Chapter 6.6. p 95-96 ex 3.

③ Let $X_1, \dots X_n \sim \text{Uniform}(0,\theta)$ and let $\hat{\theta} = 2\bar{X}_n$. Find the bias, se and MSE of this estimator

So, $a = 0 \quad b = \theta$

$$E(X_i) = \frac{a+b}{2} = \frac{0+\theta}{2} = \frac{\theta}{2}$$

$$E(\bar{X}_n) = E\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{1}{n} \cdot E(X_1 + \dots X_n) = \frac{1}{\cancel{n}} \cdot \cancel{n} \cdot E(X_i) =$$

$$= \frac{\theta}{2}$$

$$E(\hat{\theta}) = E(2\bar{X}_n) = 2E(\bar{X}_n) = \cancel{2} \cdot \frac{\theta}{\cancel{2}} = \theta$$

$$\text{bias}(\hat{\theta}) = E(\hat{\theta}) - \theta = \theta - \theta = 0$$

$$se = \sqrt{Var(\hat{\theta})}$$

$$Var(X_i) = \frac{(b-a)^2}{12} = \frac{\theta^2}{12}$$

$$Var(\hat{\theta}) = Var(2 \cdot \bar{X}_n) = 4 \cdot Var(\bar{X}_n) = 4 \cdot Var\left(\frac{X_1 + \dots + X_n}{n}\right) =$$

$$= 4 \cdot \frac{1}{n^2} \cdot Var(X_1 + \dots + X_n) = \frac{4}{n^{\cancel{2}}} \cdot \cancel{n} \cdot Var(X_i) = \frac{\cancel{4}}{n} \cdot \frac{\theta^2}{\cancel{12}_{3}} = \frac{\theta^2}{3n}$$

$$se = \sqrt{Var(\hat{\theta})} = \sqrt{\frac{\theta^2}{3n}} = \frac{\theta}{\sqrt{3n}}$$

$$MSE = E(\hat{\theta}_n - \theta)^2 = \text{bias}^2(\hat{\theta}_n) + Var(\hat{\theta}_n) = 0 + \frac{\theta^2}{3n} = \frac{\theta^2}{3n}$$

Answer:

bias = 0 $\qquad MSE = \frac{\theta^3}{3n}$

$se = \frac{\theta}{\sqrt{3n}}$

Chapter 7.4 p 104-105 ex 2, 9

② $X_1, \ldots, X_n \sim Bernoulli(p)$

$Y_1, \ldots, Y_m \sim Bernoulli(q)$

a) Find the plug-in estimator and estimated standard error for p.

$\hat{\theta} = \overline{X}_n$

$E(X_i) = p \qquad E(Y_i) = q$

$Var(X_i) = p(1-p) \qquad Var(Y_i) = q(1-q)$

$E(\hat{\theta}) = E(\overline{X}_n) = E\left(\frac{X_1 + \ldots + X_n}{n}\right) = E(X_i) = p$

$Var(\hat{\theta}) = Var(\overline{X}_n) = Var\left(\frac{X_1 + \ldots + X_n}{n}\right) = \frac{Var(X_i)}{n} = \frac{p(1-p)}{n}$

$se = \sqrt{Var(\hat{\theta})} = \sqrt{\frac{p(1-p)}{n}}$

b) Find an approximate 90 percent confidence interval for p.

$P(X_1 \leq p \leq X_2) = 90\% \qquad \alpha = 0,1$

$Z_{\frac{\alpha}{2}} = Z_{0,05} = \Phi^{-1}(0,05) = 1,65$

$P(-1,65 \leq Z \leq 1,65) = 90\% \qquad Z = \frac{\hat{\theta} - \theta}{\sqrt{Var(\hat{\theta})}}$

$-1,65 \leq \frac{p - \overline{X}_n}{\sqrt{Var(\hat{\theta})}} \leq 1,65$

$-1,65 \cdot \sqrt{\frac{p(1-p)}{n}} \leq p - \overline{X}_n \leq 1,65 \cdot \sqrt{\frac{p(1-p)}{n}}$

$\overline{X}_n - 1,65 \cdot \sqrt{\frac{p(1-p)}{n}} \leq p \leq \overline{X}_n + 1,65 \cdot \sqrt{\frac{p(1-p)}{n}}$

$$\left[\overline{X}_n - 1{,}65\cdot\sqrt{\frac{p(1-p)}{n}}\ ;\ \overline{X}_n + 1{,}65\cdot\sqrt{\frac{p(1-p)}{n}}\right]$$

c) Find the plug-in estimator and estimated standard error for $p-q$.

$$\hat{\theta} = \overline{X}_n - \overline{Y}_m$$

$$E(\hat{\theta}) = p-q$$

$$Var(\hat{\theta}) = \frac{p(1-p)}{n} - \frac{q(1-q)}{m}$$

$$se(\hat{\theta}) = \sqrt{\frac{p(1-p)}{n} - \frac{q(1-q)}{m}}$$

d) Find an approximate 90 percent confidence interval for $p-q$.

