Homework#6

Name: 范又方 Fan, You-fang

Student number: 106225022

Example 10. 1.21 [P479]

For a sample of size n = 24, we compute the Delta Method variance estimate and the bootstrap variance estimate of $\hat{p}(1-\hat{p})$ using $B = 1000$.

Algorithm 1 Bootstrap algorithm

Step 1. Set B = 1000 times, initial value a = NULL and x1 is the vector of fail times and success times.

Step 2. Repeat Random Resampling x1 with replacement B times and average is b.

$$\mathrm{b} = \widehat{p_i^*} = \frac{x_{1i}^* + \cdots + x_{ni}^*}{n} \quad, x_{1i}^* = 0{,}1\,, i = 1, \ldots, B \quad \text{, be the bootstrap versions of } \hat{p}$$

Step 3. Save b in a with sampling 1000 times, $\mathrm{a} = \mathrm{g}(\widehat{p_b^*}) = \widehat{p_b^*}(1-\widehat{p_b^*}) = b(1-b)$

Step 4. $\mathrm{Var}_B^* = \frac{1}{B-1}\sum_{i=1}^{n}(\widehat{p_b^*}(1-\widehat{p_b^*}) - \overline{\widehat{p_b^*}(1-\widehat{p_b^*})})$

	$\hat{p} = \frac{1}{4}$	$\hat{p} = \frac{1}{2}$	$\hat{p} = \frac{2}{3}$
Bootstrap	0.00189263	0.0002120429	0.001129048

[Delta Method]

Now that we want to estimate the variance of the Bernoulli distribution, $\hat{p}(1-\hat{p})$.

$$\Rightarrow \hat{p} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Let $g(\hat{p}) = \hat{p}(1-\hat{p}),\ g'(\hat{p}) = 1-2\hat{p}$

By the delta method, using the first-order the Delta Method variance,

$$Var(g(\hat{p})) \approx g'(\hat{p})^2 Var(\hat{p})$$

$$\Rightarrow Var(\hat{p}(1-\hat{p})) \approx (1-2\hat{p})^2\left(\frac{\hat{p}(1-\hat{p})}{n}\right) \quad , if\ \hat{p} \neq \frac{1}{2}$$

When $\hat{p} = \frac{1}{4}$,

$$\Rightarrow Var(\hat{p}(1-\hat{p})) \approx \left(1-2\left(\frac{1}{4}\right)\right)^2\left(\frac{\frac{1}{4}\left(1-\frac{1}{4}\right)}{24}\right) = 0.001953125$$

When $\hat{p} = \frac{2}{3}$,

$$\Rightarrow Var(\hat{p}(1-\hat{p})) \approx \left(1-2\left(\frac{2}{3}\right)\right)^2 \left(\frac{\frac{2}{3}\left(1-\frac{2}{3}\right)}{24}\right) = 0.001028807$$

Furthermore, using the second-order Delta Method variance estimate if $\hat{p} = \frac{1}{2}$,

By Theorem 5.5.26, let Y_n be a sequence of random variables that satisfies $\sqrt{n}(Y_n - \theta) \to n(0, \sigma^2)$ in distribution. For a given function g and a specific value of θ, suppose that $g'(\theta) = 0$ and $g''(\theta)$ exists and is not 0.
If $g'(\theta) = 0$, we take one more term in the Taylor expansion to get

$$g(Y_n) = g(\theta) + g'(\theta)(Y_n - \theta) + \frac{g''(\theta)}{2}(Y_n - \theta)^2 + Remainder.$$

$$\Rightarrow g(Y_n) - g(\theta) = \frac{g''(\theta)}{2}(Y_n - \theta)^2 + Remainder$$

Then implies that

$$\frac{n(Y_n - \theta)^2}{\sigma^2} \to \chi_1^2$$

$$\Rightarrow n\big(g(Y_n) - g(\theta)\big) \to N(0, \frac{g''(\theta)}{2}\sigma^2\chi_1^2) \quad in\ distribution$$

When $\hat{p} = \frac{1}{2}$,

$$Var\left(n\big(g(Y_n) - g(\theta)\big)\right) = Var(\frac{g''(\theta)}{2}\sigma^2\chi_1^2)$$

$$= 2Var\left(\frac{-2}{2}\sigma^2\right)$$

$$= 2Var(-\sigma^2)$$

$$= 2\sigma^4$$

$$= 2\big(\hat{p}(1-\hat{p})\big)^2$$

$$= 2\left(\frac{1}{2}\left(1-\frac{1}{2}\right)\right)^2 = \frac{1}{8}$$

$$\Rightarrow Var\left(\big(g(Y_n) - g(\theta)\big)\right) = \frac{\frac{1}{8}}{24^2} = 0.000217013$$

