Chapter 5. Minimax Analysis (CI)

Example 3: Assume $X \sim N(\theta, \sigma^2)$, σ^2 known, is observed, and that it is desired to test H_0: $\theta \leq \theta_0$ versus H_1: $\theta > \theta_0$, under "$0-K_i$" loss. A randomized decision rule in a testing situation is usually represented by a test function $\phi(x)$, where $\phi(\boldsymbol{x})$ is the probability of rejecting the null hypothesis when $X = x$ is observed. Thus,

$$\phi(x) = \delta^*(x, a_1) = 1 - \delta^*(x, a_0),$$

where a_i denotes accepting H_i.

It will be seen in Chapter 8 that only the "uniformly most powerful" tests $\phi_c(x) = I_{(c,\infty)}(x)$ need be considered. Any other test can be improved upon by the test ϕ_c, for some c. Clearly

$$R(\theta, \phi_c) = \begin{cases} K_1 P_\theta(X > c) & \text{if } \theta \leq \theta_0, \\ K_0 P_\theta(X < c) & \text{if } \theta > \theta_0. \end{cases}$$

Since

$$\beta(\theta) = P_\theta(X > c) = 1 - \Phi\left(\frac{c - \theta}{\sigma}\right)$$

is increasing in θ, it follows that

$$\begin{aligned}\sup_\theta R(\theta, \phi_c) &= \max\left\{\sup_{\theta \le \theta_0} K_1\beta(\theta), \sup_{\theta > \theta_0} K_0(1-\beta(\theta))\right\} \\ &= \max\{K_1\beta(\theta_0), K_0[1-\beta(\theta_0)]\}.\end{aligned}$$

Letting $z = \beta(\theta_0)$, and graphing the functions $K_1 z$ and $K_0(1-z)$, it becomes clear that

$$\max\{K_1 z, K_0(1-z)\}$$

is minimized when $K_1 z = K_0(1-z)$, or $z = K_0/(K_0+K_1)$. But if

$$\beta(\theta_0) = P_{\theta_0}(X > c) = K_0/(K_0+K_1),$$

it follows that c is the $1 - K_0/(K_0+K_1) = K_1/(K_0+K_1)$-fractile of the $N(\theta_0, \sigma^2)$ distribution. Denoting this by $z(K_1/(K_0+K_1))$, it can be concluded that

$$\begin{aligned}\inf_{\delta^* \in \mathcal{D}^*} \sup_\theta R(\theta, \delta^*) &= \inf_c \sup_\theta R(\theta, \delta^*) \\ &= \sup_\theta R(\theta, \phi_{z(K_1/(K_0+K_1))}).\end{aligned}$$

Hence $\phi_{z(K_1/(K_0+K_1))}$ is minimax. □

Example 4: Assume $X \sim \mathcal{E}(\theta)$ is observed, and it is desired to test H_0: $0 < \theta = 1$ versus H_1: $2 \leq \theta < \infty$. Let a_i denote accpeting H_i. The loss function is given by

$$L(\theta, a_0) = \begin{cases} 0 & \text{if } 0 < \theta \leq 1, \\ 4 & \text{if } \theta \geq 2, \end{cases}$$

$$L(\theta, a_1) = \begin{cases} 0 & \text{if } \theta \geq 2, \\ 5 - \theta & \text{if } 0 < \theta \leq 1. \end{cases}$$

It seems plausible that the least favorable prior is one which makes H_0 and H_1 as hard to distingush as possible, namely one which gives positive probability only to the points $\theta = 1$ and θ_2. Assume that π is such a prior, and let $\pi_1 = \pi(1) = 1 - \pi(2)$. We still must find the least favorable choice of π_1.

To find the Bayes rule with respect to π, note that the posterior expected losses of a_0 and a_1 are (letting $m(x)$ denote the marginal density of X)

$$E^{\pi(\theta|x)}[L(\theta, a_0)] = \frac{4(1 - \pi_1)(\frac{1}{2}e^{-x/2})}{m(x)}$$

and

$$E^{\pi(\theta|x)}[L(\theta, a_1)] = \frac{4\pi_1 e^{-x}}{m(x)}.$$

Thus the Bayes rule is to choose a_0 if

$$\frac{(1-\pi_1)(e^{-x/2})}{m(x)} < \frac{2\pi_1 e^{-x}}{m(x)},$$

which can be rewritten as

$$x < 2\log\left[\frac{2\pi_1}{(1-\pi_1)}\right] = c.$$

Therefore, we will consider the class of rules defined by the test functions $\pi_c(x) = I_{(c,\infty)}(x)$.

To find the least favorable π_1, or alternatively the least favorable c, note that

$$R(\theta, \phi_c) = \begin{cases} \int_c^\infty (5-\theta)\theta^{-1}e^{-x/\theta}dx & \text{if } 0 < \theta \le 1, \\ \int_c^\infty 4\theta^{-1}e^{-x/\theta}dx & \text{if } \theta \ge 2, \end{cases}$$

$$= \begin{cases} (5-\theta)e^{-c/\theta} & \text{if } 0 < \theta \le 1, \\ 4(1-e^{-c/\theta}) & \text{if } \theta \ge 2. \end{cases}$$

Define

$$h(c) = \sup_{0<\theta\le 1} (5-\theta)e^{-c/\theta}$$

and

$$g(c) = \sup_{2 \leq \theta < \infty} 4(1 - e^{-c/\theta}).$$

It follows that

$$\bar{R}(c) = \sup_{\theta \in \Theta} R(\theta, \theta_c) = \max\{h(c), g(c)\}.$$

Note that $h(c)$ is strictly decreasing in c with $h(0) = 5$ and $\lim_{c \to \infty} h(c) = 0$. Aslo, $g(c)$ is strictly increasing in c with $g(0) = 0$ and $\lim_{c \to \infty} g(c) = 4$. Hence $h(c)$ and $g(c)$ are equal just for one value of c, called it c_0, and $\bar{R}(c_0) = \inf_c \bar{R}(c)$. The test $\phi_{c_0}(x)$ is a good candidate for a minimax test.

