Convex Polyhedra II: Testing
Statistics 710
October 12, 2006

The Testing Problems. Again suppose that $W = I_n$ and consider a polyhedral cone in $\mathbb{R}^n$,

$$\Omega = \{\theta \in \Re^n : \langle \gamma_i, \theta \rangle \geq 0,\ i = 1, \cdots, m\}, \tag{1}$$

where $\gamma_1, \cdots, \gamma_m \in \Re^n$ are linearly independent; let $L = \text{span}\{\gamma_1, \cdots, \gamma_m\}$; and suppose that $y \sim \text{Normal}[\theta, \sigma^2 I_n]$. The following three hypotheses are considered: $H_0 : \theta \in L^\perp$, $H_1 : \theta \in \Omega$, and $H_2 : \theta \in \mathbb{R}^n$. For example, in monotone regression, H_0 is the hypothesis that the regression function is constant; in convex regression, it is the hypothesis that the regression function is linear.

First consider H_0 vs. $H_1 - H_0$. The log-likelihood function is

$$\ell(\theta, \sigma^2 | y) = -\frac{1}{2\sigma^2}\|y - \theta\|^2 - \frac{1}{2} n \log(\sigma^2), \tag{2}$$

and the least squares estimators of θ are the maximum likelihood estimators. So, the maximum likelihood estimator under H_0 is $\hat{\theta}^o = \Pi_{L^\perp} y$ and the unconditional maximum likelihood estimator is $\hat{\theta} = \Pi_\Omega y$. If σ^2 is known, then the log-likelhood ratio statistics is

$$\Lambda_{01} = 2\left[\ell(\hat{\theta}, \sigma^2) - \ell(\hat{\theta}^o, \sigma^2)\right] = \frac{1}{\sigma^2}\left[\|y - \hat{\theta}^o\|^2 - \|y - \hat{\theta}\|^2\right].$$

Here $y - \hat{\theta}^o = y - \hat{\theta} + \hat{\theta} - \hat{\theta}^o$, and $\|y - \hat{\theta}^o\|^2 = \|y - \hat{\theta}\|^2 + 2\langle y - \hat{\theta}, \hat{\theta} - \hat{\theta}^o \rangle + \|\hat{\theta} - \hat{\theta}^o\|^2 = \|y - \hat{\theta}\|^2 + \|\hat{\theta} - \hat{\theta}^o\|^2$ and, therefore,

$$\Lambda_{01} = \frac{1}{\sigma^2}\|\hat{\theta} - \hat{\theta}^o\|^2.$$

If σ^2 is unknown, then the maximum likelihood estimators are

$$\hat{\sigma}^2 = \frac{\|y - \hat{\theta}\|^2}{n} \qquad \text{and} \qquad \hat{\sigma}_0^2 = \frac{\|y - \hat{\theta}^o\|^2}{n},$$

and the likelihood ratio statistics is

$$\Lambda_{01} = 2\left[\ell(\hat\theta, \hat\sigma^2) - \ell(\hat\theta^o, \hat\sigma_0^2)\right] = n \log\left[\frac{\|y-\hat\theta^o\|^2}{\|y-\hat\theta\|^2}\right] = \log\left[\frac{\|\hat\theta^o-\hat\theta\|^2 + \|y-\hat\theta\|^2}{\|y-\hat\theta\|^2}\right].$$

Of course, an equivalent test is to reject if

$$\frac{\|\hat\theta^o-\hat\theta\|^2}{\|\hat\theta-\hat\theta^o\|^2 + \|y-\hat\theta\|^2}$$

is large.

Next, consider testing H_1 vs H_2, when σ^2 is known. For H_2, the maximum likelihood estimator is y, and

$$\Lambda_{12} = 2\left[\ell(y, \sigma^2) - \ell(\hat\theta, \sigma^2)\right] = \frac{1}{\sigma^2}\|y-\hat\theta\|^2.$$

If σ^2 unknown, then an independent estimate is required,

Least Favorable Configurations. Since both null hypotheses are composite, the dependence of the test statistics on parameters, under the hypotheses must be assessed. For H_0 vs. H_1 this is simple. *The distributions of* $\|\hat\theta^o-\hat\theta\|^2$ *and* $\|y-\hat\theta\|^2$ *are the same for all* $\theta \in L^\perp$. This is a simple consequence of the following: if $z \in \mathbb{R}^n$ and $\theta \in L^\perp$, then

$$\hat\theta(z+\theta) = \hat\theta(z) + \theta \qquad \text{and} \qquad \hat\theta^o(z+\theta) = \hat\theta^o(z) + \theta. \tag{3}$$

