Homework #9

MSSC 6010- Computational Probability

Henri Medeiros Dos Reis

November 28, 2023

‘

Question 1. (6.7.18) From book. The plastic tubes produced by company X for the irrigation system used in golf courses have a mean length of 1.5 meters and a standard deviation of 0.1 meter. The plastic tubes produced by company Y have a mean length of 1 meter and a standard deviation of 0.09 meter. Suppose that both tube lengths follow a normal distributions.

- *(a)* Calculate the probability that a random sample of 15 tubes from company X has a mean length at least 0.45 meter greater than the mean length of a random sample of size 20 from company Y.

$$Y \sim N(1, 0.09), X \sim N(1.5, 0.1)$$
$$\bar{Y} \sim N(1, \frac{0.09}{\sqrt{20}}), \bar{X} \sim N(1.5, \frac{0.1}{\sqrt{15}})$$

$$\begin{aligned} \bar{X} - \bar{Y} &\sim N(\mu_x - \mu_y, \sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}) \\ &\sim N(1.5 - 1, \sqrt{\frac{0.1^2}{15} + \frac{0.09^2}{20}}) \\ \mathbb{P}(\bar{X} - \bar{Y} > 0.45) &= \end{aligned}$$

```
1-pnorm(0.45,0.5,sqrt(0.1**2/15+0.09**2/20))
```

```
[1] 0.9366637
```

- *(b)* Suppose that the population variances are unknown but equal, $S_x = 0.1$ and $S_y = 0.09$. Calculate the probability that a random sample of 15 plastic tubes from company X has a mean length at least 0.45 meter greater than the mean length of a random sample of 20 plastic tubes from company Y

We use the same approach, but with samples standard deviations and t distribution

$$\frac{[(\bar{X}-\bar{Y})-(\mu_x-\mu_y]}{\sqrt{\frac{(n_x-1)S_x^2+(n_y-1)S_y^2}{n_x+n_y-2}(\frac{1}{n_x}+\frac{1}{n_y})}} \sim t_{n_x+n_y-2}$$

then

$$\mathbb{P}\left(\frac{[(\bar{X}-\bar{Y})-(1.5-1]}{\sqrt{\frac{(15-1)0.1^2+(20-1)0.09^2}{15+20-2}(\frac{1}{15}+\frac{1}{20})}} > 0.45\right) =$$
$$\mathbb{P}\left((\bar{X}-\bar{Y} > 0.45\sqrt{\frac{(15-1)0.1^2+(20-1)0.09^2}{15+20-2}(\frac{1}{15}+\frac{1}{20})}+0.5\right) \sim t_{15+20-2} =$$

```
1-pt(0.45*sqrt(((15-1)*0.1**2+(20-1)*0.09**2)
               /(15+20-2)*(1/15+1/20))+0.5,15+20-2)
```

```
[1] 0.3051638
```

Question 2. (7.4.4) Let X be a $Bin(n,\pi)$ random variable.

- *(a)* Find the mean squared error of the π parameter estimators $T_1 = X/n$ and $T_2 = (X+1)/(n+2)$

$$\begin{aligned} MSE(T_1) &= var(T_1) + (E[T_1]-\pi)^2 \\ &= Var(X/n) + (E[X/n]-\pi)^2 \\ &= \frac{1}{n^2}(n\pi(1-\pi)) + (\frac{1}{n}n\pi-\pi)^2 \\ &= \frac{1}{n^2}(n\pi-\pi)^2 \\ &= \frac{n-\pi^2}{n} \end{aligned}$$

$$\begin{aligned} MSE(T_2) &= var(\frac{X+1}{n+2}) + (E[\frac{X+1}{n+2}]-\pi)^2 \\ &= \frac{1}{(n+2)^2}var(X) + (\frac{1}{n+2}E[X]+\frac{1}{n+2}-\pi)^2 \\ &= \frac{1}{(n+2)^2}(n\pi(1-\pi)) + (\frac{n\pi}{n+2}+\frac{1}{n+2}-\pi)^2 \end{aligned}$$

