ISyE 6412 **Theoretical Statistics** **Meen-Sung Kim**

<HW #1> - Due Wed, Aug 30. **gtID : 903813716**

Problem 1. Suppose that X is a normal random variable with variance 1 and unknown mean θ. It is desired to guess the value of unknown mean θ. Since the experimenter feels the loss is roughly like square error $(d-\theta)^2$ when the true θ is small but is like squared *relative* error $(\theta^{-1}d-1)^2$ when $|\theta|$ is large, he or she chooses loss function $(\theta-d)^2/(1+\theta^2)$ to reflect this behavior.

(a) Specify S, Ω, D, and L (i.e., the sample space, the set of all possible distribution functions, the decision space, and the loss function).

Sample space $S = \{x : -\infty < x < \infty\}$

Set of all d.f's $\Omega = \{N(\theta, 1) : \theta \in \mathbb{R}\}$

Decision space $D = \{d : -\infty < d < \infty\}$

Loss function $L(\theta, d) = (\theta - d)^2/(1+\theta^2)$ as given

(b) Determine and plot on the same graph the risk function of the 6 procedures δ_i defined by

$$\delta_1(X) = X; \qquad \delta_2(X) = (1+X)/2; \qquad \delta_3(X) = X/2;$$
$$\delta_4(X) = 2X; \qquad \delta_5(X) = 0; \qquad \delta_6(X) = 1;$$

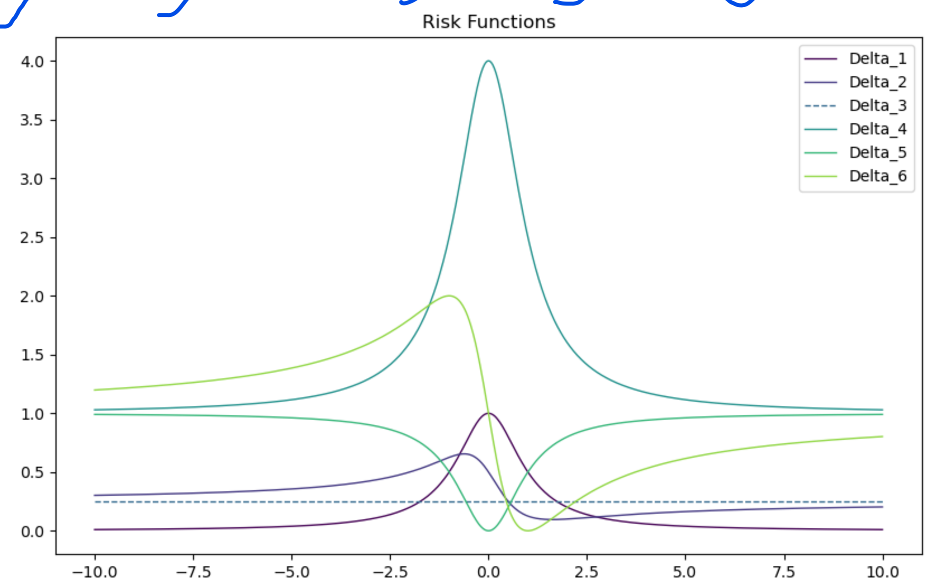
[You can save time by working (e) first but may find it easier to work (b) first. Your calculation will be made simpler if you first compute the risk function of a general procedure of the form $\delta(X) = a + bX$. A check: $R_{\delta_4}(\theta) = (\theta^2+4)/(1+\theta^2)$.]

