Statistics 581, Problem Set 10 Solutions

Wellner; 12/6/2000

1. Ferguson, ACLST, page 149, problem 2 modified as follows:
(a) Find the LR test statistic of the null hypothesis $H_0 : \mu = c\theta$ for any fixed number $c > 0$, and find the asymptotic distribution of the LR statistic under H_0.
(b) Does the theory of our chapter 4 (or Ferguson's chapter 22) apply directly?
(c) Does the local asymptotic power of your test depend on c?

Solution: (b) First, allow me to slightly re-name the parameters: I will assume that $X_1, \ldots, X_n$ are i.i.d. $\exp(\lambda)$ and $Y_1, \ldots, Y_n$ are i.i.d. $\exp(\mu)$, so that $\theta = (\lambda, \mu)$. Furthermore, we can recast the problem into the context of chapter 4 by considering the pairs of observations (X_i, Y_i), $i = 1, \ldots, n$ as i.i.d. with density

$$p_\theta(x, y) = p_{(\lambda,\mu)}(x, y) = \lambda e^{-\lambda x} 1_{(0,\infty)}(x) \mu e^{-\mu y} 1_{(0,\infty)}(y) \, .$$

Now we are testing $H_0 : \mu = c\lambda$ versus $H_1 : \mu \neq c\lambda$. By a reparametrizaton, we can put this exactly in the setting of Section 4.2: if the original parameter is $\theta = (\lambda, \mu)$, then the new parameters $\gamma = (\gamma_1, \gamma_2)$ where $\gamma_1 \equiv \lambda$, $\gamma_2 \equiv \mu - c\lambda$. Then the null hypothesis H_0 becomes $H_0 : \gamma_2 = 0, \gamma_1 =$ anything.
(a) The MLE $\widehat{\theta}$ of $\theta = (\lambda, \mu)$ under H_1 is $\widehat{\theta} = (\widehat{\lambda}, \widehat{\mu})$ where $\widehat{\lambda} = 1/\overline{X}$ and $\widehat{\mu} = 1/\overline{Y}$. The MLE $\widehat{\theta}^0$ under H_0 is $(\widehat{\lambda}^0, c\widehat{\lambda}^0)$ where

$$\widehat{\lambda}^0 = 2/(\overline{X} + c\overline{Y}) \, .$$

Now

$$l_n(\theta) = l_n(\lambda, \mu) = \sum_{i=1}^n \{\log \lambda - \lambda X_i + \log \mu - \mu Y_i\} = n \log \lambda + n \log \mu - n\overline{X}\lambda - n\overline{Y}\mu \, .$$

Thus the LR statistic for testing H_0 versus H_1 is given by

$$\begin{aligned} 2(l_n(\widehat{\theta}) - l_n(\widehat{\theta}^0)) &= 2n \left\{ 2 \log \left(\frac{\overline{X} + c\overline{Y}}{2} \right) - \log(\overline{X}) - \log(c\overline{Y}) \right\} \\ &\to_d \chi_1^2 \end{aligned}$$

under H_0.
(c) To compute the local asymptotic power of the LR test, we can reparametrize the problem by $\gamma \equiv (\gamma_1, \gamma_2)$ where $\gamma_1 \equiv \lambda$, $\gamma_2 \equiv \mu - c\lambda$. Then the null hypothesis H_0 becomes $H_0 : \gamma_2 = 0, \gamma_1 =$ anything. Then the problem fits in the context of Theorem 4.2.7: under P_{γ_n} with $\gamma_n = \gamma_0 + tn^{-1/2}$ for $\gamma_0 = (\gamma_{10}, 0)$ in the null hypothesis, we have

$$2 \log \lambda_n \to_d \chi_1^2(\delta)$$

where the non-centrality parameter δ is given by $t_2^2 I_{22\cdot 1}(\gamma_0)$, and it remains only to compute $I_{22\cdot 1}$. By straightforward computation the information matrix for γ is given by

$$I(\gamma) = \begin{pmatrix} \frac{1}{\gamma_1^2} + \frac{c^2}{(c\gamma_1+\gamma_2)^2} & \frac{c}{(c\gamma_1+\gamma_2)^2} \\ \frac{c}{(c\gamma_1+\gamma_2)^2} & \frac{1}{(c\gamma_1+\gamma_2)^2} \end{pmatrix} .$$

Thus, under the null hypothesis $H_0 : \gamma_2 = 0$ we find that

$$I_{22\cdot 1}(\gamma_0) = I_{22}(\gamma_0) - I_{21}(\gamma_0) I_{11}^{-1}(\gamma_0) I_{12}(\gamma_0) = \frac{1/2}{c^2\gamma_1^2}$$

which does depend on c: the noncentrality power of the limiting distribution *decreases* as c^{-2} as c increases.

