Homework#5 Statistical Inference II

Name: Jia-Han Shih

Problem 1.9 [p.391]

In Example 1.7, let $\delta^*(X) = X/n$ with probability $1-\varepsilon$ and $\delta^*(X) = 1/2$ with probability ε. Determine the risk function of δ^* and show that for $\varepsilon = 1/(n+1)$, its risk is constant and less than $\sup R(p, X/n)$.

Solution:

The risk function of δ^* with $\varepsilon = 1/(n+1)$ is

$$R(p, \delta^*) = E(\delta^* - p)^2 = E\left(\frac{X}{n} - p\right)^2 \frac{n}{n+1} + E\left(\frac{1}{2} - p\right)^2 \frac{1}{n+1}$$

$$= \frac{E(X - np)^2}{n^2} \cdot \frac{n}{n+1} + \frac{p^2 - p + \frac{1}{4}}{n+1} = \frac{p - p^2 + p^2 - p + \frac{1}{4}}{n+1} = \frac{1}{4(n+1)}.$$

The risk function of X/n is

$$R\left(p, \frac{X}{n}\right) = E\left(\frac{X}{n} - p\right)^2 = \frac{E(X - np)^2}{n^2} = \frac{p(1-p)}{n}.$$

Since

$$\frac{d}{dp} R\left(p, \frac{X}{n}\right) = 0 \Rightarrow 1 - 2p = 0 \Rightarrow p = \frac{1}{2}, \text{ and } \frac{d^2}{dp^2} R\left(p, \frac{X}{n}\right) = -\frac{2}{n} < 0.$$

Therefore, we have

$$\sup R\left(p, \frac{X}{n}\right) = \left.\frac{p(1-p)}{n}\right|_{p=\frac{1}{2}} = \frac{1}{4n}.$$

Thus, we have shown the desired result

$$R(p, \delta^*) = \frac{1}{4(n+1)} < \frac{1}{4n} = \sup R\left(p, \frac{X}{n}\right).$$

Problem 1.10 [p.391]

Find the bias of minimax estimator (1.11) and discuss its direction.

Solution:

The minimax estimator (1.11) is

$$\delta = \frac{X + \sqrt{n}/2}{n + \sqrt{n}}.$$

The bias of δ is

$$E\left(\frac{X + \sqrt{n}/2}{n+\sqrt{n}} - p\right) = \frac{1}{n+\sqrt{n}} E\left(X + \frac{\sqrt{n}}{2} - np - \sqrt{n}p\right) = \frac{\sqrt{n}/2 - \sqrt{n}p}{n+\sqrt{n}} = \frac{1/2 - p}{1+\sqrt{n}}.$$

Then we have the following results,

$$\frac{d}{dp} Bias(\delta) = \frac{-1}{1+\sqrt{n}} < 0, \text{ and } Bias(\delta) = \frac{1/2 - p}{1+\sqrt{n}} \to 0, \text{ as } n \to \infty.$$

Furthermore, we have the following conclusions:

1. If $p > 1/2$, then $Bias(\delta) < 0$.

2. If $p < 1/2$, then $Bias(\delta) > 0$.

Problem 1.11 [p.391]

In Example 1.7,

(a) determine c_n and show that $c_n \to 0$, as $n \to \infty$,

Solution:

To solve c_n, we have

$$R\left(p, \frac{X}{n}\right) = R(p, \delta) \Rightarrow \frac{p(1-p)}{n} = \frac{1}{4(1+\sqrt{n})^2} \Rightarrow n = (p - p^2)(2+2\sqrt{n})^2$$

$$\Rightarrow (2+2\sqrt{n})^2 p^2 - (2+2\sqrt{n})^2 p + n = 0$$

$$\Rightarrow p^2 - p + \frac{n}{(2+2\sqrt{n})^2} = 0$$

$$\Rightarrow p = \frac{1 \pm \sqrt{1 - \frac{4n}{(2+2\sqrt{n})^2}}}{2} \Rightarrow p = \frac{1 \pm \sqrt{1 - \frac{n}{n+2\sqrt{n}+1}}}{2}$$

$$\Rightarrow p = \frac{1 \pm \frac{\sqrt{1+2\sqrt{n}}}{1+\sqrt{n}}}{2}.$$

Therefore, we obtain

$$c_n = \frac{\sqrt{1+2\sqrt{n}}}{1+\sqrt{n}}.$$

Taking limit for n goes to infinite,

$$\lim_{n\to\infty} c_n = \lim_{n\to\infty} \sqrt{1 - \frac{n}{n+2\sqrt{n}+1}} = \sqrt{1-1} = 0.$$

Then we have shown that

$$c_n \to 0, \text{ as } n \to \infty.$$

(b) show that $R_n(1/2)/r_n \to 1$, as $n \to \infty$.

