Econometric Theory, **17**, 2001, 1157–1160. Printed in the United States of America.

PROBLEMS AND SOLUTIONS

PROBLEMS

01.6.1. *Minimax Median*, proposed by Stephen Portnoy and G.W. Bassett. Consider a random sample of size n, and write the data as an $r = r_n$ by $c = c_n$ matrix, $\{X_{ij} : i = 1,\ldots,r_n; j = 1,\ldots,c_n\}$. Clearly,

$$n = r_n c_n. \tag{1}$$

To specify notation, $\{X_{ij}\}$ are i.i.d. with c.d.f. $F(x)$ and (continuous) density $f(x)$. Let β denote the median ($F(\beta) = 0.5$). Define an estimator that may be called the "minmax-median" by

$$\hat{\beta} = \min_j \left\{ \max_i \{X_{ij}\} \right\}. \tag{2}$$

(a) Calculate the c.d.f. of $\hat{\beta}$, and give a condition on r_n and c_n for median-unbiasedness (that is, so that β is also the median for the distribution of $\hat{\beta}$). Note that this condition, together with (1), determines r_n and c_n. Show that

$$r_n = \log(n)(1/\log(2) + o(1)) = \log_2(n)(1 + o(1)). \tag{3}$$

(b) Show that $(\hat{\beta} - \beta)/r_n$ converges in distribution, and find the limiting distribution function.

Remarks: (1) Note that the minmax-median has superior computational properties: it can be computed by only $n + \log(n)$ comparisons, which is much faster than sorting the sample and is even faster than the $2n$ comparisons required by the Floyd–Rivest $\mathcal{O}(n)$-algorithm (see Floyd and Rivest, 1975). Furthermore, it requires only two storage locations (one for the running column mins and one for the running maxes of the mins). However, the result above shows that $\hat{\beta}$ converges only at rate $\log(n)$; and so computational advantages in speed and storage come at a considerable cost in efficiency.

(2) The second author was led to consider the minmax-median for its computational storage and speed advantages. The idea was sparked by Pearl (1981) who introduced median estimators involving alternating mins and maxes in a tree-structured manner somewhat akin to the "remedian" of Rousseeuw and Bassett (1990).

REFERENCES

Floyd, R.W. & R.L. Rivest (1975) Expected time bounds for selection. *Communications of the ACM* 18, 165–173.

Pearl, J. (1981) A space-efficient on-line method of computing quantile estimates. *Journal of Algorithms* 2, 164–177.

© 2001 Cambridge University Press 0266-4666/01 $9.50

https://doi.org/10.1017/S0266466601006065 Published online by Cambridge University Press

Rousseeuw, P. & G. Bassett (1990) The remedian: A robust averaging method for large data sets, *Journal of the American Statistical Association* 85, 97–104.

01.6.2. *Identification of Parameters in Two Competing Risks Models*, proposed by S.K. Sapra. Consider the following *dependent* competing risks models in Problem 00.4.1 for which $T = \min(T_1, T_2)$, and $\delta = 1$ if $T = T_1$, and 0 if $T = T_2$. T_1 and T_2 are latent durations of risks 1 and 2, respectively. Suppose that (T_1, T_2) have the joint survivor functions

$$\textit{Model 1}: S(t_1, t_2) = \exp(-(\lambda_1 t_1 + \lambda_2 t_2)^{\alpha}), \qquad \alpha > 0, \lambda_1, \lambda_2 > 0, \tag{1}$$

$$\textit{Model 2}: S(t_1, t_2) = 1/(\lambda_1 t_1 + \lambda_2 t_2 + 1)^{\alpha+1}, \qquad \alpha > -1, \lambda_1, \lambda_2 > 0. \tag{2}$$

Write down the likelihood function for each model if (a) both T and δ are observed, and (b) only T is observed. Examine the identifiability of parameters for each model in cases (a) and (b).

SOLUTIONS

00.6.1. *First and Second Moments of Newey and West's HAC Covariance Matrix Estimator under Normality*—Solution[1] proposed by Heinz Neudecker.

(1) Basic is $\hat{\varepsilon} = M\varepsilon$, with $M = I_n - X(X'X)^{-1}X'$. For its ith component we have $\hat{\varepsilon}_i = m_i'\varepsilon = \varepsilon' m_i$, where m_i is the ith column of M. This yields $\hat{\omega}_{ij} = \kappa_{i-j} m_i' \varepsilon\varepsilon' m_{i-j} = \kappa_{i-j}\varepsilon' m_i m_{i-j}'\varepsilon = \kappa_{i-j}\varepsilon' m_{i-j} m_i'\varepsilon = \kappa_{i-j}\varepsilon' M_{ij}\varepsilon$, with $M_{ij} = \frac{1}{2}(m_i m_{i-j}' + m_{i-j} m_i')$.

