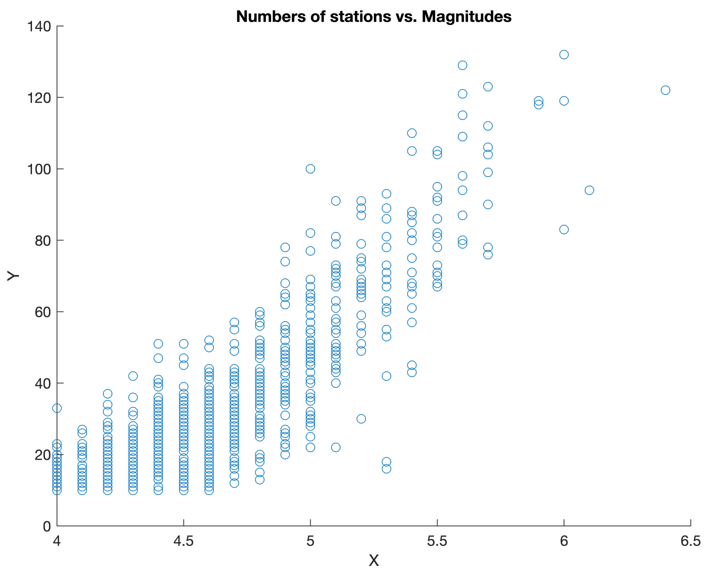
1 Boostrap Method

Part a

```
fiji = importdata("fiji.txt");
X = fiji(:,5); % magnitudes
Y = fiji(:,6); % numbers of stations
scatter(X,Y)
title("Numbers of stations vs. Magnitudes")
xlabel("X")
ylabel("Y")
```



Part b

Given the representation of θ, let us calculate the following plug-in estimate $\hat{\theta}_n$:

$$\hat{\theta}_n = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2}} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

Let us solve with $n1 = \sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y}),\quad d1 = \sum_{i=1}^{n}(X_i - \bar{X})^2,\quad d2 = \sum_{i=1}^{n}(Y_i - \bar{Y})^2$ such that $\hat{\theta}_n = \frac{n1}{\sqrt{d1 \cdot d2}}$

```
n = 10^3;
n1 = 0;
```

```
d1 = 0;
d2 = 0;
for i = 1:n
    n1 = n1 + ((X(i) - mean(X)) * (Y(i) - mean(Y)));
    d1 = d1 + (X(i) - mean(X))^2;
    d2 = d2 + (Y(i) - mean(Y))^2;
end
plugin = n1/sqrt(d1*d2);
disp("Plug-in estimate: "); disp(plugin);
```

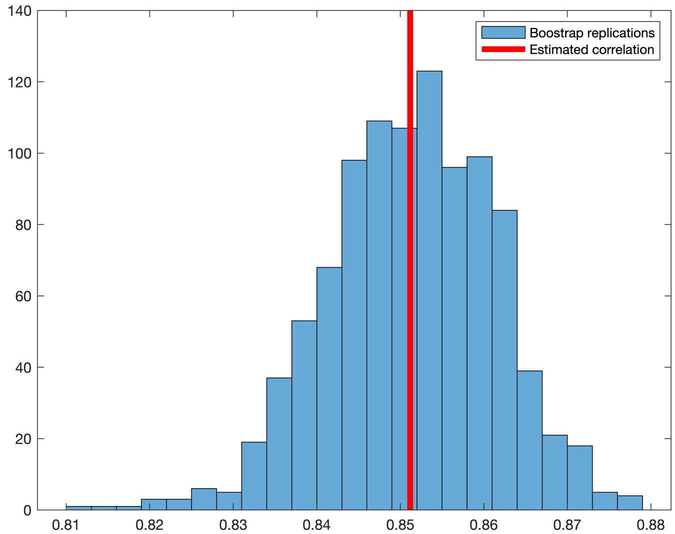
```
Plug-in estimate:
    0.8512
```

Part c

```
B = 10^3;
theta_hat = zeros(B, 1);
for i = 1:B
    samp = datasample([X Y], n);
    Xsamp = samp(:,1);
    Ysamp = samp(:,2);
    theta_hat(i) = corr(Xsamp, Ysamp);
end
bs_se = sqrt(var(theta_hat));
disp("Boostrap estimate of standard error: "); disp(bs_se);
```

```
Boostrap estimate of standard error:
    0.0099
```

```
figure
histogram(theta_hat)
hold on
line([plugin, plugin], ylim, 'LineWidth', 4, 'Color', 'red');
legend('Boostrap replications', 'Estimated correlation');
hold off;
```



We notice a decently strong bell-shaped normal distribution for the boostrap replications of correlation estimates. Furthermore, we clearly see the plug-in estimate lines essentially directly through the center of the distribution of values.