To establish the first of these assertions, it suffices to show that $\hat\theta(z)+\theta$ satisfies the necessary and sufficient conditions for $\hat\theta(z+\theta)$. Clearly, $\hat\theta(z)+\theta \in \Omega$ and

$$\langle z+\theta-[\hat\theta(z)+\theta], \xi\rangle = \langle z-\hat\theta(z), \xi\rangle \le 0$$

for all $\xi \in \Omega$. Also,

$$\langle z+\theta-[\hat\theta(z)+\theta], \hat\theta(z)+\theta\rangle = \langle z-\hat\theta(z), \hat\theta(z)+\theta\rangle = 0,$$

since $\hat\theta(z) \pm \theta \in \Omega$. The second assertion in (3) may be established similarly (and more easily). To complete the argument, observe that if $y \sim \text{Normal}(\theta, I_n)$, where $\theta \in L^\perp$, then y has the same distribution as $z+\theta$, where $z \sim \Phi^n$. It follows that

$$[\|\hat\theta(y)-\hat\theta^o(y)\|^2, \|y-\hat\theta(y)\|^2] =^{\text{d}} [\|\hat\theta(z)-\hat\theta^o(z)\|^2, \|z-\hat\theta(z)\|^2].$$

The situation is slightly more complicated for testing H_1 vs H_2, since the distribution of $\|y-\hat\theta(y)\|^2$ does depend on $\theta \in \Omega$, but a bound can be derived. If $y = z+\theta$, where $z \in \mathbb{R}^n$ and $\theta \in \Omega$, then $\hat\theta(z)+\theta \in \Omega$, so that

$$\|y-\hat\theta(y)\|^2 = \inf_{\xi\in\Omega}\|y-\xi\|^2 \le \|z+\theta-[\hat\theta(z)+\theta]\|^2 = \|z-\hat\theta(z)\|^2.$$

So,

$$\max_{\theta\in\Omega} P_\theta[\|y-\hat{\theta}(y)\|^2 > u] \le P[\|z-\hat{\theta}(z)\|^2 > u] = P_0[\|y-\hat{\theta}(y)\|^2 > u].$$

The Null Distribution. The main result is that if $\theta \in L$, then

$$P_\theta[\frac{1}{\sigma^2}\|\hat{\theta}-\hat{\theta}^o\|^2 \le u, \frac{1}{\sigma^2}\|y-\hat{\theta}\|^2 \le v] = \sum_{k=m}^{n} P[\chi^2_{k-m} \le u]P[\chi^2_{n-k} \le v]q(n,k), \qquad (*)$$

where

$$q(n,k) = P_0[D = k].$$

Two preliminary results are need to establish this. First, recall the relation $\hat{\theta} = \Pi_{L_{\hat{J}}}y + \Pi_{L^\perp}y$, where $\hat{J} = \{j \le m : \langle\gamma_j, \hat{\theta}\rangle > 0\}$. Recall too the definitions of Γ_J and Δ_J and observe that $\Gamma_J'\Pi_{L_J} = (\Delta_J'\Delta_J)^{-1}\Delta_J'$ and $\Delta_j'\Pi_{K_J} = (\Gamma_J'\Gamma_J)^{-1}\Gamma_J$. It follows easily that

$$\{y : hatJ(y) = J\} = \{y \in I\!R^n : \Gamma_J'\Pi_{L_J}y > 0 \text{ and } \Delta_{J^c}'\Pi_{K_{Jc}^\perp}y \le 0\}.$$

Next, recall that if $z \sim \Phi^n$, then $\|z\|$ and $z/\|z\|$ are independent. In fact, if $Q \ne 0$ is any projection matrix, then $\|Qz\|$ and $Qz/\|QZ\|$ are independent. To see this recall that the eigen values of a projection matrix are either 0 or 1, so that Q may be written as $Q = C\text{diag}[I_k, 0]C'$, where $1 \le k \le n$ and C is orthogonal. Then $Cz \sim \text{Normal}[0, \text{diag}(I_k, 0)]$, so that $Cz = [w', 0, \cdots, 0]'$, where $w \sim \Phi^k$. The independence of $\|z\|$ and $z/\|z\|$ now follows easily from that of $\|w\|$ and $w/\|w\|$.

