Math 281C Homework 8 Solutions

1. Let $X_1, \ldots, X_n$ be i.i.d. random variables from the beta distribution with pdf $\theta x^{\theta-1}\mathbb{1}(0 < x < 1)$, and independent of X_i's, let $Y_1, \ldots, Y_m$ be i.i.d. from the beta distribution with pdf $\mu x^{\mu-1}\mathbb{1}(0 < x < 1)$. For testing $H_0 : \theta = \mu$ versus $H_1 : \theta \neq \mu$, find the forms of the LR test, Wald's test, and Rao's score test.

   Solution: The log-likelihood function (omitting the indicator) is

$$\log \ell(\theta, \mu) = n \log \theta + m \log \mu + (\theta - 1) \sum_{i=1}^{n} \log x_i + (\mu - 1) \sum_{i=1}^{m} \log y_i.$$

   The restricted MLEs are $\widetilde{\theta} = \widetilde{\mu} = -(m+n)/(\sum_{i=1}^n \log x_i + \sum_{i=1}^m \log y_i)$, and the unrestricted MLEs are $\widehat{\theta} = -n/(\sum_{i=1}^n \log x_i)$, $\widehat{\mu} = -m/(\sum_{i=1}^m \log y_i)$.
   <u>LR test</u>:

$$\lambda_{m,n}(\boldsymbol{x}, \boldsymbol{y}) = (\widetilde{\theta}/\widehat{\theta})^n (\widetilde{\mu}/\widehat{\mu})^m \Big(\prod_{i=1}^{n} x_i\Big)^{\widetilde{\theta}-\widehat{\theta}} \Big(\prod_{i=1}^{m} y_i\Big)^{\widetilde{\mu}-\widehat{\mu}},$$

   and we reject H_0 if $\lambda_{m,n}(\boldsymbol{x}, \boldsymbol{y}) < C$.
   <u>Wald's test</u>: $R(\theta, \mu) = \theta - \mu$, and $R'(\theta, \mu) = (1, -1)^\top$. The Fisher information is $I_{m,n}(\theta, \mu) = \operatorname{diag}(n/\theta^2, m/\mu^2)$.

$$W_{m,n}(\boldsymbol{x}, \boldsymbol{y}) = \frac{(\widehat{\theta} - \widehat{\mu})^2}{\widehat{\theta}^2/n + \widehat{\mu}^2/m}.$$

   We reject H_0 if $W_{m,n}(\boldsymbol{x}, \boldsymbol{y}) > C$.
   <u>Rao's score test</u>: $g(\theta) = (\theta, \theta)^\top$. We omit the computational details of score function.

$$R_{m,n}(\boldsymbol{x}, \boldsymbol{y}) = \frac{\widetilde{\theta}^2}{n}\Big(n/\widetilde{\theta} + \sum_{i=1}^{n} \log x_i\Big)^2 + \frac{\widetilde{\theta}^2}{m}\Big(m/\widetilde{\theta} + \sum_{i=1}^{m} \log y_i\Big)^2.$$

   H_0 is rejected if $R_{m,n}(\boldsymbol{x}, \boldsymbol{y}) > C$.

2. Suppose that $X = (X_1, \ldots, X_k)^T$ has the multinomial distribution with a known size n and an unknown probability vector $\mathbf{p} = (p_1, \ldots, p_k)^T$. Consider the problem of testing

$$H_0 : \mathbf{p} = \mathbf{p}_0 \quad \text{versus} \quad H_1 : \mathbf{p} \neq \mathbf{p}_0,$$

   where $\mathbf{p}_0 = (p_{01}, \ldots, p_{0k})^T$ is a known probability vector, that is, $\sum_{j=1}^k p_{0j} = 1$ and $p_{0j} \in (0, 1)$. Find the forms of Wald's test and Rao's score test.

   Solution: Denote $\boldsymbol{\theta} = (p_1, \ldots, p_{k-1})^\top$, it can be calculated that

$$\begin{aligned}
\log \ell(\boldsymbol{\theta}) &= \sum_{i=1}^{k-1} x_i \log p_i + \left(n - \sum_{i=1}^{k-1} x_i\right) \log\left(1 - \sum_{i=1}^{k-1} p_i\right), \\
\nabla \log \ell(\boldsymbol{\theta})_i &= \frac{x_i}{p_i} - \frac{x_k}{1 - \sum_{i=1}^{k-1} p_i}, \text{ for } i \in \{1, 2, \ldots, k-1\}, \\
\nabla^2 \log \ell(\boldsymbol{\theta})_{ii} &= -\frac{x_i}{p_i^2} - \frac{x_k}{\left(1 - \sum_{i=1}^{k-1} p_i\right)^2}, \text{ for } i \in \{1, 2, \ldots, k-1\}, \\
\nabla^2 \log \ell(\boldsymbol{\theta})_{ij} &= -\frac{x_k}{\left(1 - \sum_{i=1}^{k-1} p_i\right)^2}, \text{ for } i \neq j \in \{1, 2, \ldots, k-1\}.
\end{aligned}$$