2. Ferguson, ACLST, page 150, problem 3. Does the theory in our chapter 4 (or Ferguson's chapter 22) apply directly? For $i = 1, \ldots, k$, let $X_{i,1}, \ldots, X_{i,n}$ be independent samples from Poisson distributions, Poisson(θ_i) respectively. Find the likelihood ratio test and its asymptotic distribution, for testing $H_0 : \theta_1 = \ldots = \theta_k$.

   Solution: Let $\underline{Y}_j \equiv (X_{1,j}, \ldots, X_{k,j})$ for $j = 1, \ldots, n$. These random vectors are i.i.d. with common density

$$p_\theta(\underline{y}) = P_\theta(\underline{Y} = \underline{y}) = \prod_{i=1}^{k} e^{-\theta_i} \frac{\theta_i^{y_i}}{y_i!} \, .$$

   Thus our theory from Chapter 4 will apply (there *is* some difficulty if the sample sizes n_i, $i = 1, \ldots, k$, differ). The MLE of $\theta = (\theta_1, \ldots, \theta_k)$ under the general hypothesis is $\underline{\widehat{\theta}} = (\widehat{\theta}_1, \ldots, \widehat{\theta}_k)$ where

$$\widehat{\theta}_i = \frac{1}{n} \sum_{j=1}^{n} X_{i,j} \equiv \overline{X}_{i\cdot} \, .$$

   The MLE of $\theta = (\theta_1, \ldots, \theta_1) \in H_0$ is given by $\widehat{\theta}^0 = (\widehat{\theta}_1^0, \ldots, \widehat{\theta}_1^0)$ where $\widehat{\theta}_1^0 = \overline{X}_{\cdot,\cdot} \equiv (nk)^{-1} \sum_{j=1}^{n} \sum_{i=1}^{k} X_{i,j}$. The log-likelihood ratio statistic is:

$$2\{l_n(\widehat{\theta}) - l_n(\widehat{\theta}^0)\} = 2n \left\{ \sum_{i=1}^{k} \overline{X}_{i\cdot} \log \overline{X}_{i\cdot} - k\overline{X}_{\cdot\cdot} \log \overline{X}_{\cdot\cdot} \right\} \to_d \chi^2_{k-1}$$

   under H_0.

3. Suppose that $(Y|Z) \sim$ Poisson$(\lambda e^{\gamma Z})$, and $Z \sim Bernoulli(\eta)$, and $\theta = (\lambda, \gamma, \eta)$. Let $X = (Y, Z)$, and suppose that we observe $X_1, \ldots, X_n$ i.i.d. as X.
   Consider testing the hypothesis $H : \gamma = 0$ versus $K : \gamma \neq 0$. (Note that the null hypothesis is *not simple*, but *composite*; the values of λ and η are not specified by the hypothesis H.)
   (a) Propose three different test statistics for testing H versus K, and briefly discuss how you would compute them.
   (b) Do our results in Chapter 4 apply to the (asymptotic) distribution under H of the test statistics you proposed in (a)?
   (c) Consider local alternatives of the form $\gamma_n = tn^{-1/2}$ for $t \in R$ fixed. Give an expression for the local asymptotic power of the tests you proposed in (a) for these alternatives.
   (d) Suppose that $\gamma_1 \neq 0$ is the "true" value of the parameter γ. Show that the test statistics you proposed in (a), when appropriately normalized, converge in probability to positive constants, and identify these constants as explicitly as possible.