(2) For *any* distribution $\varepsilon \sim (0, \Omega)$ we get the expected value

$$E(\hat{\omega}_{ij}) = \kappa_{i-j} m_i' E(\varepsilon\varepsilon') m_{i-j} = \kappa_{i-j} m_i' \Omega m_{i-j} = \kappa_{i-j} \operatorname{tr} M_{ij}\Omega.$$

(3) For the *normal* distribution $\varepsilon \sim N(0, \Omega)$ we find the covariance

$$C(\hat{\omega}_{ij}, \hat{\omega}_{kl}) = \kappa_{i-j}\kappa_{k-l} C(\varepsilon' M_{ij}\varepsilon, \varepsilon' M_{kl}\varepsilon) = 2\kappa_{i-j}\kappa_{k-l} \operatorname{tr} M_{ij}\Omega M_{kl}\Omega$$

$$= \kappa_{i-j}\kappa_{k-l}(m_i'\Omega m_{k-l} m_k'\Omega m_{i-j} + m_i'\Omega m_k m_{i-j}'\Omega m_{k-l})$$

by Corollary 4.1(i) in Magnus and Neudecker (1979).

(4) For *any* distribution $\varepsilon \sim (0, I_n)$ with independent ε_i and ε_j and $E(\varepsilon_i^4) = \varphi$ we have

$$E(\varepsilon\varepsilon' \otimes \varepsilon\varepsilon') = I_{n^2} + K_{nn} + (vec I_n)(vec I_n)' + (\varphi - 3)\tilde{K}_{nn},$$

where K_{nn} is a commutation matrix, $\hat{K}_{nn} = \sum_{t=1}^{n}(E_{tt} \otimes E_{tt})$. The square matrix E_{tt} has zero elements everywhere except in the ttth position which is unity. For $E(\varepsilon\varepsilon' \otimes \varepsilon\varepsilon')$ under normality (where $\varphi = 3$) see Theorem 4.1(i) in Magnus and Neudecker (1979). We then find

https://doi.org/10.1017/S0266466601006065 Published online by Cambridge University Press

$$
\begin{aligned}
&E(\varepsilon' M_{ij}\varepsilon\varepsilon' M_{kl}\varepsilon)\\
&\quad = E(\varepsilon\otimes\varepsilon)'(M_{ij}\otimes M_{kl})(\varepsilon\otimes\varepsilon) = \operatorname{tr}(M_{ij}\otimes M_{kl})E(\varepsilon\varepsilon'\otimes\varepsilon\varepsilon')\\
&\quad = \operatorname{tr}(M_{ij}\otimes M_{kl}) + \operatorname{tr}(M_{ij}\otimes M_{kl})K_{nn}\\
&\qquad + \operatorname{tr}(M_{ij}\otimes M_{kl})(\operatorname{vec} I_n)(\operatorname{vec} I_n)' + (\varphi-3)\operatorname{tr}(M_{ij}\otimes M_{kl})\tilde{K}_{nn}\\
&\quad = (\operatorname{tr} M_{ij})\operatorname{tr} M_{kl} + 2\operatorname{tr} M_{ij}M_{kl} + (\varphi-3)\operatorname{tr}(M_{ij}\otimes M_{kl})\tilde{K}_{nn},
\end{aligned}
$$

and further

$$
\begin{aligned}
C(\hat{\omega}_{ij},\hat{\omega}_{kl}) &= \kappa_{i-j}\kappa_{k-l}C(\varepsilon' M_{ij}\varepsilon, \varepsilon' M_{kl}\varepsilon)\\
&= 2\kappa_{i-j}\kappa_{k-l}\operatorname{tr} M_{ij}M_{kl} + (\varphi-3)\kappa_{i-j}\kappa_{k-l}\operatorname{tr}(M_{ij}\otimes M_{kl})\tilde{K}_{nn},
\end{aligned}
$$

(5) For *any* distribution $\varepsilon \sim (0,\Omega)$ derived as $\Omega^{1/2}\tilde{\varepsilon}$, where $\tilde{\varepsilon}$ indicates ε as in 4, we find

$$
\begin{aligned}
C(\hat{\omega}_{ij},\hat{\omega}_{kl}) &= \kappa_{i-j}\kappa_{k-l}C(\varepsilon' M_{ij}\varepsilon, \varepsilon' M_{kl}\varepsilon)\\
&= 2\kappa_{i-j}\kappa_{k-l}\operatorname{tr} M_{ij}\Omega M_{kl}\Omega\\
&\quad + (\varphi-3)\kappa_{i-j}\kappa_{k-l}\operatorname{tr}(\Omega^{1/2}M_{ij}\Omega^{1/2}\otimes\Omega^{1/2}M_{kl}\Omega^{1/2})\tilde{K}_{nn}\\
&= 2\kappa_{i-j}\kappa_{k-l}(m_i'\Omega m_{k-l}m_k'\Omega m_{i-j} + m_i'\Omega m_k m_{i-j}'\Omega m_{k-l})\\
&\quad + (\varphi-3)\kappa_{i-j}\kappa_{k-l}\sum_{t=1}^{n} a_{tt}b_{tt},
\end{aligned}
$$

where a_{tt} (respectively b_{tt}) is the ttth element of $A = \Omega^{1/2}M_{ij}\Omega^{1/2}$ (respectively $\Omega^{1/2}M_{kl}\Omega^{1/2}$).