Solution:

In Example 1.7, we have

$$R_n\left(\frac{1}{2}\right) = \frac{1}{4n}, \text{ and } r_n = \frac{1}{4(1+\sqrt{n})^2}.$$

Therefore, we have

$$\frac{R_n\left(\frac{1}{2}\right)}{r_n} = \frac{4(1+\sqrt{n})^2}{4n} = \frac{n+2\sqrt{n}+1}{n}.$$

Taking limit for n goes to infinite,

$$\lim_{n\to\infty} \frac{R_n\left(\frac{1}{2}\right)}{r_n} = \lim_{n\to\infty} \frac{n+2\sqrt{n}+1}{n} = 1.$$

Then we have proven that

$$R_n(1/2)/r_n \to 1, \text{ as } n \to \infty.$$

Problem 1.12 [p.391]

In Example 1.7, graph the risk function of X/n and the minimax estimator (1.11) for $n = 1, 4, 9, 16$, and indicate the relative position of the two graphs for large values of n.

Solution:

As in problem 1.11, the risk function of X/n and the minimax estimator δ are

$$\frac{p(1-p)}{n}, \text{ and } \frac{1}{4(1+\sqrt{n})^2},$$

respectively. Then we can plot the risk function for $n = 1, 4, 9, 16$ by R (Figure 1).

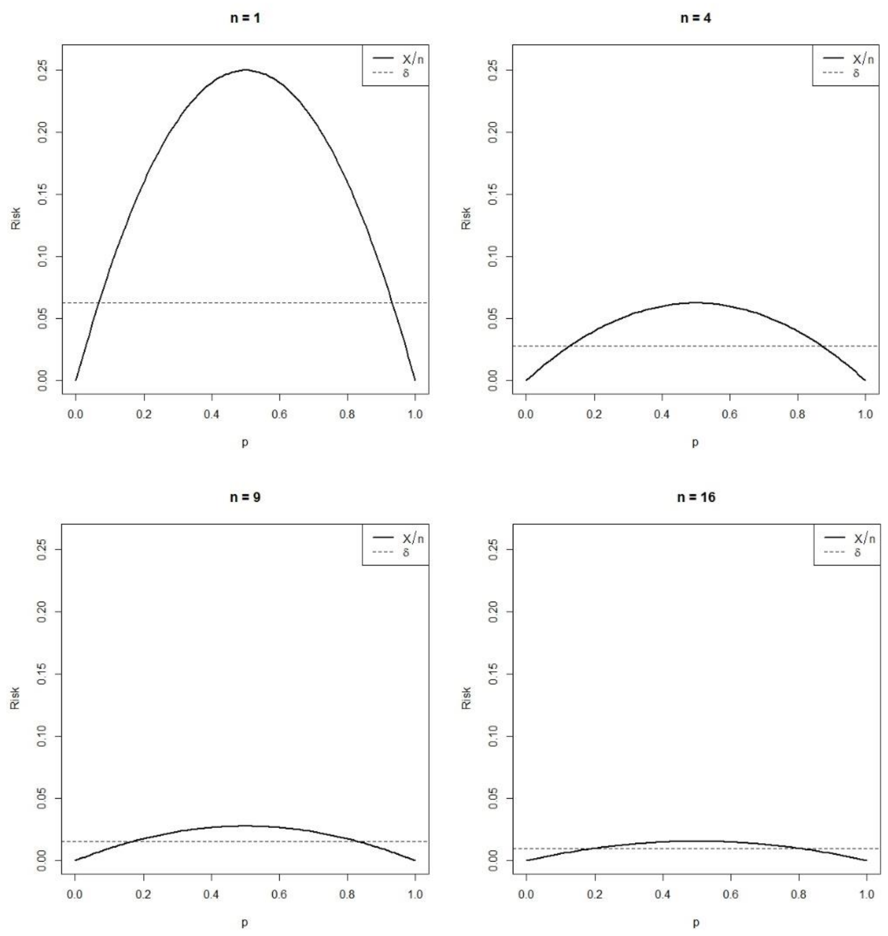

Fig. 1 The risk functions of X/n and δ for $n = 1, 4, 9, 16$.

For the case of large value of n, I set $n=10000$ (Figure 2). Figure 1 and 2 reveal that the interval of the risk of X/n greater than the risk of δ shrink as n goes larger. When n goes to infinite, the interval will shrink to the point $p=1/2$.