The risk function of a general procedure of the form $\delta(X) = a + bX$ is:

$$R_\delta(\theta) = E_\theta[L(\theta, \delta(X))] = \frac{1}{(1+\theta^2)} E_\theta[a+bX-\theta]^2 = \frac{1}{1+\theta^2}[(E_\theta[a+bX]-\theta)^2 + Var(a+bX)]$$

$$= \frac{1}{1+\theta^2}[(a+b\underbrace{E_\theta(X)}_{\theta}-\theta)^2 + b^2\underbrace{Var(X)}_{=1}] = \frac{(a+(b-1)\theta)^2+b^2}{1+\theta^2}$$

Thus, $R_{\delta_1} = \frac{1}{1+\theta^2}$, $R_{\delta_2} = \frac{(1-\theta)^2+1}{4(1+\theta^2)}$, $R_{\delta_3} = \frac{1}{4}$, $R_{\delta_4} = \frac{\theta^2+4}{1+\theta^2}$, $R_{\delta_5} = \frac{\theta^2}{1+\theta^2}$, and $R_{\delta_6} = \frac{(1-\theta)^2}{1+\theta^2}$.



(c) From these calculations, can you assert that any of these six procedures is inadmissible?

We can assert that δ_4 is inadmissible. δ_1 is better than δ_4.

(d) On the basis of the risk functions, if one of these 6 procedures must be used, which procedure would you use, and why? (*Note:* Don't consult any references in answering this. Later you will find out the precise meaning of your present intuition.)

I would probably choose δ_3 because it's constant $\forall \theta$.

(e) Suppose X is replaced by the vector $(X_1, \ldots, X_n)$ of iid normal $N(\theta, 1)$ random variables. The procedures corresponding to $\delta_1, \delta_2, \delta_3, \delta_6$ are

$$\delta_{1,n}(X_1, \ldots, X_n) = \bar{X}_n; \qquad \delta_{2,n}(X_1, \ldots, X_n) = \frac{\bar{X}_n + n^{-1}}{1 + n^{-1}};$$

$$\delta_{3,n}(X_1, \ldots, X_n) = \frac{\sqrt{n}\,\bar{X}_n}{1+\sqrt{n}}; \qquad \delta_{6,n}(X_1, \ldots, X_n) = 1.$$

Compute the risk functions of these four procedures, and plot graphs of these four risk functions (or, rather, of $nR_{\delta_i,n}$ to make the results comparable to those of part (b)) for n large (e.g., for $n = 10,000$). [Use the fact that $\bar{X}_n$ is $N(\theta, n^{-1})$ distributed. Again, you may find it is easier first to find $(1+\theta^2)^{-1}\mathbf{E}_\theta(a + b\bar{X}_n - \theta)^2$ for general a, b.]

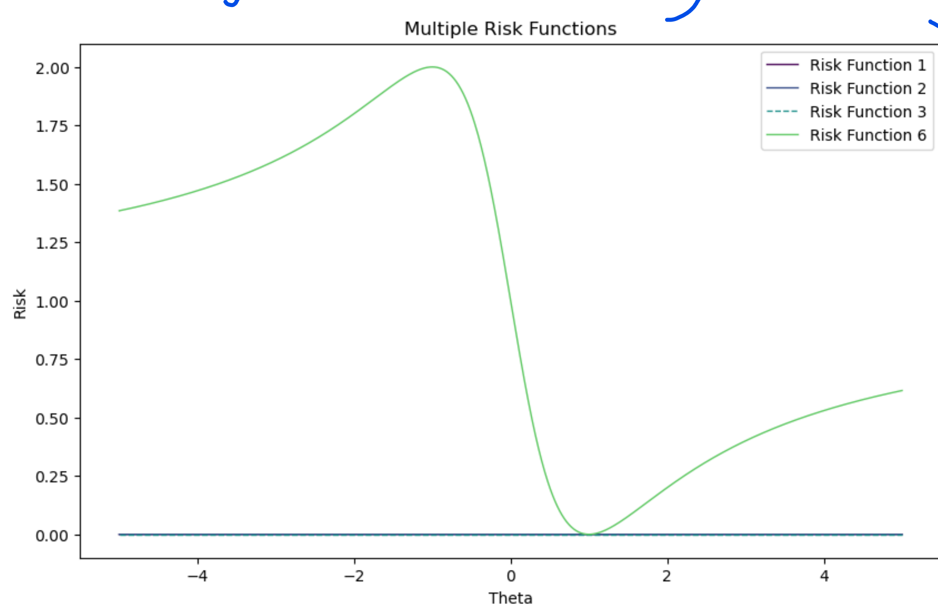
The risk function of a general procedure of the form $\delta_{a,b,n}(X_1,\ldots,X_n) = a + b\bar{X}_n$ is:

$$R_{\delta_{a,b,n}}(\theta) = E_\theta[L(\theta, \delta_{a,b,n}(X_1,\ldots,X_n))] = \frac{1}{1+\theta^2} E_\theta[a + b\bar{X}_n - \theta]^2$$