Part d

```
normal_interval = norminv([0.025 0.975], plugin, bs_se);
disp("Normal 95% confidence interval: "); disp(normal_interval);
```

```
Normal 95% confidence interval:
    0.8317    0.8707
```

```
pivotal_interval = bootci(B, @corr, X, Y);
disp("Pivotal 95% confidence interval: "); disp(transpose(pivotal_interval));
```

```
Pivotal 95% confidence interval:
    0.8278    0.8678
```

We notice that the two intervals are very similar.

2 MoM vs. MLE

Let $X_1, ..., X_n \sim \mathcal{U}[\alpha, \beta]$, where α, β are unknown parameters such that $\alpha < \beta$ and $\mathcal{U}[\alpha, \beta]$:

2.1 Part a

Let us find the method of moments estimates of α and β. With $k = 2$, let define the first and second moments:

$$\implies m_1(\alpha, \beta) = \mathbb{E}_f[X] = \frac{\alpha + \beta}{2}$$

$$\implies m_2(\alpha, \beta) = \mathbb{E}_f[X^2] = \int_\alpha^\beta x^2 f(x; \alpha, \beta) dx = \int_\alpha^\beta x^2 \cdot \frac{1}{\beta - \alpha} dx = \frac{1}{3\beta - \alpha} \cdot [x^3]_\alpha^\beta = \frac{\beta^3 - \alpha^3}{3(\beta - \alpha)}$$

$$\implies m_2(\alpha, \beta) = \frac{(\beta - \alpha)(\beta^2 + ab + \alpha^2)}{3(\beta - \alpha)} = \frac{\beta^2 + \alpha\beta + \alpha^2}{3}$$

Now let us define the sample moments as follows:

$$\implies \hat{m}_1 = \frac{1}{n} \sum_{i=1}^n X_i \qquad \implies \hat{m}_2 = \frac{1}{n} \sum_{i=1}^n X_i^2$$

So, we know the MOM estimates $(\hat{\alpha}_{MOM}, \hat{\beta}_{MOM})$ are the solution to the following system of 2 equations:

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

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

$$\implies \beta^2 + 2\beta\hat{m}_1 - \beta^2 + 4\hat{m}_1^2 - 4\beta\hat{m}_1 + \beta^2 = 3\hat{m}_2$$

$$\implies \beta^2 - 2\beta\hat{m}_1 = \frac{3}{n}\hat{m}_2 - 4\hat{m}_1^2 \implies \beta^2 - 2\hat{m}_1\beta + (4\hat{m}_1^2 - 3\hat{m}_2) = 0$$

$$\implies \beta = \frac{2\hat{m}_1 \pm \sqrt{4\hat{m}_1^2 - 4(4\hat{m}_1^2 - 3\hat{m}_2)}}{2} = \frac{2\hat{m}_1 \pm \sqrt{12\hat{m}_2 - 12\hat{m}_1^2}}{2} = \hat{m}_1 \pm \sqrt{3(\hat{m}_2 - \hat{m}_1^2)}$$

$$\implies \alpha = 2\hat{m}_1 - (\hat{m}_1 \pm \sqrt{3(\hat{m}_2 - \hat{m}_1^2)})$$

Since $\beta > \alpha$ we find the solutions as follows:

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

where $\hat{m}_1 = \frac{1}{n}\sum_{i=1}^n X_i$ and $\hat{m}_2 = \frac{1}{n}\sum_{i=1}^n X_i^2$

2.2 Part b

Let us find the MLEs of α and β.