Solution: (a) The LR, Wald, and Rao statistics are given by

$$2\log\lambda_n = 2\{l_n(\widehat{\theta}) - l_n(\widehat{\theta}^0)\}\,,$$

$$W_n = [n^{1/2}(\widehat{\gamma}-0)]\hat{I}_{22\cdot 1}[n^{1/2}(\widehat{\gamma}-0)] = n\widehat{\gamma}^2\hat{I}_{22\cdot 1}\,,$$

and

$$R_n = [Z_n(\widehat{\theta}^0)]^T\hat{I}(\widehat{\theta}^0)[Z_n(\widehat{\theta}^0)]\,,$$

where $\widehat{\theta} = (\widehat{\lambda}, \widehat{\gamma}, \widehat{\eta})$ are the MLE's as discussed in problem set #9, and $\widehat{\theta}^0 = (\widehat{\lambda}^0, 0, \widehat{\eta}^0)$ are the MLE's of θ under H_0.

(b) Yes – the results for a composite hypothesis in section 4.2 do apply; all three converge in distribution to χ_1^2 under H_0.

(c) Under local alternatives of the form $\gamma_n = tn^{-1/2}$ the three statistics in (a) converge in distribution to $\chi_1^2(\delta)$ where $\delta = t^2 I_{22\cdot 1}$ where, as in the solution to problem 1(d), Problem Set #8,

$$\begin{aligned} I_{22\cdot 1} &= \lambda\left\{E\{Z^2e^{\gamma Z}\} - \frac{E\{Ze^{\gamma Z}\}^2}{E\{e^{\gamma Z}\}}\right\} \\ &= \lambda E\{e^{\gamma Z}\}Var_G(Z)\,, \end{aligned}$$

where Var_{G_γ} indicates that the variance is computed under G_γ, the $\gamma-$tilted distribution corresponding to G:

$$G_\gamma(A) = \frac{\int_A e^{\gamma z}dG(z)}{\int e^{\gamma z}dG(z)}\,.$$

(d) When $\theta = (\lambda, \gamma, \eta) \notin \Theta_0$, and A0-A4 hold at θ, then

$$n^{-1}W_n \;= \widehat{\gamma}^2\hat{I}_{22\cdot 1} \to_p \gamma^2 I_{22\cdot 1}(\theta)\,.$$

Let $\theta_*^0 \in \Theta_0$ satisfy $\inf_{\theta'\in\Theta_0} K(P_\theta, P_{\theta'}) = K(P_\theta, P_{\theta_*^0})$. Then we can write

$$\begin{aligned} 2n^{-1}\log\lambda_n &= 2n^{-1}\{l_n(\widehat{\theta}_n) - l_n(\widehat{\theta}_n^0)\} \\ &= 2n^{-1}\left\{\{l_n(\widehat{\theta}_n) - l_n(\theta)\} + \;\{l_n(\theta) - l_n(\theta_*^0)\} + \;\{l_n(\theta_*^0) - l_n(\widehat{\theta}_n^0)\}\right\} \\ &\equiv A_n + B_n + C_n\,. \end{aligned}$$

Now under P_θ we have

$$nA_n \to_d \chi_3^2$$

so that $A_n \to_p 0$; furthermore,

$$B_n = 2n^{-1}\sum_{i=1}^n \log\frac{p_\theta}{p_{\theta_*^0}}(X_i) \to_p 2K(P_\theta, P_{\theta_*^0}) \;=\; 2\inf_{\theta'\in\Theta_0} K(P_\theta, P_{\theta'})\,.$$

Finally, we claim that $C_n \to_p 0$. Note that the score equations for $\theta^0 = (\lambda^0, 0, \eta^0)$ become

$$\begin{aligned} 0 &= \frac{1}{\widehat{\lambda}^0}\sum_{i=1}^n (Y_i - \widehat{\lambda}^0) \\ 0 &= (\widehat{\eta}^0(1-\widehat{\eta}^0))^{-1}\sum_{i=1}^n (Z_i - \widehat{\eta}^0)\,, \end{aligned}$$

and hence it follows that $\widehat{\lambda}^0 = \overline{Y}$, $\widehat{\eta}^0 = \overline{Z}$. Under P_θ with $\theta = (\lambda, \gamma, \eta) \notin \Theta_0$, we have

$$\widehat{\theta}^0 = (\widehat{\lambda}^0, 0, \widehat{\eta}^0) \to_p (E_\theta(Y), 0, E_\theta(Z)) = (\lambda E(e^{\gamma Z}), 0, \eta) \equiv \theta_*^0 \in \Theta_0 \, .$$