[True variance]

$X_1, \dots, X_{24} \overset{iid}{\sim} Ber(p) \implies \hat{p} = \bar{x}$

$$\implies \mathrm{Y} = \sum_{i=1}^{24} x_i \overset{iid}{\sim} Bin(n\,,p)$$

Then $$Var\big(\hat{p}(1-\hat{p})\big) = Var\left(\frac{\sum_{i=1}^{24} x_i}{n}\left(1 - \frac{\sum_{i=1}^{24} x_i}{n}\right)\right)$$

$$= Var\left(\frac{y}{n}\left(1-\frac{y}{n}\right)\right) = Var\left(\frac{y}{n} - \frac{y^2}{n^2}\right)$$

$$= \frac{1}{n^2}Var(y) + \frac{1}{n^4}Var(y^2) - \frac{2}{n^3}Cov(y, y^2)$$

Now we calculate $Var(y),\ Var(y^2)\ and\ Cov(y, y^2)$.

$$M_Y(t) = E(e^{ty}) = \sum_{y=0}^{n} e^{ty}\binom{n}{y}p^y(1-p)^{n-y} = \sum_{y=0}^{n}\binom{n}{y}(pe^t)^y(1-p)^{n-y}$$

$$= (pe^t + 1 - p)^n$$

Let $\eta_{\mathrm{y}}(t) = \mathrm{M_Y}(lnt) = (pt + 1 - p)^n$

$$f_1 = E(Y) = \frac{d}{dt}\eta_y(t)\Big|_{t=1} = n(pt+1-p)^{n-1}p|_{t=1} = np$$

$$f_2 = E\big(Y(Y-1)\big) = \frac{d^2}{dt^2}\eta_y(t)\Big|_{t=1} = n(n-1)(pt+1-p)^{n-2}p^2|_{t=1}$$

$$= n(n-1)p^2$$

$$f_3 = E\big(Y(Y-1)(Y-2)\big) = \frac{d^3}{dt^3}\eta_y(t)\Big|_{t=1}$$

$$= n(n-1)(n-2)(pt+1-p)^{n-3}p^3|_{t=1} = n(n-1)(n-2)p^3$$

$$f_4 = E\big(Y(Y-1)(Y-2)(Y-3)\big) = \frac{d^4}{dt^4}\eta_y(t)\Big|_{t=1}$$

$$= n(n-1)(n-2)(n-3)(pt+1-p)^{n-4}p^4|_{t=1}$$

$$= n(n-1)(n-2)(n-3)p^4$$

$$\implies E(Y^2) = E(Y(Y-1) + Y) = n(n-1)p^2 + np$$

$$\implies E(Y^3) = E(Y(Y-1)(Y-2) + 3Y^2 - 2Y)$$

$$= n(n-1)(n-2)p^3 + 3(n(n-1)p^2 + np) - 2np$$

$$\Longrightarrow E(Y^4) = E(Y(Y-1)(Y-2)(Y-3) + 6Y^3 - 11Y^2 + 6Y)$$
$$= n(n-1)(n-2)(n-3)p^4$$
$$+6(n(n-1)(n-2)p^3 + 3(n(n-1)p^2) + np)$$
$$-11(n(n-1)p^2 + np) + 6np$$
$$= n(n-1)(n-2)(n-3)p^4 + 6n(n-1)(n-2)p^3 + 7n(n-1)p^2$$
$$+np$$

Therefore, we get

$$Var(Y) = E(Y^2) - \left(E(Y)\right)^2 = np(1-p)$$

$$Var(Y^2) = \mathrm{E}((\mathrm{Y}^2)^2) - \left(\mathrm{E}(\mathrm{Y}^2)\right)^2 = E(Y^4) - \left(\mathrm{E}(\mathrm{Y}^2)\right)^2$$

$$Cov(Y, Y^2) = E(Y^3) - E(Y)E(Y^2) = E(Y^3) - npE(Y^2)$$

$$\Longrightarrow \mathrm{Var}\left(\hat{p}(1-\hat{p})\right) = \frac{1}{\mathrm{n}^2}\mathrm{Var}(\mathrm{y}) + \frac{1}{\mathrm{n}^4}Var(y^2) - \frac{2}{n^3}Cov(y, y^2)$$