   Therefore,

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

and

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

For <u>Wald's test</u>, the test statistic is

$$W_n = \sum_{i=1}^{k} \frac{(x_i - np_{0i})^2}{x_i}.$$

For <u>Rao's score test</u>, 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 $[nI(\boldsymbol{\theta})]^{-1}$,

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

Therefore, the test statistics is

$$R_n = \sum_{i=1}^{k} \frac{(x_i - np_{0i})^2}{np_{0i}}.$$

3. Suppose that $X_1, \dots, X_n$ are iid from the Beta$(\mu, 1)$ distribution, and $Y_1, \dots, Y_m$ are iid from the Beta$(\theta, 1)$ distribution. Assume that X_i's and Y_j's are independent.
   (i) Find the LRT for testing $H_0 : \theta = \mu$ versus $H_1 : \theta \neq \mu$.

   Solution: Following Question 1,

$$\lambda_{m,n}(\boldsymbol{x}, \boldsymbol{y}) = (\widetilde{\theta}/\widehat{\theta})^n (\widetilde{\mu}/\widehat{\mu})^m \Big(\prod_{i=1}^{n} x_i\Big)^{\widetilde{\theta} - \widehat{\theta}} \Big(\prod_{i=1}^{m} y_i\Big)^{\widetilde{\mu} - \widehat{\mu}},$$

   and we reject H_0 if $\lambda_{m,n}(\boldsymbol{x}, \boldsymbol{y}) < C$.

   (ii) Show that the test in part (i) can be based on the statistic

$$T = \frac{\sum_{i=1}^{n} \log X_i}{\sum_{i=1}^{n} \log X_i + \sum_{j=1}^{m} \log Y_j}.$$

   Solution: The claimed result can be found by taking logarithm of $\lambda_{m,n}(\boldsymbol{x}, \boldsymbol{y})$.

   (iii) Find the distribution of T when H_0 is true, and show how to get a test of size $\alpha = 0.1$.

   Solution: By a density transformation, $-\log X_i$ and $-\log Y_i$ follow $\exp(\theta)$, so $-\sum_{i=1}^{n} \log X_i \sim$ Gamma(n, θ), $-\sum_{j=1}^{m} \log Y_j \sim$ Gamma(m, θ), and hence, $T \sim$ Beta(n, m). We reject H_0 if $T < C_1$ or $T > C_2$, where C_1, C_2 satisfy

$$C_1^n (1 - C_1)^m = C_2^n (1 - C_2)^m, \qquad \int_{C_1}^{C_2} \frac{\Gamma(m+n)}{\Gamma(m) + \Gamma(n)} x^{n-1} (1-x)^{m-1} \mathrm{d}x = 1 - \alpha.$$

4. The following example comes from genetics. There is a particular characteristic of human blood (the so-called MN blood group) that has three types: M, MN, and N. Under idealized circumstances known as Hardy-Weinberg equilibrium, these three types occur in the population with probabilities $p_1 = \pi_M^2$, $p_2 = 2\pi_M \pi_N$ and $p_3 = \pi_N^2$, respectively, where π_M is the frequency of the M allele in the population and $\pi_N = 1 - \pi_M$ is the frequency of the N allele.
   We observe data $X_1, \dots, X_n$, where X_i has one the three possible values: $(1, 0, 0)^T$, $(0, 1, 0)^T$, or $(0, 0, 1)^T$, depending on whether the ith individual has the M, MN, or N blood type. Denote the total number of individuals of each of the three types by n_1, n_2, and n_3; that is, $n_j = n\bar{X}_j$ for each j.
   If the value of π_M were known, then we already know that the Pearson χ^2 statistic converges in distribution to a chi-square distribution with 2 degrees of freedom. However, in practice we usually don't know π_M. Instead, we estimate it using the maximum likelihood estimator $\widehat{\pi}_M = (2n_1 + n_2)/(2n)$. By the invariance principle of maximum likelihood estimation, this gives $\widehat{p} = (\widehat{\pi}_M^2, 2\widehat{\pi}_M \widehat{\pi}_N, \widehat{\pi}_N^2)^T$ as the maximum likelihood estimator of $p = (p_1, p_2, p_3)^T$.