- *(b)* When n=100 and $\pi = 0.4$, which estimator, T_1 or T_2, has the smaller MSE?

$$MSE(T_1) = \frac{0.4-0.4^2}{100}$$

```
(0.4-0.4**2)/100
```

```
[1] 0.0024
```

$$MSE(T_2) = \frac{1}{102^2}(100(0.4) - 100(0.4)^2) + (\frac{100(0.4)}{102} + \frac{1}{102} - 0.4)^2$$

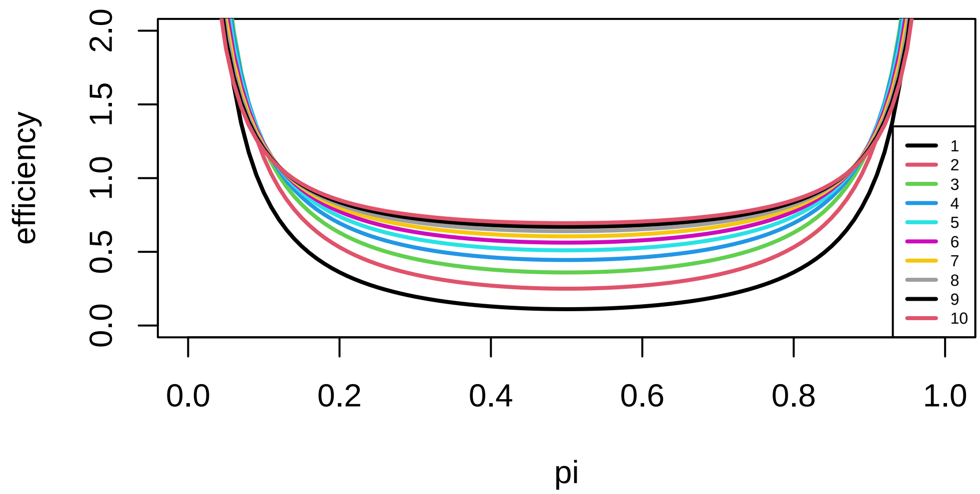
```
(1/102**2)*(100*0.4-100*(0.4**2))+(100*0.4/102+1/102-0.4)**2
```

```
[1] 0.00231065
```

- *(c)* Plot the efficiency of T_2 relative to T_1 versus π values in (0,1) for n values from 1 to 10.

```
T_1 <- function(n,p){
  T_1 <- (p-p**2)/n
}
T_2 <- function(n,p){
  T_2 <- (1/(n+2)**2)*(n*p-n*(p**2))+(n*p/(n+2)+1/(n+2)-p)**2
}
eff <- function(T_2,T_1){
  eff <- T_2/T_1
}
n <- seq(1,10)
p <- seq(0,1,0.01)
my_col<- c()
plot(0,type='n', xlim=c(0,1), ylim=c(0,2), xlab="pi",ylab='efficiency')
for( i in n)
{
  my_T_1 <- sapply(p,T_1, n=i)
  my_T_2 <- sapply(p,T_2, n=i)
  lines(p,eff(my_T_2,my_T_1), col = i, lwd=2)

  my_col <- c(my_col,i)
}
legend('bottomright', legend=as.character(1:10),
       col=my_col, lwd=2,cex=0.5)
```



Question 3. (7.4.9) from book. Verify that $Var[\frac{\partial f(X|\theta)}{\partial\theta}] = E[(\frac{\partial lnf(X|\theta)}{\partial})^2]$. (Hint: show that $E[(\frac{\partial lnf(X|\theta)}{\partial})] = 0$)

Since $Var[\frac{\partial f(X|\theta)}{\partial\theta}] = E[(\frac{\partial lnf(X|\theta)}{\partial})^2] - E[(\frac{\partial lnf(X|\theta)}{\partial})]^2$, we just need to show the hint.

