Homework#6, Statistical Inference I, 2012 Fall

1. Admissibility

Let $X \sim \mathrm{P} = \{P_\theta : \theta \in \Theta \subset R^k\}$, and let $T(X)$ be a unique minimax estimator for $\vartheta = g(\theta) \in R^1$ under some loss. Show that $T(X)$ admissible.

2. Minimaxity in nonparametric models

1) Let $X \sim F \in \Im$, where $\Im$ is a set of distributions (parametric or nonparametric). Assume that $\delta^*(X)$ is minimax in $\Im$, i.e.,

$$\sup_{F\in\Im} R_{\delta^*}(F) \le \sup_{F\in\Im} R_\delta(F) \quad \text{for} \quad \forall\delta .$$

Let $\Im \subset \Im_1$. Show that $\delta^*(X)$ is minimax in $\Im_1$ if

$$\sup_{F\in\Im} R_{\delta^*}(F) = \sup_{F\in\Im_1} R_{\delta^*}(F) .$$

2) Let $X_1,..., X_n \overset{iid}{\sim} F \in \Im$, where $\mathbf{L}_2 = \{ F \mid E_F(X_1 - E_F(X_1))^2 \le c \}$ be a class of bounded variance (L_2 space). Find a minimax estimator for $\vartheta = E_F(X_1)$ under the squared error loss in $\mathbf{L}_2$.

3) Let $X_1,..., X_n \overset{iid}{\sim} F \in \Im$, where $\mathbf{L}_{[0,1]} = \{ F \mid 0 \le X_1 \le 1, \text{ with probability } 1 \}$ be a class of finite support on $[0,1]$. Find a minimax estimator for $\vartheta = E_F(X_1)$ under the squared error loss in $\mathbf{L}_{[0,1]}$.

3. Bayes estimator

Let $X = (X_1,..., X_n) \mid \boldsymbol{\mu} = \mu \overset{iid}{\sim} N(\mu, \sigma^2)$, where σ^2 is known. Consider estimation of μ under the loss $L(\mu, a) = (\mu - a)^2$ and an improper prior $d\Pi(\mu) = I(\mu > 0)d\mu$.

4) Derive the generalized Bayes estimator $\delta(X)$.

5) Find the asymptotic distribution of $\sqrt{n}\{\delta(X) - \mu\}$.

6) Under n=10, σ^2=1, $\mu \in [0.1,\ 2.5]$, draw the graph for the risk of $\delta(X)$ and $\overline{X}$ using simulations (including codes in Appendix).

7) $X=(X_1,\ldots,X_n)\overset{iid}{\sim} N(\mu,\sigma^2)$, where $\mu>0$ is unknown and σ^2 is known. Derive the MLE under the restricted space $\Theta=\{\mu \mid \mu>0\}$, and draw the graph of the risk using simulations under the same setting as (3).

8) According to the three risks, which estimator do you prefer? Give a short discussion.

4. The James-Stein estimator (I)

Based on data $X \sim N_p(\theta, I_p)$, consider estimation of $\theta=(\theta_1,\ldots,\theta_p)'$ under the loss $\|\delta(X)-\theta\|^2$. Calculate the risk of the estimator

$$\delta(X)=X-\frac{p-3}{\|X-\bar{X}J_p\|^2}(X-\bar{X}J_p),\quad J_p=(1,\ldots,1)',$$

and then show that X is inadmissible when $p\geq 4$.

5. The James-Stein estimator (II)

Let $X \mid \boldsymbol{\theta}=\theta \sim N_p(\theta,I_p),\quad \theta\in R^p$, and $\boldsymbol{\theta}\sim N_p(c,\tau^2 I_p),\quad c\in R^p$, where $\tau^2>0$ is a known hyperparameter.

1. Derive the marginal distribution of X.
2. Derive the posterior distribution of $\boldsymbol{\theta} \mid X=x$.
3. Derive the Bayes estimator of θ, which is written in terms of $1/(1+\tau^2)$.
4. Show that $(p-2)/(\|X-c\|^2)$ is an unbiased estimator of $1/(1+\tau^2)$.
5. By the plug-in unbiased estimator of $1/(1+\tau^2)$, find the empirical Bayes estimator of θ.

Answer 1:

[Proof by contradiction] Suppose that $\tilde{T}(X)$ is better than $T(X)$. Then, $R_{\tilde{T}}(\theta) \le R_T(\theta)$ for $\forall\theta$. Thus, $\sup R_{\tilde{T}}(\theta) \le \sup R_T(\theta)$ and so $\tilde{T}(X)$ is also minimax. This contradicts to the uniqueness.