When $\hat{p} = \frac{1}{4}$, $a = \frac{np(1-p)}{n^2}$, $b = \frac{1}{n^4}Var(y^2)$, $c = \frac{Cov(y,y^2)}{n^3}$

```
> p=1/4
> n=24
> a=p*(1-p)/n
> b=(n*(n-1)*(n-2)*(n-3)*p^4+6*n*(n-1)*(n-2)*p^3+7*n*(n-1)*p^2+n*p-(n*(n-1)*p^2+n*p)^2)/n^4
> c=(n*(n-1)*(n-2)*p^3+3*n*(n-1)*p^2+n*p-n*p*(n*(n-1)*(p^2)+n*p))/n^3
> v=a+b-2*c
> v
[1] 0.001910739
```

When $\hat{p} = \frac{1}{2}$, $a = \frac{np(1-p)}{n^2}$, $b = \frac{1}{n^4}Var(y^2)$, $c = \frac{Cov(y,y^2)}{n^3}$

```
> p=1/2
> n=24
> a=p*(1-p)/n
> b=(n*(n-1)*(n-2)*(n-3)*p^4+6*n*(n-1)*(n-2)*p^3+7*n*(n-1)*p^2+n*p-(n*(n-1)*p^2+n*p)^2)/n^4
> c=(n*(n-1)*(n-2)*p^3+3*n*(n-1)*p^2+n*p-n*p*(n*(n-1)*(p^2)+n*p))/n^3
> v=a+b-2*c
> v
[1] 0.0002079716
```

When $\hat{p} = \frac{2}{3}$, a $= \frac{np(1-p)}{n^2}$, $b = \frac{1}{n^4} Var(y^2)$, $c = \frac{Cov(y,y^2)}{n^3}$

```
> p=2/3
> n=24
> a=p*(1-p)/n
> b=(n*(n-1)*(n-2)*(n-3)*p^4+6*n*(n-1)*(n-2)*p^3+7*n*(n-1)*p^2+n*p-(n*(n-1)*p^2+n*p)^2)/n^4
> c=(n*(n-1)*(n-2)*p^3+3*n*(n-1)*p^2+n*p-n*p*(n*(n-1)*(p^2)+n*p))/n^3
> v=a+b-2*c
> v
[1] 0.001109182
```

The table : Bootstrap and Delta Method variances of $\hat{p}(1-\hat{p})$. The true is calculated numerically assuming that $\hat{p} = p$.

	$\hat{p} = \frac{1}{4}$	$\hat{p} = \frac{1}{2}$	$\hat{p} = \frac{2}{3}$
Bootstrap	0.00189263	0.0002120429	0.001129048
Delta Method	0.001953125	0.000217013	0.001028807
True	0.001910739	0.0002079716	0.001109182

Example 10. 1.22 (Parametric bootstrap)[P480]

Suppose that we have a sample

$$-1.81\ , 0.63\ , 2.22\ , 2.41\ , 2.95\ , 4.16\ , 4.24\ , 4.53\ , 5.09$$

With $\bar{x} = 2.713333$ and $s^2 = 4.820575$.

```
> data=c(-1.81,0.63,2.22,2.41,2.95,4.16,4.24,4.53,5.09)
> m=mean(data)
> m
[1] 2.713333
> v=var(data)
> v
[1] 4.820575
```

If we assume that the underlying distribution is normal, then a parametric bootstrap would take samples

$$X_1^*, X_2^*, \ldots, X_n^* \sim n(2.71\ , 4.28)$$

Base on $\mathrm{B} = 1000$ samples,

Algorithm 2 Bootstrap algorithm

Step 1. Set $\mathrm{B} = 1000$ times, $\mathrm{n} = 9$

Step 2. Set a $\mathrm{B} \times 9$ matrix of all the elements are 0 and a is a vector that 1000 one inside.

Step 3. Generate the random normal distribution with mean is 2.71 and standard deviation is $4.28^{\frac{1}{2}}$. Repeat B times to calculate sample variance of each column put into b. $\mathrm{S_i^2} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\ , i = 1, \ldots, B$

Step 4. Calculate $\mathrm{Var}_B^*(S^2) = \frac{1}{B-1}\sum_{i=1}^{B}\left(\mathrm{S_i^2} - \overline{\mathrm{S_i^2}}\right)^2$