$$E[(\frac{\partial lnf(X|\theta)}{\partial})] = \int_{-\infty}^{\infty}[(\frac{\partial lnf(X|\theta)}{\partial})]f(X|\theta)d\theta$$

and $\frac{d}{dx}ln(f(x)) = \frac{1}{f(x)}f'(x)$, then

$$E[(\frac{\partial lnf(X|\theta)}{\partial})] = \int_{-\infty}^{\infty}\frac{1}{f(X|\theta)}\frac{\partial}{\partial\theta}(X)f(X|\theta)d\theta = \frac{\partial}{\partial\theta}\int_{-\infty}^{\infty}f(X|\theta)d\theta$$

And the integral of any valid cdf from negative infinity to infinity is one, then

$$E[(\frac{\partial lnf(X|\theta)}{\partial})] = \frac{\partial}{\partial\theta} = 0$$

Therefore $Var[\frac{\partial f(X|\theta)}{\partial\theta}] = E[(\frac{\partial lnf(X|\theta)}{\partial})^2] + 0^2$

Extra 1. If X has a binomial distribution with $n = 3$ and $\pi = 1/3$, find the probability distributions of

- *(a)* $Y = \frac{X}{1+X}$

$$(1+X)Y = X \Rightarrow Y = X - XY \Rightarrow X = \frac{Y}{1-Y}$$

$$f_x(x) = \binom{3}{x}(1/3)^x(2/3)^{3-x}$$

Then

$$f_y(y) = f_x(x)|J(x \to y)|, \text{ where } |J(x \to y)| = \frac{dx(y)}{dy}$$
$$= \binom{3}{\frac{Y}{1-Y}}(1/3)^{\frac{Y}{1-Y}}(2/3)^{3-\frac{Y}{1-Y}}|(1-y)^{-1} + y(1-y)^{-2}|$$

- *(b)* $Y = (X-1)^4$

$$X = Y^{1/4} + 1$$

  then

$$f_y(y) = \binom{3}{Y^{1/4}+1}(1/3)^{Y^{1/4}+1}(2/3)^{3-Y^{1/4}+1}|\frac{1}{4}y^{-3/4}|$$

Extra 2. If $X = lnY$ has a normal distribution with the mean μ and standard deviation σ, find the probability density of Y which is said to have the log-normal distribution

$$e^X = Y \qquad f_x(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\{\frac{(x-\mu)^2}{2\sigma^2}\}$$

then

$$f_y(y) = f_x(lny)|\frac{1}{y}| = \frac{1}{\sigma\sqrt{2\pi}}\exp\{\frac{(lny-\mu)^2}{2\sigma^2}\}|\frac{1}{y}|$$

Extra 3. Let X_1 and X_2 be two continuous random variables having the joint probability density

$$f(x_1, x_2) = \begin{cases} 4x_1x_2 & 0 < x_1 < 1 \quad 0 < x_2 < 1 \\ 0 & \text{elsewhere} \end{cases}$$

Find the joint probability density of $Y_1 = X_1^2$ and $Y_2 = X_1X_2$

$$X_1 = Y_1^{1/2} \qquad X_2 = \frac{Y_2}{Y_1^{1/2}}$$

and

$$|J(X \to Y)| = \begin{vmatrix} \frac{1}{2}y_1^{1/2} & 0 \\ y_2(\frac{-1}{2}y^{-3/2}) & y_1^{-1/2} \end{vmatrix} = \frac{1}{2y_1^{1/2}}\frac{1}{y_1^{1/2}} - y_2(\frac{-1}{2}y^{-3/2})0$$

Then

$$f_y(y_1, y_2) = 4y_1^{1/2}\frac{y_2}{y_1^{1/2}}|J(x \to y)| = 4y_2(\frac{1}{2y_1^{1/2}}\frac{1}{y_1^{1/2}}) = 2\frac{y_2}{y_1}$$