(para, 'Normalization', 'pdf')
title('Histogram of parametric theta MLE')
ylabel('Frequency')
xlabel('Theta MLE')
legend({'Analytical', 'Parametric'})
hold off
figure
fplot(@(x) (n.*x.^(n-1)),[0.8,1])
hold on
histogram(nonpara, 'Normalization', 'pdf')
title('Histogram of nonparametric theta MLE')
ylabel('Frequency')
xlabel('Theta MLE')
legend({'Analytical', 'Nonparametric'})
hold off

% Our analytical theta MLE is fairly close to both distributions with
 the
% parametric looking closer than the nonparametric. Some simulations
 are
% close than others depending on the whether a data point close to 1
 is
% randomly chosen.
```


Histogram of parametric theta MLE
Analytical
Parametric
Frequency
Theta MLE
0
5
10
15
20
25
30
35
40
0.8
0.82
0.84
0.86
0.88
0.9
0.92
0.94
0.96
0.98
1


Histogram of nonparametric theta MLE
Analytical
Nonparametric
Frequency
Theta MLE
0
50
100
150
200
250
300
0.8
0.82
0.84
0.86
0.88
0.9
0.92
0.94
0.96
0.98
1

Published with MATLAB® R2021a