NOTE

1. An excellent solution has been independently proposed by Jose Matos Passos, the poser of the problem.

REFERENCE

Magnus J.R. & H. Neudecker (1979) The commutation matrix: Some properties and applications. *The Annals of Statistics*, 7, 381–394.

00.6.2. *A Distributional Equality*—Solution proposed by Paolo Parulo. Let

$$
\underset{T\times T}{C} := \begin{pmatrix} 1 & & & \\ 1 & 1 & & \\ . & . & . & \\ 1 & 1 & 1 & 1 \end{pmatrix}, \qquad \underset{T\times T}{R} := \begin{pmatrix} & & & 1 \\ & & 1 & \\ & . & & \\ 1 & & & \end{pmatrix}
$$

https://doi.org/10.1017/S0266466601006065 Published online by Cambridge University Press

and let $\stackrel{\mathcal{L}}{=}$ indicate equality in distribution. Observe that $RR = I$, and write $y := (y_1,\ldots,y_T)' = C\epsilon$, $i := (1,\ldots,1)'$, and $j := (1,2,\ldots,T)' = Ci$. Note that

$$a_T = T^{-1}y'My = T^{-1}\epsilon'C'MC\epsilon,$$

where $M := (I - iT^{-1}i')$ is idempotent. Likewise, $b_T = T^{-1}Q'Q$ where $Q = y - jT^{-1}y_T = C\epsilon - CiT^{-1}i'\epsilon = CM\epsilon$, so that $b_T = T^{-1}\epsilon'MC'CM\epsilon$. Thus a_T and b_T are quadratic forms in ϵ with associated p.s.d. matrices $C'MC$ and $MC'CM$; hence a_T and b_T can be represented as weighted averages of independent χ^2 variables, with weights equal to the eigenvalues of the associated matrices. If two matrices associated with two quadratic forms in standard normal variables have the same eigenvalues, the two quadratic forms have the same distribution. Since the eigenvalues of $MC'CM$ and $CMMC' = CMC'$ are the same, this shows that $b_T = T^{-1}\epsilon'MC'CM\epsilon \stackrel{\mathcal{L}}{=} T^{-1}\epsilon'CMC'\epsilon$.

Now observe that $C' = RCR$ and $RC'R = C$; note also that $Ri = i$ and $RMR = R(I - iT^{-1}i')R = RR - iT^{-1}i' = M$. One then has $CMC' = RC'RMRCR = RC'MCR$, and hence $b_T \stackrel{\mathcal{L}}{=} T^{-1}\epsilon'RC'MCR\epsilon =: T^{-1}\tilde{\epsilon}'C'MC\tilde{\epsilon}$, where $\tilde{\epsilon} := R\epsilon$ contains the same components of the vector ϵ in reverse order. It is simple to note that $\epsilon \stackrel{\mathcal{L}}{=} \tilde{\epsilon} \stackrel{\mathcal{L}}{=} N(0,I)$, and therefore $b_T \stackrel{\mathcal{L}}{=} T^{-1}\tilde{\epsilon}'C'MC\tilde{\epsilon} \stackrel{\mathcal{L}}{=} T^{-1}\epsilon'C'MC\epsilon = a_T$, proving $\mathcal{L}(a_T) = \mathcal{L}(b_T)$.

In order to see that this implies (1), let x_1 indicate the random variable on the left-hand side of (1), and x_2 indicate the random variable on the right-hand side, and $t = [Tu]$. Note that the invariance principle plus the continuous mapping theorem applied to $T^{-1/2}y_t = T^{-1/2}\sum_{i=1}^{t}\epsilon_i \to^w B_u$ implies that $T^{-1/2}a_T \to^w x_1, T^{-1/2}b_T \to^w x_2$. This with $\mathcal{L}(a_T) = \mathcal{L}(b_T)$ shows that also $\mathcal{L}(x_1) = \mathcal{L}(x_2)$.

REFERENCE

K. Tanaka (1996) *Time Series Analysis—Nonstationary and Noninvertible Distribution Theory*. New York: Wiley.

https://doi.org/10.1017/S0266466601006065 Published online by Cambridge University Press