We calculate $Var_B^*(S^2) = 4.360439$

Base on normal theory,

The likelihood function is

$$L(\mu, \sigma^2|\boldsymbol{x}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{(x_i - \mu)^2}{2\sigma^2})$$

$$= (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{n}{2}} exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \bar{x} + \bar{x} - \mu)^2\right)$$

$$= (2\pi)^{-\frac{n}{2}}(\sigma^2)^{-\frac{n}{2}} exp(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2) exp(-\frac{n}{2\sigma^2}(\bar{x}-\mu)^2)$$

$$logL(\mu,\sigma^2|\boldsymbol{x}) = \left(-\frac{n}{2}\right)log(2\pi) - \frac{n}{2}log\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\bar{x})^2 - \frac{n}{2\sigma^2}(\bar{x}-\mu)^2$$

$$\frac{\partial}{\partial\mu}logL(\mu,\sigma^2|\boldsymbol{x}) = -\frac{n}{2\sigma^2}2(\bar{x}-\mu)(-1) \overset{set}{=} 0 \Longrightarrow \hat{\mu} = \bar{x}$$

$$\frac{\partial}{\partial\sigma^2}logL(\mu,\sigma^2|\boldsymbol{x})\bigg|_{\mu=\hat{\mu}} = \left(-\frac{n}{2}\right)\frac{1}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\bar{x})^2 + \frac{n}{2\sigma^4}(\bar{x}-\hat{\mu})^2 \overset{set}{=} 0$$

$$\Longrightarrow -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i-\bar{x})^2 = 0$$

$$\Longrightarrow \widehat{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$$

```
> data=c(-1.81,0.63,2.22,2.41,2.95,4.16,4.24,4.53,5.09)
> m=mean(data)
> s=(sum((data-mean(data))^2))/9
> s
[1] 4.284956
```

Therefore, $\widehat{\sigma^2} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n} = 4.28$

Then the variance of S^2,

Because $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

$$\Longrightarrow Var\left(\frac{(n-1)S^2}{\sigma^2}\right) = 2(n-1)$$

$$\Longrightarrow Var\left((n-1)S^2\right) = 2(n-1)\sigma^4$$

$$\Longrightarrow Var(S^2) = \frac{2\sigma^4}{n-1}$$

$$\Longrightarrow Var(S^2)|_{\sigma^2=\widehat{\sigma^2}} = \frac{2(\widehat{\sigma^2})^2}{n-1} = \frac{2(4.28)^2}{8} = 4.5796$$

The data values were actually generated from a normal distribution with variance 4,

```
> B=1000
> n=9
> b=matrix(0,B,9)
> a=rep(1,B)
> for(i in 1:1000){
+   b[i,]=rnorm(9,m,4^0.5)
+   a[i]=var(b[i,])
+  }
> var(a)
[1] 4.056825
```

So $Var(S^2) \approx 4$.

The parametric bootstrap is a better estimate here.

Appendix 1. R codes for Example 10. 1.21

R codes

############### Example 10. 1.21###############

$\hat{p} = \frac{1}{4}$

```
> set.seed(500)
> B=1000
> a=NULL
> x1=c(rep(0,18),rep(1,6))
> x1
 [1] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1
> for(i in 1:B){
+   b=mean(sample(x1,24,TRUE))
+   a[i]=b*(1-b)
+ }
> var(a)
[1] 0.00189263
```

$\hat{p} = \frac{1}{2}$

```
> set.seed(500)
> B=1000
> a=NULL
> x2=c(rep(0,12),rep(1,12))
> x2
 [1] 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1
> for(i in 1:1000){
+   b=mean(sample(x2,24,TRUE))
+   a[i]=b*(1-b)
+ }
> var(a)
[1] 0.0002120429
```

$\hat{p} = \frac{2}{3}$

```
> set.seed(500)
> B=1000
> a=NULL
> x3=c(rep(0,8),rep(1,16))
> x3
 [1] 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
> for(i in 1:1000){
+   b=mean(sample(x3,24,TRUE))
+   a[i]=b*(1-b)
+ }
> var(a)
[1] 0.001129048
```

Appendix 2. R codes for Example 10. 1.22

R codes

############## Example 10. 1.22##############

```
> data=c(-1.81,0.63,2.22,2.41,2.95,4.16,4.24,4.53,5.09)
> m=mean(data)
> s=(sum((data-mean(data))^2))/9
> s
[1] 4.284956

> B=1000
> n=9
> b=matrix(0,B,9)
> a=rep(1,B)
> for(i in 1:1000){
+   b[i,]=rnorm(9,m,s^0.5)
+   a[i]=var(b[i,])
+  }
> var(a)
[1] 4.32529
```