(a) Define $Z_n = \sqrt{n}(\bar{X} - \widehat{p})$. Use the delta method to derive the asymptotic distribution of $D^{-1/2} Z_n$, where $D = \mathrm{diag}(p_1, p_2, p_3)$.

Solution: Denote $T = (T_1, T_2, T_3) := (n_1/n, n_2/n, n_3/n)$. We know that

$$\sqrt{n}(T - p) \xrightarrow{d} N(0, \Sigma),$$

where $\Sigma_{ii} = p_i(1 - p_i)$ and $\Sigma_{ij} = -p_i p_j$ for $i \neq j$. It can be computed that

$$(T - \widehat{p})^\top = (T_1 T_3 - T_2^2/4, T_2^2/2 - 2T_1 T_3, T_1 T_3 - T_2^2/4).$$

Define $\phi(\boldsymbol{x}) : \mathbb{R}^3 \mapsto \mathbb{R}^3$ as $\phi(x_1, x_2, x_3) := (x_1 x_3 - x_2^2/4, x_2^2/2 - 2x_1 x_3, x_1 x_3 - x_2^2/4)^\top$, we have $\phi(p) = 0$, and applying Delta method gives us

$$\sqrt{n}(T - \widehat{p}) = \sqrt{n}(\phi(T) - \phi(p)) \xrightarrow{d} N(0, \nabla\phi \cdot \Sigma \cdot \nabla\phi^\top),$$

where $\nabla\phi = \boldsymbol{u} \cdot (p_3, -p_2/2, p_1)$ with $\boldsymbol{u} = (1, -2, 1)^\top$. Consequently,

$$D^{-1/2}\sqrt{n}(T - \widehat{p}) \xrightarrow{d} N(0, D^{-1/2}\nabla\phi \cdot \Sigma \cdot \nabla\phi^\top D^{-1/2}),$$

and it can be verified (with burdensome computation) that

$$D^{-1/2}\nabla\phi \cdot \Sigma \cdot \nabla\phi^\top D^{-1/2} = (p_2^2/4) \cdot D^{-1/2}\boldsymbol{u}\boldsymbol{u}^\top D^{-1/2}.$$

(b) Define $\widehat{D}$ to be the diagonal matrix with entries $\widehat{p}_1, \widehat{p}_2, \widehat{p}_3$ along its diagonal. Derive the asymptotic distribution of $\widehat{D}^{-1/2} Z_n$.

Solution: Since

$$\widehat{D}^{-1/2} Z_n = (\widehat{D}^{-1/2} - D^{-1/2}) Z_n + D^{-1/2} Z_n.$$

A random matrix converges in probability if and only if every element converges in probability, so $\widehat{D}^{-1/2} \xrightarrow{p} D^{-1/2}$. Combining Slutsky's theorem and part (a) gives us

$$\widehat{D}^{-1/2} Z_n \xrightarrow{d} N(0, (p_2^2/4) \cdot D^{-1/2}\boldsymbol{u}\boldsymbol{u}^\top D^{-1/2}).$$

(c) Derive the asymptotic distribution of the Pearson chi-square statistic

$$\chi^2 = \sum_{j=1}^{n} \frac{(n_j - n\widehat{p}_j)^2}{n\widehat{p}_j}.$$

Solution: The χ^2 statistic defined above can be written as $\chi^2 = \|\widehat{D}^{-1/2} Z_n\|_2^2$. Intuitively, the asymptotic covariance matrix in part (b) is rank-one, and the only non-zero eigenvalue is 1, so $\chi^2 \xrightarrow{d} \chi_1^2$. To prove this, denote $\boldsymbol{v} := (p_2/2) \cdot D^{-1/2}\boldsymbol{u}$, then the asymptotic covariance of $\widehat{D}^{-1/2} Z_n$ can be written as $\boldsymbol{v}\boldsymbol{v}^\top$. Equivalently, $\widehat{D}^{-1/2} Z_n \xrightarrow{d} \boldsymbol{v} Z$, where $Z \sim N(0, 1)$, and

$$\|\widehat{D}^{-1/2} Z_n\|_2^2 \xrightarrow{d} Z\boldsymbol{v}^\top \boldsymbol{v} Z = Z^2 \sim \chi_1^2.$$