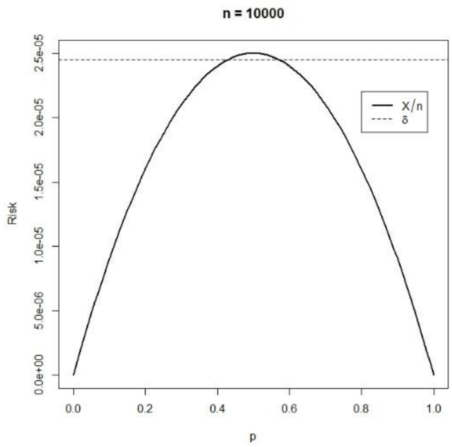

Fig. 2 The risk function of X/n and δ for $n=10000$.

Problem 1.13 [p.391]

(a) Find two points $0 < p_0 < p_1 < 1$ such that the estimator (1.11) for $n = 1$ is Bayes with respect to a distribution Λ for which $\Pr(p = p_0) + \Pr(p = p_1) = 1$.

Solution:

With $n = 1$, the estimator in (1.11) is

$$\delta = \begin{cases} 3/4 & \text{if } X = 1 \\ 1/4 & \text{if } X = 0 \end{cases}.$$

Also, we have $X \mid p \sim \text{Ber}(p)$, that is

$$\Pr(X = 1 \mid p) = p \quad \text{and} \quad \Pr(X = 0 \mid p) = 1 - p.$$

Now, consider a discrete prior

$$\Pr(p = p_0) = a \quad \text{and} \quad \Pr(p = p_1) = 1 - a \quad \text{with} \quad 0 < p_0 < p_1 < 1, \quad 0 < a < 1.$$

To obtain the posterior distribution, we need calculate the following conditional probabilities

$$\begin{aligned} \Pr(p = p_0 \mid X = 1) &= \frac{\Pr(p = p_0, X = 1)}{\Pr(X = 1)} \\ &= \frac{\Pr(X = 1 \mid p = p_0)\Pr(p = p_0)}{\Pr(X = 1 \mid p = p_0)\Pr(p = p_0) + \Pr(X = 1 \mid p = p_1)\Pr(p = p_1)} \\ &= \frac{ap_0}{ap_0 + (1-a)p_1}. \end{aligned}$$

$$\begin{aligned} \Pr(p = p_1 \mid X = 1) &= \frac{\Pr(p = p_1, X = 1)}{\Pr(X = 1)} \\ &= \frac{\Pr(X = 1 \mid p = p_1)\Pr(p = p_1)}{\Pr(X = 1 \mid p = p_0)\Pr(p = p_0) + \Pr(X = 1 \mid p = p_1)\Pr(p = p_1)} \\ &= \frac{(1-a)p_1}{ap_0 + (1-a)p_1}. \end{aligned}$$

$$\Pr(p=p_0 \mid X=0)=\frac{\Pr(p=p_0, X=0)}{\Pr(X=0)}$$
$$=\frac{\Pr(X=0\mid p=p_0)\Pr(p=p_0)}{\Pr(X=0\mid p=p_0)\Pr(p=p_0)+\Pr(X=0\mid p=p_1)\Pr(p=p_1)}$$
$$=\frac{a(1-p_0)}{a(1-p_0)+(1-a)(1-p_1)}.$$
$$\Pr(p=p_1 \mid X=0)=\frac{\Pr(p=p_1, X=0)}{\Pr(X=0)}$$
$$=\frac{\Pr(X=0\mid p=p_1)\Pr(p=p_1)}{\Pr(X=0\mid p=p_0)\Pr(p=p_0)+\Pr(X=0\mid p=p_1)\Pr(p=p_1)}$$
$$=\frac{(1-a)(1-p_1)}{a(1-p_0)+(1-a)(1-p_1)}.$$

Then let the posterior mean equal to δ, we have

$$E(p\mid X=1)=\frac{3}{4}\Rightarrow\frac{ap_0^2+(1-a)p_1^2}{ap_0+(1-a)p_1}=\frac{3}{4},$$
$$E(p\mid X=1)=\frac{3}{4}\Rightarrow\frac{a(1-p_0)p_0+(1-a)(1-p_1)p_1}{a(1-p_0)+(1-a)(1-p_1)}=\frac{1}{4}.$$

Solve these two equation, we have

$$\begin{cases} ap_0^2+(1-a)p_1^2=3/4 \\ ap_0+(1-a)p_1=1/2 \end{cases}.$$

For all discrete prior that I defined previously satisfy the above two equation then its Bayes estimator is equal to δ. For example, let $a=1/2$, we have

$$\begin{cases} p_0^2+p_1^2=3/2 \\ p_0+p_1=1 \end{cases}.$$

The solution is

$$p_0=\frac{2-\sqrt{2}}{4} \quad \text{and} \quad p_1=\frac{2+\sqrt{2}}{4}$$

Hence we have found two points $0<p_0<p_1<1$ such that the estimator (1.11) for $n=1$ is Bayes with respect to a distribution Λ for which

$\Pr(p = p_0) + \Pr(p = p_1) = 1$.

