Review for Final Exam

Chapter 8. Optimal Tests of Hypothesis

§ 8.1 Most Powerful Tests

- **Hypothesis testing** (general setup): Let $X_1, \ldots, X_n$ be i.i.d.$\sim f(x;\theta)$, $\theta \in \Omega = \Omega_0 \bigcup \Omega_1$, where $\Omega_0 \bigcap \Omega_1 = \emptyset$. Let $\mathcal{S}$ be the support of $\mathbf{X} = (X_1, \ldots, X_n)'$. We want to test the *null hypothesis* $H_0 : \theta \in \Omega_0$ versus the *alternative hypothesis* $H_1 : \theta \in \Omega_1$.
(1) Critical region (rejection region): C$\subset \mathcal{S}$ such that,
we reject H_0 if and only if $\mathbf{x} = (x_1, \ldots, x_n)' \in \mathcal{C}$C .
(2) Size of the test (significance level, Type I error):
$\alpha = \max_{\theta \in \Omega_0} P_\theta(\mathbf{X} \in \mathcal{C})$.
(3) Power function: $\gamma_{\mathcal{C}}(\theta) = P_\theta(\mathbf{X} \in \mathcal{C})$, $\theta \in \Omega_1$.

- **Best critical region**: To test $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$, let C be the critical region, which is a subset of $\mathcal{S} \subset R^n$. We say that $\mathcal{C}$ is a *best critical region of size* α, $0 < \alpha < 1$ if
(1) $P_{\theta_0}(\mathbf{X} \in \mathcal{C}) = \alpha$;
(2) For any other critical region $A \subset \mathcal{S}$ of the same size α, we must have
$P_{\theta_1}(\mathbf{X} \in \mathcal{C}) \geq P_{\theta_1}(\mathbf{X} \in \mathcal{A})$.
In other words, $\mathcal{C}$ is the most powerful critical region of size α . The test based on $\mathcal{C}$ is called the *most powerful test of size* α .

- **Theorem (Neyman-Pearson)**: Let $X_1, \ldots, X_n$ be i.i.d.$\sim f(x;\theta)$, $\theta \in \{\theta_0, \theta_1\}$. The likelihood function $L(\theta;\mathbf{x}) = \prod_{i=1}^n f(x_i;\theta)$, for $\mathbf{x} = (x_1, \ldots, x_n)' \in \mathcal{S}$. Let $\mathcal{C}$ be a subset of $\mathcal{S}$ and let k be a positive number such that
(a) $L(\theta_0;\mathbf{x})/L(\theta_1;\mathbf{x}) \leq k$ for each $\mathbf{x} \in \mathcal{C}$;
(b) $L(\theta_0;\mathbf{x})/L(\theta_1;\mathbf{x}) \geq k$ for each $\mathbf{x} \notin \mathcal{C}$;
(c) $\alpha = P_{H_0}(\mathbf{X} \in \mathcal{C})$.
Then $\mathcal{C}$ is a best critical region of size α for testing the simple hypothesis $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$.
Note: (1) The conditions (a), (b), and (c) are also necessary for region $\mathcal{C}$ to be a best critical region of size α .
(2) In the case of continuous distributions, the best critical region $\mathcal{C}$ of size α is unique in the probability sense.

- **Theorem (power of test)**: Let $\mathcal{C}$ be the best critical region of size α for testing $H_0 : \theta = \theta_0$ versus $H_1 : \theta = \theta_1$. Let $\gamma_{\mathcal{C}}(\theta_1) = P_{\theta_1}(\mathbf{X} \in \mathcal{C})$ denote the power of the test based on $\mathcal{C}$. Then $\gamma_{\mathcal{C}}(\theta_1) \geq \alpha$.
  In other words, a lower bound of the power of the most powerful test of size α is α .

- **Theorem (nonparametric case)**: Let $X_1, \ldots, X_n$ be an arbitrary sample. It is desired to test the simple hypothesis "H_0 : the joint pdf (or pmf) is $g(x_1, \ldots, x_n)$" versus "H_1 : the joint pdf (or pmf) is $h(x_1, \ldots, x_n)$" . Then $\mathcal{C} \subset R^n$ is a best critical region of size α if, for $k > 0$,
  (1) $g(x_1, \ldots, x_n)/h(x_1, \ldots, x_n) \leq k$ for $(x_1, \ldots, x_n)' \in \mathcal{C}$;
  (2) $g(x_1, \ldots, x_n)/h(x_1, \ldots, x_n) \geq k$ for $(x_1, \ldots, x_n)' \notin \mathcal{C}$
  (3) $\alpha = P_{H_0}[(X_1, \ldots, X_n)' \in \mathcal{C}]$.