$$\implies \mathcal{L}(\alpha, \beta | X_1, ..., X_n) = \prod_{i=1}^{n} f(X_i; \alpha, \beta) = \prod_{i=1}^{n} \frac{1}{\beta - \alpha} = (\frac{1}{\beta - \alpha})^n$$

We know that $\hat{\alpha}_n$ and $\hat{\beta}_n$ are defined such that $\mathcal{L}$ is maximized. Let us also note that we are assuming that the data falls within $[\alpha, \beta]$ since the likelihood would be 0 otherwise. Therefore, with our definition of $\mathcal{L}$, we can see that we want to minimize the denominator $(\beta - \alpha)$ under the conditions that $X_1, ..., X_n \in [\alpha, \beta]$ with $\beta > \alpha$:

$$\implies \hat{\alpha}_{MLE} = \min(X_1, ..., X_n) \qquad \hat{\beta}_{MLE} = \max(X_1, ..., X_n)$$

3 Plug-In vs. MLE

Let $X_1, ..., X_n \sim \mathcal{U}[\alpha, \beta]$ *and let the parameter of interest be the mean of the distribution* $\mu = \int x dF(x)$*:*

3.1 Part a

Let us find the MLE of μ. From the Equivariance section in Lecture 10, since we are interested in estimating the parameter μ (which is a function of α, β), we simply plug in $\hat{\alpha}_{MLE}$ and $\hat{\beta}_{MLE}$ (from 2b) into the known formula for the mean of uniform distribution:

$$\implies \mu = \frac{\alpha + \beta}{2} \implies \hat{\mu}_{MLE} = \frac{\hat{\alpha}_{MLE} + \hat{\beta}_{MLE}}{2} = \frac{\min(X_1, ..., X_n) + \max(X_1, ..., X_n)}{2}$$

3.2 Part b

Let us first find the MSE of $\widehat{\mu}_{MSE}$ by simulation:

```
n = 10;
a = 1;
B = 3;
m = 10^4;
mu = (a + B)/2;
sum = 0;
for i = 1:m
    X = a + (B-a) * rand(n,1);
    sum = sum + (((min(X) + max(X))/2) - mu)^2;
end
MSE_sim = sum/m;
disp("MSE of MLE of mu"); disp(MSE_sim);
```

```
MSE of MLE of mu
    0.0152
```

Let us now find the MSE of the plug-in estimate $\widehat{\mu}_n = \overline{X}_n$ of μ analytically:

We know $MSE[\widehat{\mu}_n] = bias[\widehat{\mu}_n]^2 + se[\widehat{\mu}_n]^2$:

$$\rightarrow bias[\widehat{\mu}_n] = \mathbb{E}[\widehat{\mu}_n] - \mu = \mu - \mu = 0$$

$$\rightarrow se[\widehat{\mu}_n] = \frac{\sigma}{\sqrt{n}} = \frac{1}{\sqrt{n}} \cdot \sqrt{\frac{(\beta-\alpha)^2}{12}} = \frac{1}{\sqrt{10}} \cdot \sqrt{\frac{(3-1)^2}{12}} = \frac{2}{\sqrt{120}} \qquad \text{(ignore sampling fraction)}$$

$$\rightarrow MSE[\widehat{\mu}_n] = 0^2 + (\frac{2}{\sqrt{120}})^2 = \frac{1}{30} \approx 0.0333$$

Therefore, we see that the MSE of $\widehat{\mu}_n$ is about 2 times greater than the simulated MSE of $\widehat{\mu}_{MSE}$.

4 MLE and CI

Suppose $X_1, ..., X_n \sim \mathcal{N}(\theta, 1)$. Let us define

$$Y_i = \begin{cases} 1 & \text{if } X_i > 0 \\ 0 & \text{if } X_i \leq 0 \end{cases}$$

Also, let $\psi = \mathbb{E}[Y_1]$:

4.1 Part a

Let us express ψ in terms of θ and find the MLE of ψ based on the data $X_1, ..., X_n$.

$$\implies \mathbb{E}[Y_1] = 1 \cdot \mathbb{P}(X_1 > 0) + 0 \cdot \mathbb{P}(X_1 \leq 0) = \mathbb{P}(X_1 > 0)$$

Since we know $X_1 \sim \mathcal{N}(\theta, 1)$, we can simply determine this expression in terms of θ:

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

By Eq. (29) from Lecture 10, we know that $\hat{\theta}_{MLE} = \bar{X}_n$ in this case. So, with the equivariance property (also from Lecture 10) we simply plug $\hat{\theta}_{MLE}$ into ψ in order ot find the MLE of ψ:

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

4.2 Part b

Let us find an approximate 95% confidence interval for ψ from the data $X_1, ..., X_n$:

$$\implies \mathcal{I}_n = \hat{\theta}_n \pm z_{\frac{\alpha}{2}} se[\hat{\theta}_n] = \bar{X}_n \pm (-1.96) \frac{1}{\sqrt{n}}$$

$$\implies \mathcal{I}_\psi = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-(\bar{X}_n \pm (-1.96)\frac{1}{\sqrt{n}}))^2}{2}} dx$$

5 Parametric vs. Non-Parametric Bootstrap

Let $X_1, ..., X_n \sim \mathcal{U}[0, \theta]$. From Lecture 9, we know the MLE of θ is $\hat{\theta}_{MLE} = X_{(n)}$:

5.1 Part a

Let us find the probability density function of $\hat{\theta}_{MLE} = X_{(n)}$:

$$\text{CDF} \implies F(x) = \mathbb{P}(X_{(n)} \leq x) = \mathbb{P}(\max(X_1, ..., X_n) \leq x) = \mathbb{P}(X_1 \leq x, X_2 \leq x, ..., X_n \leq x)$$

$$\implies \mathbb{P}(X_{(n)} \leq x) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq x) = \prod_{i=1}^{n} \frac{x-0}{\theta-0} = (\frac{x}{\theta})^n$$

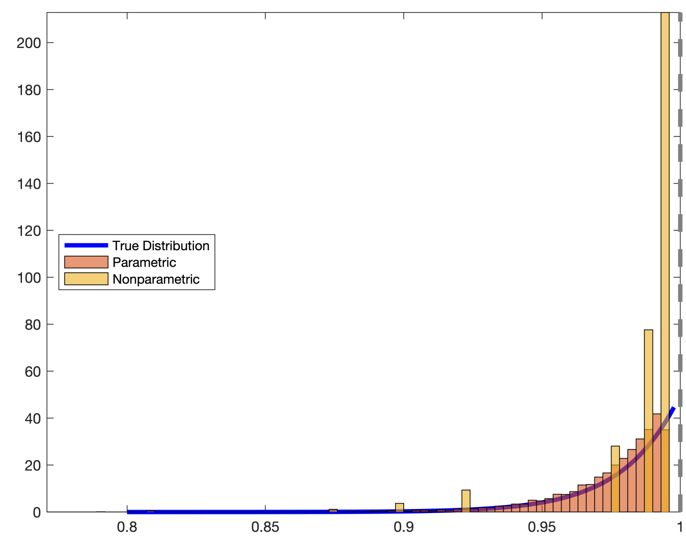
$$\text{PDF} \implies f(x) = \frac{d}{dx} F(x) = \frac{d}{dx} (\frac{x}{\theta})^n = \frac{nx^{n-1}}{\theta^n}$$

5.2 Part b

```
n = 50;
theta = 1;
B = 10^4;
sample = unifrnd(0,1,n,1);
theta_hat_MLE = max(sample);
p = zeros(B,1);
np = zeros(B,1);

for i = 1:B
    p(i) = max(unifrnd(0,theta_hat_MLE,n,1));
    np(i) = max(datasample(sample,n));
end

figure
fplot(@(x) (n.*x.^(n-1)), [0.8,1], 'LineWidth',3,'Color','blue')
hold on
histogram(p, 'Normalization', 'pdf')
histogram(np, 'Normalization', 'pdf')
legend({'True Distribution','Parametric', 'Nonparametric'},'Location', 'west')
hold off
```



We see the bootstrap replications follow a similar exponential trend along the general path of the true distribution (analytical PDF). Clearly, the parametric PDF more closely matches the true distribution of $\widehat{\theta}_{MLE}$ in comparison to the nonparamteric PDF. Comparing the two histograms, the nonparametric PDF is much more sparse which makes sense due to its construction.