For the proof of (*), we may suppose that $\theta = 0$ and $\sigma = 1$. Then

$$P_0\left[\|\hat{\theta}-\hat{\theta}^o\|^2 \le u, \|y-\hat{\theta}\|^2 \le v\right] = \sum_J P\left[\hat{J}(y) = J,\ \|\Pi_{L_J}y\|^2 \le u, \|\Pi_{K_J^\perp}y\|^2 \le v\right]$$

Here $\Pi_{L_J}y$ and $\Pi_{K_{Jc}^\perp}y$ are independent. So,

$$\begin{aligned} P[\hat{J}(y) = J,\ &\|\Pi_{L_J}y\|^2 \le u, \|\Pi_{K_J^\perp}y\|^2 \le v] \\ &= P[\Gamma_J'\Pi_{L_J}y > 0,\ \Delta_{J^c}\Pi_{K_{Jc}^\perp} \le 0,\ \|\Pi_{L_J}y\|^2 \le u, \|\Pi_{K_J^\perp}y\|^2 \le v] \\ &= P[\Gamma_J'\Pi_{L_J}y > 0,\ \|\Pi_{L_J}y\|^2 \le u] \times P[\Delta_{J^c}\Pi_{K_{Jc}^\perp} \le 0, \|\Pi_{K_J^\perp}y\|^2 \le v] \end{aligned}$$

Next, using the independence of norms and angles

$$P[\Gamma_J'\Pi_{L_J}y > 0,\ \|\Pi_{L_J}y\|^2 \le u] = P[\Gamma_J'\Pi_{L_J}y > 0]P[\|\Pi_{L_J}y\|^2 \le u]$$

and

$$P[\Delta_{J^c}\Pi_{K_{Jc}^\perp} \le 0, \|\Pi_{K_J^\perp}y\|^2 \le v].$$

So, letting $k = \#J$

$$\begin{aligned} P[\hat{J}(y) = J,\ \|\Pi_{L_J} y\|^2 \le u, \|\Pi_{K_J^\perp} y\|^2 \le v] \\ &= P[\Gamma'_J \Pi_{L_J} y > 0] \times P[\|\Pi_{L_J} y\|^2 \le u] \times P[\Delta_{J^c} \Pi_{K_{J^c}^\perp} \le 0] \times P[\|\Pi_{K_J^\perp} y\|^2 \le v] \\ &= P[\chi^2_{k-m} \le u] P[\chi^2_{n-k} \le v] P[\hat{J} = J] \end{aligned}$$

in which the independence of $\Pi_{L_J} y$ and $\Pi_{K_J^\perp}$ has been used again. Relation (*) then follows by writing

$$\sum_J = \sum_{k=0}^{n-m} \sum_{\#J=k} .$$

So, for the case of known σ^2,

$$P_\theta\left[\Lambda_{01} > c\right] = P_0\left[\frac{1}{\sigma^2}\|\hat{\theta} - \hat{\theta}^o\|^2 > c\right] = \sum_{k=m}^{n} P[\chi^2_{k-m} > c] q(n,k)$$

for all $\theta \in L^\perp$, and this may set equal to any given α, by appropriate choice of c. For unknown σ^2, recall that if U and V are independent chi-squared variables with r and s degrees of freedom, then

$$\frac{U}{U+V} \sim \beta(\frac{r}{2}, \frac{s}{2}).$$

So,

$$P_\theta\left[\frac{\|\hat{\theta}^o - \hat{\theta}\|^2}{\|\hat{\theta}^o - \hat{\theta}\|^2 + \|y - \hat{\theta}\|^2} > c\right] = \sum_{k=m}^{n} P\left[\beta(\frac{k-m}{2}, \frac{n-k}{2}) > c\right] q(n,k)$$

for all $\theta \in L^\perp$.

Remark. This material is adapted from [1].

References

[1] Robertson, Tim, Farrell Wright, and Richard Dykstra (1988). *Order Restricted Inference*. Wiley.