We claim that this point θ_*^0 is the point in Θ_0 satisfying $\inf_{\theta' \in \Theta_0} K(P_\theta, P_{\theta'}) = K(P_\theta, P_{\theta_*^0})$. To see this, note that from part (b) of problem 2, Problem set #9 we have, using $\gamma' = 0$ for $\theta' \in \Theta_0$,

$$\begin{aligned} K(P_\theta, P_{\theta'}) &= E_\theta \left(\lambda e^{\gamma Z} \log \left(\frac{\lambda e^{\gamma Z}}{\lambda' e^{\gamma' Z}} \right) \right) + \eta(\lambda' e^{\gamma'} - \lambda e^{\gamma}) + (1 - \eta)(\lambda' - \lambda) \\ &\qquad + \eta \log \frac{\eta}{\eta'} + (1 - \eta) \log \frac{1 - \eta}{1 - \eta'} \\ &= E_\theta \left(\lambda e^{\gamma Z} \log \left(\frac{\lambda e^{\gamma Z}}{\lambda'} \right) \right) + \eta(\lambda' - \lambda e^{\gamma}) + (1 - \eta)(\lambda' - \lambda) \\ &\qquad + \eta \log \frac{\eta}{\eta'} + (1 - \eta) \log \frac{1 - \eta}{1 - \eta'} \, . \end{aligned}$$

To minimize this with respect to λ' and η' we compute

$$\frac{\partial}{\partial \lambda'} K(P_\theta, P_{\theta'}) = E_\theta \left(\lambda e^{\gamma Z} \left(\frac{-1}{\lambda'} \right) \right) + 1 = 0 \, ,$$

and

$$\frac{\partial}{\partial \eta'} K(P_\theta, P_{\theta'}) = \left(\frac{\eta}{\eta'} - \frac{1 - \eta}{1 - \eta'} \right) = 0 \, ,$$

and hence we find that $\lambda_*^0 = \lambda E_\theta(e^{\gamma Z})$ and $\eta_0^* = \eta$. It is not hard to check that these are indeed the minimizing values. Finally we complete the argument by showing that $C_n \to_p 0$:

$$\begin{aligned} C_n &= 2n^{-1}\{l_n(\theta_*^0) - l_n(\widehat{\theta}_n^0)\} \\ &= 2n^{-1} \dot{l}_{n,\theta}(\theta_n^{0*})(\theta_*^0 - \widehat{\theta}_n^0) \\ &\to_p E_\theta\{\dot{l}_\theta(\theta_*^0)\} \cdot 0 = 0 \end{aligned}$$

since

$$\widehat{\theta}_n^0 = \text{argmax}_{\theta' \in \Theta_0} n^{-1} l_n(\theta') \to_p \text{argmax}_{\theta' \in \Theta_0} l(\theta') = \text{argmin}_{\theta' \in \Theta_0} K(P_\theta, P_{\theta'}) = \theta_*^0$$

where

$$l(\theta') \equiv E_\theta \log p_{\theta'}(X_1) = E_\theta \log p_\theta(X_1) - K(P_\theta, P_{\theta'}) \, .$$

Hence we conclude that

$$2n^{-1} \log \lambda_n \to_p 2K(P_\theta, P_{\theta_*^0}) \, .$$

where $\theta_*^0 \in \Theta_0$ satisfies $\inf_{\theta' \in \Theta_0} K(P_\theta, P_{\theta'}) = K(P_\theta, P_{\theta_*^0})$.

Finally, we consider R_n: this is relatively easy based on our above proof that $\widehat{\theta}_n^0 \to_p \theta_*^0$:

$$\begin{aligned} n^{-1} R_n &= n^{-1/2} Z_n(\widehat{\theta}_n^0)^T I(\widehat{\theta}_n^0)^{-1} n^{-1/2} Z_n(\widehat{\theta}_n^0) \\ &\to_p E_\theta\{\dot{l}_\theta(\theta_*^0)\}^T I(\theta_*^0)^{-1} E_\theta\{\dot{l}_\theta(\theta_*^0)\}^T \, . \end{aligned}$$

assuming that $E_\theta|\dot{l}_i(\theta_*^0)| < \infty$, which is easily checked in this case.