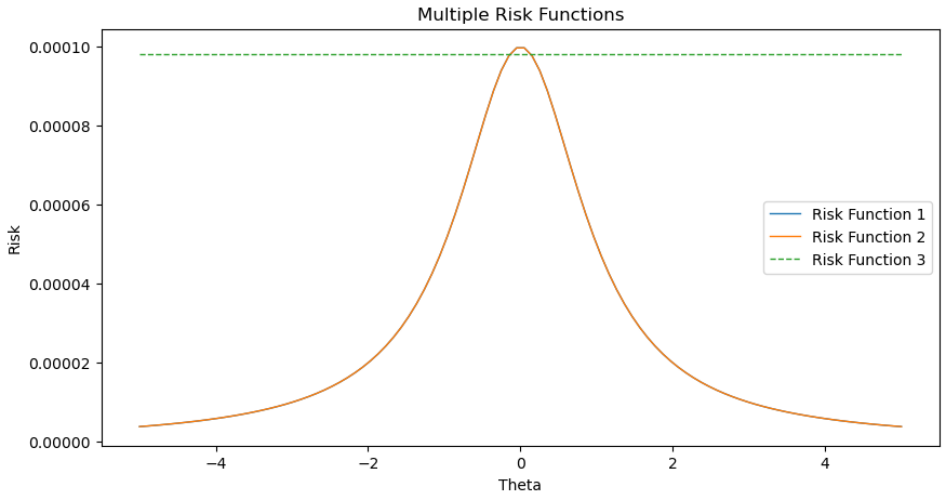
$$= \frac{1}{1+\theta^2}[(a + bE_\theta(\bar{X}_n) - \theta)^2 + Var(a + b\bar{X}_n)]$$

$$= \frac{(a+(b-1)\theta)^2 + b^2/n}{1+\theta^2}$$

Thus, $R_{\delta_{1,n}}(\theta) = \frac{1}{n(1+\theta^2)}$, $R_{\delta_{2,n}}(\theta) = \frac{(1-\theta)^2 + n}{(1+\theta^2)(n+1)^2}$, $R_{\delta_{3,n}}(\theta) = \frac{1}{(1+\sqrt{n})^2}$, and $R_{\delta_{6,n}}(\theta) = \frac{(1-\theta)^2}{1+\theta^2}$.



Below the figure is a plot without $R_{\delta_{6,n}}$, just to look closer into the rest three that look like the same line in the above graph.



(f) If n is large, which of the four procedures of part (e) would you use, and why? (Your answer to this last may differ from the answer to part (d) for the case $n = 1$; does it?)

In part (d), I chose a constant procedure δ_3. Here, the risk of $\delta_{3,n}$ is not significantly different from $\delta_{1,n}$ or $\delta_{2,n}$ when n is large. We could use either one.

(g) Suppose the statistician decides to restrict consideration to procedures $\delta_{a,b,n} = a + b\bar{X}_n$ of the form mentioned at the end of (e). He or she is concerned about the behavior of the risk function when $|\theta|$ is large. Show that the risk function approaches 0 as $|\theta| \to \infty$ if and only if $b = 1$. In addition, *among procedures with* $b = 1$, show that the choice $a = 0$ gives uniformly smallest risk function.

[This justification of the procedure $\delta_{1,n} = \bar{X}_n$ under the restriction to procedures of the form $\delta_{a,b,n}$ will seem more sensible to many people than a justification in terms of the "unbiasedness" criterion to be discussed later].