Answer 2:

1) **[Proof by contradiction]** Suppose that $\delta^*(X)$ is not minimax under $\Im_1$. That is, there exist an estimator $\delta(X)$ with $\sup_{F\in\Im_1} R_\delta(F) < \sup_{F\in\Im_1} R_{\delta^*}(F)$. Then,

$$\sup_{F\in\Im} R_\delta(F) \le \sup_{F\in\Im_1} R_\delta(F) < \sup_{F\in\Im_1} R_{\delta^*}(F) = \sup_{F\in\Im} R_{\delta^*}(F).$$

This contradicts that $\delta^*(X)$ is minimax in $\Im$.

[Proof by direct calculation] For $\forall T(X)$,

$$\sup_{F\in\Im_1} R_{\delta^*}(F) = \sup_{F\in\Im} R_{\delta^*}(F) \le \sup_{F\in\Im} R_T(F) \le \sup_{F\in\Im_1} R_T(F).$$

Therefore, $\delta^*(X)$ is minimax in $\Im_1$.

2) Set $\Im = \{ F_{N(\mu,\sigma^2)} \mid (\mu,\sigma^2) \in (-\infty,\infty)\times(0,c]\}$, $\Im_1 = \mathbf{L}_2$ and $\delta^*(X) = \bar{X}$. Then, $\Im \subset \Im_1$ and $\sup_{F\in\Im} R_{\delta^*}(F) = \sup_{F\in\Im_1} R_{\delta^*}(F) = c/n$. Since $\bar{X}$ is a minimax estimator for $\mu = E(X_1)$ in $\Im$, it is minimax also in $\mathbf{L}_2$.

3) Set $\Im = \{ F \mid X_1 \sim Bin(1,p),\ p\in(0,1) \}$, $\Im_1 = \mathbf{L}_{[0,1]}$ and $\delta^*(X) = \frac{n}{n+\sqrt{n}}\bar{X} + \frac{\sqrt{n}}{n+\sqrt{n}}\frac{1}{2}$ that is minimax in $\Im$ with the constant risk $R_{\delta^*}(p) = \frac{1}{4(1+\sqrt{n})^2}$. Then, $\Im \subset \Im_1$ and it remains to show $\sup_{F\in\Im_1} R_{\delta^*}(F) = \frac{1}{4(1+\sqrt{n})^2}$. By direct calculations,

$$R_{\delta^*}(F) = E_F\left(\frac{n}{n+\sqrt{n}}\bar{X} + \frac{\sqrt{n}}{n+\sqrt{n}}\frac{1}{2} - \vartheta\right)^2 = \frac{1}{(1+\sqrt{n})^2}\left[Var_F(X_1) + \left(\frac{1}{2}-\vartheta\right)^2\right].$$

Since $Var_F(X_1) = E_F(X_1^2) - \vartheta^2 \le E_F(X_1) - \vartheta^2 = \vartheta - \vartheta^2$, $R_{\delta^*}(F) \le \frac{1}{4(1+\sqrt{n})^2}$.

Hence, $\sup_{F\in\Im_1} R_{\delta^*}(F) = \frac{1}{4(1+\sqrt{n})^2}$ holds.

Answer 3:

1. The (improper) posterior density is

$$f_\mu(x)d\Pi(\mu) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2\right\} I(\mu>0)d\mu$$

$$\propto \exp\left\{-\frac{n}{2\sigma^2}(\mu-\bar{x})^2\right\} I(\mu>0)d\mu.$$

The generalized Bayes action is the minimizer of

$$\varphi(a) = \int (\mu-a)^2 f_\mu(x)d\Pi(\mu)$$

$$\propto \int_0^\infty (\mu-a)^2 \exp\left\{-\frac{n}{2\sigma^2}(\mu-\bar{x})^2\right\} d\mu.$$