(b) For $n = 1$, show that (1.11) is a minimax estimator of p even if it is known that $p_0 \le p \le p_1$.

Solution:

The risk function of δ is

$$\begin{aligned} E\left(\frac{2X+1}{4} - p\right)^2 &= E\left(\frac{4X^2+4X+1}{16} + p^2 - \left(X + \frac{1}{2}\right)p\right)^2 \\ &= \frac{E(X^2)}{4} + \frac{E(X)}{4} + \frac{1}{4} + p^2 - E(X)p - \frac{p}{2} \\ &= \frac{p}{4} + \frac{p}{4} + \frac{1}{4} + p^2 - p^2 - \frac{p}{2} \\ &= \frac{1}{4}. \end{aligned}$$

The estimator δ has constant risk. By (a), we have already shown that it is a Bayes estimator. Therefore, δ is a minimax estimator.

(c) In (b), find the values p_0 and p_1 for which $p_1 - p_0$ is as small as possible.

Solution:

We need to find p_0 and p_1 minimize $p_1 - p_0$ and satisfy equations in (a), that is

$$\begin{cases} ap_0^2 + (1-a)p_1^2 = 3/4 \\ ap_0 + (1-a)p_1 = 1/2 \end{cases}.$$

By the second equation, we have

$$p_0 = \frac{1-2(1-a)p_1}{2a}.$$

Taking it into the first equation, we have

$$\begin{aligned} &\frac{1+4(1-a)^2 p_1^2 - 4(1-a)p_1}{4a} + (1-a)p_1^2 = \frac{3}{4} \\ &\Rightarrow 1+4(1-a)^2 p_1^2 - 4(1-a)p_1 + 4a(1-a)p_1^2 = 3a \\ &\Rightarrow 4(1-a)p_1^2 - 4(1-a)p_1 + 1 - 3a = 0 \\ &\Rightarrow p_1^2 - p_1 + \frac{1-3a}{4-4a} = 0. \end{aligned}$$

Therefore, the solution of p_1 is

$$p_1 = \frac{1 \pm \sqrt{1 - 4 \cdot 1 \cdot \frac{1-3a}{4-4a}}}{2} = \frac{1 \pm \sqrt{\frac{2a}{1-a}}}{2}.$$

Since $p_0 \le p_1$, we obtain

$$p_1 = \frac{1+\sqrt{\frac{2a}{1-a}}}{2} \quad \text{and} \quad p_0 = \frac{1-\sqrt{\frac{2a}{1-a}}}{2}$$

Thus,

$$p_1 - p_0 = 2\sqrt{\frac{2a}{1-a}} \to 0, \text{ as } a \to 0.$$

Then $p_1 \to 1/2^+$ and $p_0 \to 1/2^-$.

Problem 6.12

Show that the efficiency (6.27) tends to 0 as $|a-\theta| \to \infty$.

Solution:

Equation 6.27 is

$$e_{2,1} = \frac{\phi^2(a-\theta)}{\Phi(a-\theta)\{1-\Phi(a-\theta)\}}.$$

Define $x = a-\theta$, and letting $x \to \infty$, we have

$$\lim_{(a-\theta)\to\infty} e_{2,1} = \lim_{x\to\infty} e_{2,1} = \lim_{x\to\infty} \frac{\phi^2(x)}{\Phi(x)\{1-\Phi(x)\}}$$
$$= \lim_{x\to\infty} \frac{\phi(x)}{\Phi(x)} \times \lim_{x\to\infty} \frac{\phi(x)}{1-\Phi(x)}.$$

Consider the limits separately, the first term is

$$\lim_{x\to\infty} \frac{\phi(x)}{\Phi(x)} = \lim_{x\to\infty} \phi(x) = \lim_{x\to\infty} e^{-\frac{x^2}{2}} = 0.$$

By L'Hospital rule, the second term is

$$\lim_{x\to\infty} \frac{\phi(x)}{1-\Phi(x)} = \lim_{x\to\infty} \frac{-x\phi(x)}{-\phi(x)} = \lim_{x\to\infty} x = \infty.$$

The speed of the first term goes to zero (exponential) is faster than the speed of second term goes to infinite. Therefore, we have

$$\lim_{(a-\theta)\to\infty} e_{2,1} = \lim_{x\to\infty} e_{2,1} = 0.$$

Similarly, we can show that $\lim_{(a-\theta)\to-\infty} e_{2,1} = \lim_{x\to-\infty} e_{2,1} = 0$.

Thus, we have proven that

$$\lim_{|a-\theta|\to\infty} e_{2,1} = 0.$$