$$-1{,}65 \leq \frac{(p-q)-(\overline{X}_n - \overline{Y}_m)}{se} \leq 1{,}65$$

$$\overline{X}_n - \overline{Y}_m - 1{,}65\cdot\sqrt{\frac{p(1-p)}{n} - \frac{q(1-q)}{m}} \leq p-q \leq \overline{X}_n - \overline{Y}_m + 1{,}65\sqrt{\frac{p(1-p)}{n} - \frac{q(1-q)}{m}}$$

$$\left[\overline{X}_n - \overline{Y}_m - 1{,}65\sqrt{\frac{p(1-p)}{n} - \frac{q(1-q)}{m}}\ ;\ \overline{X}_n - \overline{Y}_m + 1{,}65\sqrt{\frac{p(1-p)}{n} - \frac{q(1-q)}{m}}\right]$$

9) n(standard antibiotic treat) $-p_1$
m (new antibiotic treat) $-p_2$

$$p_1 = \frac{90}{100} \qquad p_2 = \frac{85}{100}$$

$$\theta = p_1 - p_2 = 0{,}9 - 0{,}85 = 0{,}05$$

$$\hat{\theta} = \hat{p}_1 - \hat{p}_2 = 0{,}05$$

$$se = \sqrt{Var(\hat{\theta})} = \sqrt{\frac{p_1(1-p_1)}{n} - \frac{q_1(1-q_1)}{m}}$$

$$Se(\theta) = \sqrt{\frac{0,9 \cdot 0,1}{100} + \frac{0,85 \cdot 0,15}{100}} = \sqrt{0,0009 + 0,001275} = 0,047$$

$P(x_1 \leq \hat{\theta} \leq x_2) = 95\%$

$P(-1,96 \leq Z \leq 1,96) = 95\%$ $\quad Z = \frac{\hat{\theta} - \theta}{se}$

$-1,96 \leq \frac{\hat{\theta} - \theta}{Se} \leq 1,96$

$\theta - 1,96 \cdot se \leq \hat{\theta} \leq \theta - 1,96 \cdot se$

$-0,042 \leq \hat{\theta} \leq 0,142$

$[-0,042; 0,142]$

$P(x_1 \leq \theta \leq x_2) = 80\%$ $\quad \alpha = 0,2$

$Z_{0,1} = \cancel{\varphi^{-1}(0,05)} = \varphi^{-1}(0,1) = 1,28$

$[0,05 - 1,28 \cdot 0,047 \leq \hat{\theta} \leq 0,05 + 1,28 \cdot 0,047]$

$[0,010; 0,110]$

```
import math
import numpy as np
from scipy.stats import norm
import matplotlib.pyplot as plt

#  Create a data set (using μ = 5) consisting of n=100 observations.
X = norm.rvs(loc = 5, scale = 1, size = 100)
print(X)
teta = np.exp(5)
teta_hat = np.exp(X.mean())

# Use the bootstrap to get the se and 95 percent confidence interval for θ
B = 10000
bootstr = np.empty(B)
n = len(X)
for i in range(B):
    xx = np.random.choice(X, n, replace=True)
    bootstr[i] = np.exp(xx.mean()) #adding mean of each simulation to our array

# calculating standard error
se = bootstr.std()
# confidence interval
Z = 1.96 # for 95%
confidence_int = [teta_hat - Z * se, teta_hat + Z * se]
print("Theta hat: ", teta_hat)
print("Standard error: ", se)
print("Confidence interval for 95%: ", confidence_int)

#Plot a histogram of the bootstrap replications. This is an estimate of the
distribution of θ
bins = np.linspace(50, 250, 500)
plt.hist(bootstr, bins, label='bootstrap', color='green', histtype='step',
density=True)
plt.axvline(x = teta, color ='black', label = 'θ')
plt.legend(loc='upper left')
plt.show()



Answer:
[5.0325117  5.73385788 5.61654575 4.94202658 4.2343521  3.76700884
 4.49945795 5.80167786 4.56449873 4.41215103 3.67076783 6.17402172
 4.8831115  7.04200272 3.83985942 6.55481927 5.31212254 5.53764894
 5.33106784 5.75830224 5.01888397 6.00223234 5.4657577  4.00295516
 4.0604167  4.71207165 5.44541857 4.04272752 5.33094196 4.37547593
 5.86304092 4.9104957  4.25762128 4.00404148 5.51911244 5.28186802
 5.4422817  5.09790009 3.52142126 5.53136988 4.86780934 6.10664059
 3.89144126 5.79010635 4.83811403 6.2063548  4.09177083 4.30352212
 2.98839113 4.00207589 4.92626594 3.90730675 7.16715754 3.52758752
 4.75968495 5.73064642 5.55425175 4.21140441 4.45774259 5.89803845
 5.43536596 6.62618157 5.95530332 4.50703617 6.30766075 6.10790165
```

```
 5.83187622 6.14156645 5.36147044 6.77432924 4.26646554 3.98427304
 6.24254805 6.36466717 5.18086022 5.38464762 3.83972306 4.62045351
 4.77056058 5.16570386 5.57856985 5.56980316 4.21805758 4.51693184
 4.4850125  4.12915814 3.02602485 5.79109415 4.56423385 3.98005559
 3.51713088 4.07269989 3.7768845  3.46924142 5.51046176 6.23041944
 4.13611082 4.67028487 4.3146555  5.46702469]
Theta hat:  145.05690164266073
Standard error:  13.575125237182917
Confidence interval for 95%:  [118.44965617778222, 171.66414710753924]
```