Math 281A Homework 7 Solution

1. Let $\{X_i\}_{i=1}^n$ be i.i.d. from Pareto distribution with density
$$f(x) = \frac{\theta c^\theta}{x^{\theta+1}} I\{x \geq c\},$$
for known constant $c > 0$ and parameter $\theta > 0$. Derive the Wald, Rao, and likelihood ratio tests of $\theta = \theta_0$ against a two-sided alternative.

Solution: We have
$$\begin{aligned}
\ell(\theta) &= n \log \theta + n\theta \log c - (\theta + 1) \sum_{i=1}^n \log X_i, \\
\ell'(\theta) &= \frac{n}{\theta} + n \log c - \sum_{i=1}^n \log X_i, \\
\ell''(\theta) &= -\frac{n}{\theta^2}.
\end{aligned}$$
Hence, $\hat{\theta}_n = 1\big/\overline{\log(X/c)}$ and $nI(\theta) = n/\theta^2$.

- Wald test:
$$W_n = \frac{\sqrt{n}}{\theta_0}\Big(1\big/\overline{\log(X/c)} - \theta_0\Big),$$
and we reject H_0 when $|W_n| \geq z_{\alpha/2}$.
- Rao test:
$$R_n = \sqrt{n}\theta_0\left(\frac{1}{\theta_0} - \overline{\log(X/c)}\right),$$
and we reject H_0 when $|R_n| \geq z_{\alpha/2}$.
- Likelihood ratio test:
$$\Delta_n = n\Big(\theta_0\overline{\log(X/c)} - \log\Big(\theta_0\overline{\log(X/c)}\Big) - 1\Big),$$
and we reject H_0 when $2\Delta_n \geq z_{\alpha/2}^2$.

2. Suppose $\mathbf{X} \sim \text{multinomial}(n, \mathbf{p})$, where $\mathbf{p} \in \mathbb{R}^k$, and define $\boldsymbol{\theta} = (p_1, \ldots, p_{k-1})$. Suppose we wish to test $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$.

(a) Prove that the Wald and score tests are the same as the usual Pearson χ^2 test;

Solution: Pearson χ^2 test statistic is
$$C_n^2 = \sum_{i=1}^k \frac{(X_{n,i} - np_i)^2}{np_i} = \sum_{i=1}^{k-1} \frac{(X_{n,i} - np_i)^2}{np_i} + \frac{\left(\sum_{i=1}^{k-1} X_{n,i} - n\sum_{i=1}^{k-1} p_i\right)^2}{n\left(1 - \sum_{i=1}^{k-1} p_i\right)}. \tag{1}$$
Then we derive Wald and score test statistics.
$$\begin{aligned}
\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{k-1} X_{n,i} \log p_i + \left(n - \sum_{i=1}^{k-1} X_{n,i}\right) \log\left(1 - \sum_{i=1}^{k-1} p_i\right), \\
\nabla \ell(\boldsymbol{\theta})_i &= \frac{X_{n,i}}{p_i} - \frac{X_{n,k}}{1 - \sum_{i=1}^{k-1} p_i}, \text{ for } i \in \{1, 2, \ldots, k-1\}, \\
\nabla^2 \ell(\boldsymbol{\theta})_{ii} &= -\frac{X_{n,i}}{p_i^2} - \frac{X_{n,k}}{\left(1 - \sum_{i=1}^{k-1} p_i\right)^2}, \text{ for } i \in \{1, 2, \ldots, k-1\}, \\
\nabla^2 \ell(\boldsymbol{\theta})_{ij} &= -\frac{X_{n,k}}{\left(1 - \sum_{i=1}^{k-1} p_i\right)^2}, \text{ for } i \neq j \in \{1, 2, \ldots, k-1\}.
\end{aligned}$$

Therefore,

$$\hat{p}_i = \frac{X_{n,i}}{n}, \text{ for } i \in \{1, 2, \ldots, k-1\}$$

and

$$[nI(\boldsymbol{\theta})]_{ii} = \frac{n}{p_i} + \frac{n}{\left(1 - \sum_{i=1}^{k-1} p_i\right)}, \text{ for } i \in \{1, 2, \ldots, k-1\},$$
$$[nI(\boldsymbol{\theta})]_{ij} = \frac{n}{\left(1 - \sum_{i=1}^{k-1} p_i\right)}, \text{ for } i \neq j \in \{1, 2, \ldots, k-1\}.$$

For Wald test, the test statistic is

$$\begin{aligned} W_n &= \sum_{i=1}^{k-1} \left(\frac{X_{n,i}}{n} - p_i\right)^2 \frac{n}{p_i} + \frac{n}{\left(1 - \sum_{i=1}^{k-1} p_i\right)} \left(\sum_{i=1}^{k-1} \left(\frac{X_{n,i}}{n} - p_i\right)\right)^2 \\ &= \sum_{i=1}^{k-1} \frac{(X_{n,i} - np_i)^2}{np_i} + \frac{\left(\sum_{i=1}^{k-1} (X_{n,i} - np_i)\right)^2}{n\left(1 - \sum_{i=1}^{k-1} p_i\right)}. \end{aligned} \tag{2}$$