From (e), $R_{\delta_{a,b,n}}(\theta) = \dfrac{(a+(b-1)\theta)^2 + b^2/n}{1+\theta^2}$. When $b=1$, $R_{\delta_{a,1,n}}(\theta) = \dfrac{a^2 + 1/n}{1+\theta^2} \to 0$ as $|\theta| \to \infty$.

Similarly, as $|\theta| \to \infty$, $R_{\delta_{a,b,n}} \to (b-1)^2$. This is zero when $b = 1$.

Moreover, $R_{\delta_{a,1,n}}(\theta) = \dfrac{a^2 + 1/n}{1+\theta^2} \ge \dfrac{1/n}{1+\theta^2} = R_{\delta_{0,1,n}}$. $R_{\delta_{a,1,n}}$ is uniformly smallest when $a = 0$.

(h) Show that the procedure $\delta_{6,n}$, defined by $\delta_{6,n}(X_1, \ldots, X_n) \equiv 1$, is admissible for each n. [Hints: how can another procedure δ' satisfy $R_{\delta'}(\theta) \le R_{\delta_{6,n}}(\theta)$ when $\theta = 1$?]

Suppose $\delta_{6,n}(\underline{X}) = 1$ is inadmissible and there is another procedure δ' which is better than $\delta_{6,n}$. We then have $0 \le R_{\delta'}(\theta) \le R_{\delta_{6,n}}(\theta) = \dfrac{(1-\theta)^2}{1+\theta^2}$ $\forall\theta$, i.e. $E_\theta[L(\theta, \delta'(\underline{X}))] = E_\theta \dfrac{(\theta - \delta'(\underline{X}))^2}{1+\theta^2} \le E_\theta \dfrac{(\theta - \delta_{6,n}(\underline{X}))^2}{1+\theta^2}$

$\Rightarrow E_\theta(\theta - \delta'(\underline{X}))^2 \le E_\theta(\theta - \delta_{6,n}(\underline{X}))^2$ ...

incomplete

Problem 2. Assume that we observe a binomial random variable X with parameter (n, θ), i.e., the probability mass function of X is given by $\mathbf{P}(X = i) = \binom{n}{i}\theta^i(1-\theta)^{n-i}$ for $i = 0, 1, \ldots, n$, where $n \geq 1$ is a known integer and $0 \leq \theta \leq 1$ is unknown. Consider the problem of estimating θ under the so-called "absolute deviation" loss function defined by $L(\theta, d) = |\theta - d|$.

(a) Specify S, Ω, D, and L (i.e., the sample space, the set of all possible distribution functions, the decision space, and the loss function).

Sample space $S = \{0, 1, 2, \ldots, n\}$

Set of all d.f's $\Omega = \{Bin(n, \theta) : 0 \leq \theta \leq 1\}$

Decision space $D = \{d : 0 \leq d \leq 1\}$

Loss function $L(\theta, d) = |\theta - d|$ as given

(b) When $n = 20$, graph and compare the risk functions of the following three procedures:

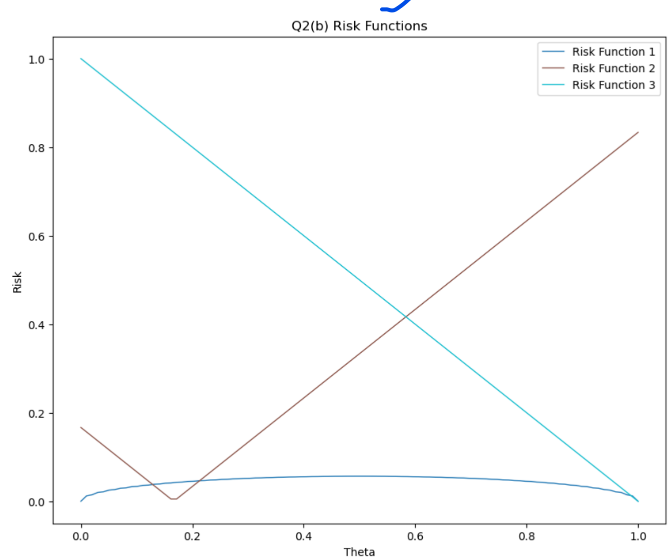
$$\delta_1(X) = \frac{X}{n}, \qquad \delta_2(X) = \frac{1}{3}, \qquad \text{and} \qquad \delta_3(X) = 1.$$