Hence,

$$a = \frac{\int_0^\infty \mu \exp\left\{-\frac{n}{2\sigma^2}(\mu-\bar{x})^2\right\} d\mu}{\int_0^\infty \exp\left\{-\frac{n}{2\sigma^2}(\mu-\bar{x})^2\right\} d\mu} = \frac{\int_{\sqrt{n}\bar{x}/\sigma}^\infty \left(\bar{x}+\frac{\sigma}{\sqrt{n}}z\right)\exp\left\{-\frac{z^2}{2}\right\} d\mu}{\int_{\sqrt{n}\bar{x}/\sigma}^\infty \exp\left\{-\frac{z^2}{2}\right\} d\mu}$$

$$= \bar{x} + \frac{\sigma}{\sqrt{n}} \frac{\int_{\sqrt{n}\bar{x}/\sigma}^\infty z\exp\left\{-\frac{z^2}{2}\right\} d\mu}{\int_{\sqrt{n}\bar{x}/\sigma}^\infty \exp\left\{-\frac{z^2}{2}\right\} d\mu} = \bar{x} + \frac{\sigma}{\sqrt{n}} \frac{\Phi'\left(\frac{\sqrt{n}\bar{x}}{\sigma}\right)}{1-\Phi\left(-\frac{\sqrt{n}\bar{x}}{\sigma}\right)}.$$

The Bayes action is

$$\delta(X) = \bar{X} + \frac{\sigma}{\sqrt{n}} \frac{\Phi'\left(\frac{\sqrt{n}\bar{X}}{\sigma}\right)}{1-\Phi\left(-\frac{\sqrt{n}\bar{X}}{\sigma}\right)}.$$

2. Note that $\Phi\left(-\frac{\sqrt{n}\bar{X}}{\sigma}\right) \to_P 0$ and $\Phi'\left(\frac{\sqrt{n}\bar{X}}{\sigma}\right) \to_P 0$. It follows that

$$\sqrt{n}\{\delta(X)-\mu\} = \sqrt{n}(\bar{X}-\mu) + \sigma \frac{\Phi'\left(\frac{\sqrt{n}\bar{X}}{\sigma}\right)}{1-\Phi\left(-\frac{\sqrt{n}\bar{X}}{\sigma}\right)}$$

$$= \sqrt{n}(\bar{X}-\mu) + o_P(1) \to_d N(0,\sigma^2).$$

3. 4. 5. See below

Overall, the Bayes action provides the smallest MSE while the sample mean has the largest MSE.

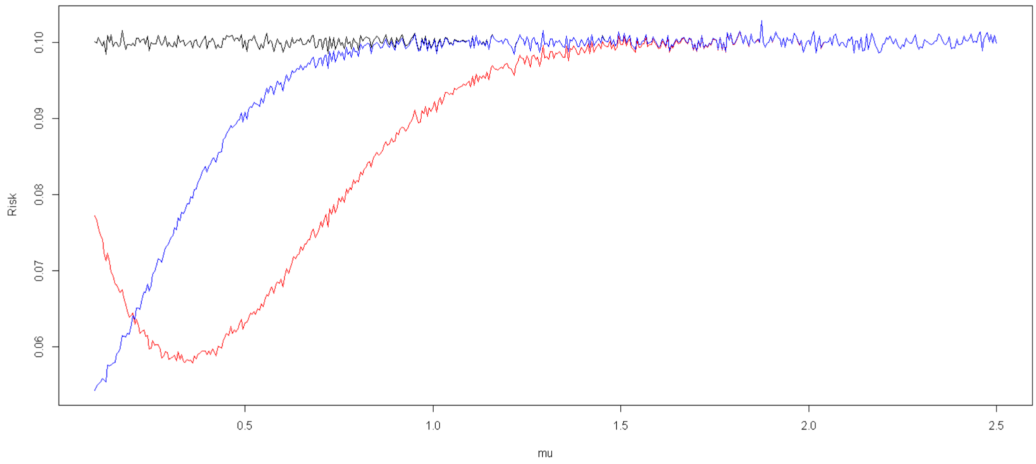

Fig. ----- $\bar{X}$; ------ $\delta(X)$; ----- $\max(\bar{X},0)$

```
n=10
R=50000
N=500
mu_vec=seq(0.1,2.5,length=N)
R_bar=R_Bayes=R_MLE=numeric(N)

for(j in 1:N){

   mu=mu_vec[j]

   bar=Bayes=MLE=numeric(R)
   for(i in 1:R){
      x=rnorm(n,mean=mu)
      x_bar=mean(x)
      bar[i]=x_bar
      Bayes[i]=x_bar+dnorm(sqrt(n)*x_bar)/(1-pnorm(-sqrt(n)*x_bar))/sqrt(n)
      MLE[i]=max(x_bar,0)
   }

   R_bar[j]=mean((bar-mu)^2)
   R_Bayes[j]=mean((Bayes-mu)^2)
   R_MLE[j]=mean((MLE-mu)^2)

}

Min=min(R_bar,R_Bayes,R_MLE)
Max=max(R_bar,R_Bayes,R_MLE)
plot(mu_vec,R_bar,type="l",xlab="mu",ylab="Risk",ylim=c(Min,Max))
points(mu_vec,R_Bayes,type="l",col="red")
points(mu_vec,R_MLE,type="l",col="blue")
```

