STAT 517

Minimax and Classification Procedures

Prof. Michael Levine

March 31, 2016

Definition

- Now is the time to consider decision theory as it is used in testing...
- Task: to find a minimax procedure that yields a best test of H_0 vs. H_1
- Let $X_1, \dots, X_n \sim f(x;\theta)$ with the likelihood $L(\theta)$
- The hypotheses are simple: $H_0 : \theta = \theta^{'}$ vs. $\theta = \theta^{"}$...so $\Omega = \{\theta : \theta = \theta^{'}, \theta^{"}\}$
- A decision function is $\delta : \delta = \theta^{'}$ or $\delta = \theta^{"}$

Loss function

- The loss function is $\mathcal{L}(\theta, \delta)$ is defined so that $\mathcal{L}(\theta^{'}, \theta^{'}) = \mathcal{L}(\theta^{''}, \theta^{''}) = 0$
- On the contrary, $\mathcal{L}(\theta^{'}, \theta^{"}) > 0$ and $\mathcal{L}(\theta^{'}, \theta^{"}) > 0$
- Think in terms of choosing a subset C of the sample space s.t. if $\mathbf{x} \in C$, we decide $\delta = \theta^{"}$ and if $\mathbf{x} \in C^{'}$, we decide that $\delta = \theta^{"}$

A minimax solution

- Need to find a critical region C s. t. $\max[R(\theta^{'}, C), R(\theta^{''}, C)]$ is minimized
- Remember the "best of the worst" adage...
- The solution is the region

$$C = \left\{(x_1, \ldots, x_n) : \frac{L(\theta^{'}; x_1, \ldots, x_n)}{L(\theta^{''}; x_1, \ldots, x_n)} \leq k\right\}$$

- $k > 0$ is selected so that $R(\theta^{'}, C) = R(\theta^{''}, C)$

Risk function

- Define

$$R(\theta, C) = R(\theta, \delta) = \int_{C \cup C'} \mathcal{L}(\theta, \delta) L(\theta)$$

- Note that this is the same as

$$R(\theta, C) = \int_{C} \mathcal{L}(\theta, \theta^{"}) L(\theta) + \int_{C'} \mathcal{L}(\theta, \theta^{'}) L(\theta)$$

- Simple algebra suggests that

$$R(\theta^{'}, C) = \mathcal{L}(\theta^{'}, \theta^{"}) \alpha$$

and

$$R(\theta^{"}, C) = \mathcal{L}(\theta^{"}, \theta^{'}) \beta$$

Example

- Take $X_1, \ldots, X_{100} \sim N(\theta, 100)$
- Test $H_0 : \theta = 75$ vs. $H_1 : \theta = 78$
- Find the minimax solution with $\mathcal{L}(75, 78) = 3$ and $\mathcal{L}(78, 75) = 1$
- $L(75)/L(78) \leq k$ is the same as $\bar{x} \geq c$...so find c and k s.t.

$$3P(\bar{X} \geq c; \theta = 75) = P(\bar{X} < c; \theta = 78)$$

- Given that $\bar{X} \sim N(\theta, 1)$ the equivalent is

$$3[1 - \Phi(c - 75)] = \Phi(c - 78)$$

Classification

- **classification** problem is having to place an item into one of several categories
- The decision is made based on several measurements on the item; assume just two, X and Y,for simplicity
- The joint pdf is $f(x,y;\theta)$ with two values $H_0:\theta=\theta^{'}$ vs. $H_a:\theta=\theta^{''}$
- By Neyman-Pearson, a best decision rule is to reject H_0 if

$$\frac{f(x,y;\theta^{'})}{f(x,y;\theta^{''})}\le k$$

Choice of k

- If π is the probability of the first choice and $\pi^{''}$ is that of a second, we have $\pi^{'} + \pi^{''} = 1$
- Can be shown that the optimal classification rule is to choose $k = \frac{\pi^{''}}{\pi^{'}}$
- In practice, the common choice is $k = 1$

Example

- Take (X, Y) to be bivariate normal with parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$
- Assume σ_1^2,σ_2^2, and ρ are known; the decision is between $(\mu_1^{'}, \mu_2^{'})$ and $(\mu_1^{''}, \mu_2^{''})$
- The decision rule is linear: if

$$ax + by \leq c$$

for some a,b, and c