Note that the risk functions may not have simple expressions, and it will be OK to use some computer software to plot the risk functions.

$$R_{\delta_1}(\theta) = E_\theta L(\theta, \delta_1(X)) = E_\theta |\theta - X/n| = \sum_{i=0}^{n} P_\theta(X=i)|\theta - i/n| = \sum_{i=0}^{n} \binom{n}{i}\theta^i(1-\theta)^{n-i}|\theta - i/n|$$

$$R_{\delta_2}(\theta) = E_\theta L(\theta, \tfrac{1}{3}) = |\theta - \tfrac{1}{3}| = \begin{cases} \theta - \frac{1}{3}, & \text{if } \frac{1}{3} \leq \theta \leq 1 \\ \frac{1}{3} - \theta, & \text{if } 0 \leq \theta < \frac{1}{3} \end{cases}$$

$$R_{\delta_3}(\theta) = |\theta - 1| = 1 - \theta \quad \text{for} \quad 0 \leq \theta \leq 1.$$



(c) Show that for any given integer $n \geq 1$, the procedure $\delta_2(X) = 1/3$ is admissible. [Hints: how can another procedure δ' satisfy $R_{\delta'}(\theta) \leq R_{\delta_2}(\theta)$ when $\theta = 1/3$?]

Suppose $\delta_2(x) = 1/3$ is inadmissible and there is another procedure δ' which is better than δ_2. We then have $0 \leq R_{\delta'}(\theta) \leq R_{\delta_2}(\theta) = |\theta - \frac{1}{3}|$ $\forall \theta$.

When $\theta = 1/3$, $R_{\delta'}(\theta) = E_\theta L(\theta, \delta'(X)) = E_\theta |\theta - \delta'(x)| = \sum_{i=0}^{n} P_\theta(X=i)|\theta - \delta'(x)|$

$= \sum_{i=0}^{n} \binom{n}{i} \theta^i (1-\theta)^{n-i} |\theta - \delta'(x)|$

is $\sum_{i=0}^{n} \underbrace{\binom{n}{i} \frac{2^{n-i}}{3^n}}_{>0} \underbrace{|\frac{1}{3} - \delta'(x)|}_{\geq 0} \leq 0$.

From this, we have $\delta'(x) = \frac{1}{3}$ $\forall i$ $\Rightarrow R_{\delta'}(\theta) = R_{\delta_2}(\theta)$

This contradicts with our assumption that δ' is better than δ_2.

We conclude that $\delta_2(X) = \frac{1}{3}$ is admissible.

(d) Show that when $n = 2$, the procedure $\delta_3(X) = 1$ is admissible.

incomplete

Remarks: Parts (c) and (d) suggest that an admissible estimator may not be appealing. Of course, it is clear that inadmissible estimators are definitely not desirable.

In Part (b), the following R code can be used to plot the risk functions. For more information about the free statistical software R, please see the website <http://www.r-project.org/>.

```
theta <- seq(0,1,0.0001);
R1 <- 0;
for (i in 0:20){
 R1 <- R1+choose(20,i)*(theta^i)*((1-theta)^(20-i))*abs(i/20 - theta);
}
R2 <- abs(1/3  - theta);
R3 <- abs(1  - theta);
plot(theta, R1,"l", ylab="Risk Function", ylim=c(0,1));
lines(theta, R2, col="red");
lines(theta, R3, col="blue")
```

final remarks: 1h and 2d not answered. Rest are correct.
Feedback: "B", 1.6 points