To find c_0, note first that $4(1 - e^{-c/\theta})$ is decreasing in θ, so that $g(c) = 4(1 - e^{-c/2})$. Next observe that

$$\begin{aligned} \frac{d}{d\theta}[(5 - \theta)e^{-c/2}] =& - e^{-c/\theta} + (5 - \theta)c\theta^{-2}e^{-c/\theta} \\ =& \theta^{-2}e^{-c/\theta}(-\theta^2 - \theta c + 5c). \end{aligned}$$

This derivative is positive for $0 < \theta \leq 1$, providing $\theta^2 + \theta c - 5c < 0$. The roots of the equation $\theta^2 + \theta c - 5c = 0$ are

$$\frac{1}{2}(-c \pm [c^2 + 20c]^{1/2}),$$

one of which is negative, while the other is larger than

1 for $c > \frac{1}{4}$. Hence if $\frac{1}{4}$ and $0 < \theta \leq 1$, it follows that $\theta^2 + \theta c - 5c < 0$, and $(5 - \theta)e^{-c/2}$ is maximized at $\theta = 1$. Thus $h(c) = 4e^{-c}$ for $c > \frac{1}{4}$.

Let's assume the solution to $h(c) = g(c)$ is some $c > 1/4$. Then we want to solve the equation

$$4e^{-c} = 4(1 - e^{-c/2}).$$

The solution is $e^{-c/2} \approx 0.618$, which corresponds to $c \approx 0.96$. Since $0.96 > 3/4$, this is indeed the unique solution. Note that $c = 0.96$ corresponds to $\pi_1 \approx 0.45$.

Observe finally that for $\pi_1 = 0.45$,

$$\begin{aligned} r(\pi) =& r(\pi, \phi_{0.96}) = \pi_1 R(1, \phi_{0.96}) + (1 - \pi_1)R(2, \phi_{0.96}) \\ =& \pi_1 h(0.96) + (1 - \pi_1)g(0.96) = h(0.96). \end{aligned}$$

Also, $R(\theta, \phi_{0.96}) \leq h(0.96)$ for all θ. Hence, Theorem 3 can be used to conclude that $\phi_{0.96}$ is a minimax and π is least favorable. $\Box$.

• Minimax Estimators of a Normal Mean Vector

Suppose $\boldsymbol{X} \sim N_p(\boldsymbol{\theta}, \Sigma)$, Σ known, is observed, and that $L(\boldsymbol{\theta}, \boldsymbol{\delta}) = (\boldsymbol{\theta} - \boldsymbol{\delta})'Q(\boldsymbol{\theta} - \boldsymbol{\delta})$, Q a known $(p \times p)$ positive definite matrix.

The standard estimator is $\boldsymbol{\delta}^0(\boldsymbol{x}) = \boldsymbol{x}$. It is easy to see that $\boldsymbol{\delta}^0(\boldsymbol{x})$ has a constant risk. It can also be shown that $\boldsymbol{\delta}^0(\boldsymbol{x})$ is minimax.

For $p = 1$ and $p = 2$, $\boldsymbol{\delta}^0(\boldsymbol{x})$ is admissible. But, for $p > 2$, it is not admissible, so that other minimax estimators exist.

It can be shown that

$$\boldsymbol{\delta}(\boldsymbol{x}) = \boldsymbol{x} - \frac{r(||\boldsymbol{x} - \boldsymbol{\mu}||^2)}{|\boldsymbol{x} - \boldsymbol{\mu}||^2} Q^{-1}\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu})$$

is minimx, where r is a continuous, positive, piecewise differentiable function which is bounded by $2(p-2)$.

Among the interesting examples of minimax $\boldsymbol{\delta}$ are

$$\boldsymbol{\delta}(\boldsymbol{x}) = \boldsymbol{x} - \frac{p-2}{||\boldsymbol{x} - \boldsymbol{\mu}||^2} Q^{-1}\Sigma^{-1}(\boldsymbol{x} - \boldsymbol{\mu}),$$

which reduces to the James-Stein estimator when $\boldsymbol{\mu} = 0$ and $\Sigma = Q = I$.

- **A Brief Discussion on Rationality and the Minimax Principle**

Discuss Figure 5.8 and Example 23.

- **Minimax Regret**

Let

$$L^*(\theta, a) = L(\theta, a) - \inf_{a \in \mathcal{A}} L(\theta, a).$$

The function L^* is called *regret loss* (since it gives the regret we have for not using the best action.

Regret loss can be dealt with exactly as is a usual loss. Defining

$$R^*(\theta, \delta^*) = E_\theta[L^*(\theta, \delta^*(\boldsymbol{X}, \cdot))],$$

where for a randomized rule δ^*

$$L^*(\theta, \delta^*(\boldsymbol{x}, \cdot)) = E^{\delta^*(x, \cdot)}[L^*(\theta, a)].$$

Minimax Regret Principle: A rule δ_1^* is preferred to δ_2^* if

$$\sup_\theta R^*(\theta, \delta_1^*) < \sup_\theta R^*(\theta, \delta_2^*).$$

Definition: A minimax regret decision rule δ^{*MR} is a rule which is minimax for the loss L^*.