Answer 4:

One can write $\delta(X) = X - g(X)$ where $g(X) = (g(X_1),..., g(X_p))'$ and

$g_i(X) = \dfrac{p-3}{\| X - \bar{X}J_p \|}(X_i - \bar{X})$. Then,

$$\frac{\partial}{\partial X_i} g_i(X)$$

$$= (p-3)\left[\frac{(1-1/p)}{\| X - \bar{X}J_p \|^2} - \frac{(X_i - \bar{X}J_p)}{(\| X - \bar{X}J_p \|^2)^2}\left\{-\frac{2}{p}\sum_j (X_j - \bar{X}) + 2(X_i - \bar{X})\right\}\right]$$

$$= (p-3)\left[\frac{(1-1/p)}{\| X - \bar{X}J_p \|^2} - \frac{2(X_i - \bar{X}J_p)}{(\| X - \bar{X}J_p \|^2)^2}\right]$$

and

$$\sum_{i=1}^{p} \frac{\partial}{\partial X_i} g_i(X) = \frac{(p-3)^2}{\| X - \bar{X}J_p \|^2}.$$

By Corollary 7.2 of Lehmann & Casella (p.273 of the book), we have

$$R_\delta(\theta) = E_\theta \| X - g(X) \|^2 = p + E_\theta \| g(X) \|^2 - 2\sum_{i=1}^{p} E_\theta\left[\frac{\partial}{\partial X_i} g_i(X)\right]$$

$$= p - (p-3)^2 E_\theta\left[\frac{1}{\| X - g(X) \|^2}\right]$$

Hence, if $p \geq 4$, then $R_\delta(\theta) < p = R_X(\theta)$

Answer 5:

1. $X \sim N_p(c, (1+\tau^2)I_p)$.

   __Proof__: Since $X_1, ..., X_p$ are iid, we only need to derive the marginal distribution of X_j. It can be shown (calculations omitted) that,

$$f_{X_j}(x) = \int f_{X_j|\theta_j}(x_j) f_{\theta_j}(\theta_j) d\theta_j = \frac{1}{\sqrt{2\pi(\tau^2+1)}} \exp\left\{-\frac{(x_j - c_j)^2}{2(\tau^2+1)}\right\}.$$

   Hence, $X_j \sim N(c_j, (1+\tau^2))$.

2. $\boldsymbol{\theta} \mid X = x \sim N_p\left(\frac{\tau^2 x + c}{1+\tau^2}, \frac{\tau^2}{1+\tau^2} I_p\right)$.

$$f_{\boldsymbol{\theta}}(x)d\Pi_\tau(\theta)$$

$$\propto \exp\left(-\frac{1}{2}\|x-\theta\|^2\right)\exp\left(-\frac{1}{2\tau^2}\|\theta-c\|^2\right) \propto \exp\left\{-\frac{1}{2}\left(\|x-\theta\|^2+\frac{1}{\tau^2}\|\theta-c\|^2\right)\right\}$$

$$\propto \exp\left[-\frac{1}{2}\left(\left(1+\frac{1}{\tau^2}\right)\|\theta\|^2 - 2\left(x+\frac{c}{\tau^2}\right)'\theta\right)\right] \propto \exp\left\{-\frac{1}{2\left(1+\frac{1}{\tau^2}\right)^{-1}}\left(\|\theta\|^2 - 2\frac{\left(x+\frac{c}{\tau^2}\right)'\theta}{\left(1+\frac{1}{\tau^2}\right)}\right)\right\}$$

$$\propto \exp\left[-\frac{1}{2\left(\frac{\tau^2}{1+\tau^2}\right)}\left(\left\|\theta-\frac{\tau^2 x+c}{1+\tau^2}\right\|^2\right)\right]$$

3. $\delta^B(X) = \dfrac{\tau^2 X + c}{1+\tau^2} = X - \dfrac{1}{1+\tau^2}(X-c)$

4. Since $X \sim N_p(c,(1+\tau^2)I_p)$, $\dfrac{\|X-c\|^2}{1+\tau^2} \sim \chi_p^2$ and $E\left(\dfrac{p-2}{\|X-c\|^2}\right) = \dfrac{1}{1+\tau^2}$.

5. $\delta^{EB}(X) = X - \dfrac{p-2}{\|X-c\|^2}(X-c)$ (the James-Stein estimator)