It can be found that (1) and (2) are equal.
For score test, recall the Sherman-Morrison formula

$$(A + uv^\top)^{-1} = A^{-1} - \frac{A^{-1} uv^\top A^{-1}}{1 + v^\top A^{-1} u}.$$

With this formula, we can derive the expression of $[nI(\boldsymbol{\theta})]^{-1}$,

$$[nI(\boldsymbol{\theta})]^{-1}_{ii} = \frac{p_i - p_i^2}{n}, \text{ for } i \in \{1, 2, \ldots, k-1\},$$
$$[nI(\boldsymbol{\theta})]^{-1}_{ij} = \frac{p_i p_j}{n}, \text{ for } i \neq j \in \{1, 2, \ldots, k-1\}.$$

Therefore, the test statistics is

$$R_n = \sum_{i=1}^{k-1} \left(\frac{X_{n,i}}{p_i} - \frac{X_{n,k}}{1 - \sum_{i=1}^{k-1} p_i}\right)^2 \frac{p_i}{n} + \left(\sum_{i=1}^{k-1} \left(\frac{X_{n,i}}{p_i} - \frac{X_{n,k}}{1 - \sum_{i=1}^{k-1} p_i}\right) p_i\right)^2 \Big/ n \tag{3}$$

It can be calculated that (1) and (3) are equal.

(b) Derive the likelihood ratio statistic.

Solution: The likelihood ratio statistic is

$$\Delta_n = \sum_{i=1}^{k-1} X_{n,i} \log \frac{X_{n,i}}{np_i} + X_{n,k} \log \frac{1 - \sum_{i=1}^{k-1} X_{n,i}/n}{1 - \sum_{i=1}^{k-1} p_i}.$$

3. Show that under sufficient regularity conditions, the mean integrated square error (MISE) of a kernel estimator $\hat{f}_n$ with a bandwidth parameter h satisfies

$$\text{MISE}_f(\hat{f}_n) \sim \frac{1}{nh} \int K^2(y) \mathrm{d}y + \frac{h^4}{4} \int f''(x)^2 \mathrm{d}x \left(\int y^2 K(y) \mathrm{d}y\right)^2.$$

What does this imply for an optimal choice of the bandwidth h?

Solution: We assume that $f(x)$ is twice differentiable, and every integral on the right-hand side is finite. Based on the bias-variance decomposition of MISE, we have

$$\text{MISE}_f(\hat{f}_n) = \int \text{Var}_f\, \hat{f}(x) \mathrm{d}x + \int (\mathbb{E}_f \hat{f}(x) - f(x))^2 \mathrm{d}x.$$

For the first term, by u-substitution and Taylor expansion, we have

$$\begin{aligned} \text{Var}_f\, \hat{f}(x) &= \frac{1}{nh^2} \left[\mathbb{E} K^2\left(\frac{X - x}{h}\right) - \left(\mathbb{E} K\left(\frac{X - x}{h}\right)\right)^2\right] \\ &\sim \frac{1}{nh} \int K^2(y) f(x + hy) \mathrm{d}y - \frac{1}{n}\left(f(x) + \frac{f''(x)}{2} h^2 \int y^2 K(y) \mathrm{d}y\right)^2. \end{aligned}$$

For small $h > 0$, taking integral with respect to x on both sides with Fubini theorem gives

$$\int \mathrm{Var}_f \hat{f}(x)\mathrm{d}x \sim \frac{1}{nh} \int K^2(y)\mathrm{d}y. \tag{4}$$

Then for the second term, with u-substitution and Taylor expansion again, we have

$$\begin{aligned}\mathbb{E}_f \hat{f}(x) - f(x) &= \int K(y)(f(x - hy) - f(x))\mathrm{d}y \\ &\sim \frac{h^2}{2} f''(x)\left(\int y^2 K(y)\mathrm{d}y\right),\end{aligned}$$

so

$$\int (\mathbb{E}_f \hat{f}(x) - f(x))^2 \mathrm{d}x \sim \frac{h^4}{4} \int f''(x)^2 \mathrm{d}x \left(\int y^2 K(y)\mathrm{d}y\right)^2. \tag{5}$$

Combining (4) and (5) completes the proof. Based on this, the optimal choice of the bandwidth h is

$$h = \left(\frac{\int K^2(y)\mathrm{d}y}{n \int f''(x)^2 \mathrm{d}x (\int y^2 K(y)\mathrm{d}y)^2}\right)^{1/5}.$$