- For practice: Example 8.1.2

§ 8.2 Uniformly Most Powerful Tests

- **UMP critical region**: A critical region $\mathcal{C}$ is called a *uniformly most powerful* (UMP) critical region of size α for testing $H_0 : \theta \in \Omega_0$ against $H_1 : \theta \in \Omega_1$ if, for each $\theta_1 \in \Omega_1$, $\mathcal{C}$ is a best critical region of size α for testing H_0 against $H_1' : \theta = \theta_1$.
  The test based on the UMP critical region $\mathcal{C}$ is called a UMP test.

- **Monotone likelihood ratio**: The likelihood function $L(\theta; \mathbf{x})$, $\mathbf{x} = (x_1, \ldots, x_n)'$, is said to have *monotone likelihood ratio* (mlr) in the statistic $Y = u(X_1, \ldots, X_n)$ if $L(\theta_1; \mathbf{x})/L(\theta_2; \mathbf{x})$ is a monotone function of $y = u(x_1, \ldots, x_n)$ as long as $\theta_1 < \theta_2$.

- Two-step standard procedure for finding a UMP test of size α for testing $H_0 : \theta = \theta_0$ against $H_1 : \theta \in \Omega_1$, where Ω_1 might be $\theta > \theta_0$, $\theta < \theta_0$, or $\theta \neq \theta_0$:
  Step 1: For each fixed $\theta_1 \in \Omega_1$, find a best critical region $\mathcal{C}$ of size α for testing H_0 against $H_1' : \theta = \theta_1$ based on the Neyman-Pearson theorem.
  Step 2: Check if $\mathcal{C}$ depends on θ_1 . If it does not, then $\mathcal{C}$ is a UMP critical region of size α for testing H_0 against H_1; otherwise there is no UMP test for this case.

- Theorem: If $L(\theta; \mathbf{x})$ has mlr in the statistic $Y = u(\mathbf{X})$, then a UMP test for $H_0 : \theta \leq \theta_0$ against $H_1 : \theta > \theta_0$ exists. Furthermore,
  (1) if it is monotone increasing, the UMP critical region takes the form of $\{(x_1, \ldots, x_n) : u(x_1, \ldots, x_n) \leq \mathcal{C}\}$;
  (2) if it is monotone decreasing, the UMP critical region takes the form of $\{(x_1, \ldots, x_n) : u(x_1, \ldots, x_n) \geq \mathcal{C}\}$.
  Note: The case of $H_0 : \theta \geq \theta_0$ against $H_1 : \theta < \theta_0$ is similar.

- Theorem: Let $X_1, \ldots, X_n$ be i.i.d.$\sim f(x;\theta)$, where

$$f(x;\theta) = \exp\{p(\theta)K(x) + S(x) + q(\theta)\}$$

belongs to the regular exponential class. If $p(\theta)$ is monotone, then the likelihood function $L(\theta;\mathbf{x})$ has mlr in $Y = \sum_{i=1}^{n} K(X_i)$.
For example, if $p(\theta)$ is monotone increasing, then $L(\theta;\mathbf{x})$ has monotone decreasing likelihood ratio in Y .

- For practice: Example 8.2.1, Example 8.2.2, Example 8.2.5

§ 8.3 Likelihood Ratio Tests

- **Unbiased test**: A test for $H_0 : \theta \in \Omega_0$ against $H_1 : \theta \in \Omega_1$ is said to be *unbiased*, if its power never falls below the significance level. In other words, if $\alpha = \max_{\theta\in\Omega_0} P_\theta[\text{ reject } H_0]$, then $P_\theta[\text{ reject } H_0] \geq \alpha$ for each $\theta \in \Omega_1$.

- **Likelihood ratio test**: For testing $H_0 : \theta \in \Omega_0$ against $H_1 : \theta \in \Omega_1$, the *likelihood ratio test* statistic is

$$\Lambda = \frac{\max_{\theta\in\Omega_0} L(\theta;\mathbf{x})}{\max_{\theta\in\Omega} L(\theta;\mathbf{x})} \ ,$$

where $\Omega = \Omega_0 \cup \Omega_1$.
Note that $0 < \Lambda \leq 1$. If H_0 is true, Λ should be close to 1; if H_1 is true, Λ should be smaller.

- **Likelihood ratio principle**: Reject H_0 if and only if $\Lambda \leq \lambda_0$, where $\lambda_0 < 1$ is a constant determined by the significance level α such that $P_{\theta_0}(\Lambda \leq \lambda_0) = \alpha$, where θ_0 is the boundary point of Ω_0 and Ω_1.

- ***p*-value**: The so-called *p-value* is the probability that the test statistic under H_0 is at least as extreme as the particular observed value.
A small enough p-value indicates the rejection of H_0 .

- For practice: Example 8.3.1, Example 8.3.2