The University of Sydney
School of Mathematics and Statistics

Solutions to Tutorial Week 12

STAT3023: Statistical Inference — Semester 2, 2022

Lecturers: Linh Nghiem, Michael Stewart and Rachel Wang

1. Suppose $X \sim B(n, \theta)$ and that $\tilde{d}(X)$ is the Bayes procedure based on a $U[\theta_0, \theta_1]$ prior under squared-error loss. Suppose also that for all $\theta_0 < \theta < \theta_1$,

$$\lim_{n\to\infty} nE_\theta \left\{ \left[\tilde{d}(X) - \theta \right]^2 \right\} \to \theta(1-\theta) .$$

Use the Asymptotic Minimax Lower Bound Theorem to show that the maximum likelihood estimator of θ is asymptotically minimax (over any interval $[a, b]$ for $0 < a < b < 1$).

Solution: Using the information above, according to the AMLB theorem, for any procedure (sequence) $\{d_n(\cdot)\}$ and $0 < a < b < 1$,

$$\lim_{n\to\infty} \max_{a\le\theta\le b} nE_\theta \left\{ \left[d_n(X) - \theta \right]^2 \right\} \ge \max_{a\le\theta\le b} \theta(1-\theta) . \tag{1}$$

The maximum likelihood estimator $\hat{\theta}_{\text{ML}} = X/n$ is unbiased so the risk is just the variance:

$$E_\theta \left\{ \left[\hat{\theta}_{\text{ML}} - \theta \right]^2 \right\} = \text{Var}_\theta \left(\hat{\theta}_{\text{ML}} \right) = \text{Var}_\theta \left(\frac{X}{n} \right) = \frac{1}{n^2}\text{Var}_\theta(X) = \frac{n\theta(1-\theta)}{n^2} = \frac{\theta(1-\theta)}{n} .$$

Thus the maximum (rescaled) risk over $[a, b] \subset [0, 1]$ is

$$\max_{a\le\theta\le b} nE_\theta \left\{ \left[\hat{\theta}_{\text{ML}} - \theta \right]^2 \right\} = \max_{a\le\theta\le b} n \left(\frac{\theta(1-\theta)}{n} \right) = \max_{a\le\theta\le b} \theta(1-\theta)$$

which attains the lower bound (1) above. Thus $\hat{\theta}_{\text{ML}}$ is asymptotically minimax.

2. Suppose $\mathbf{X} = (X_1, \ldots, X_n)$ consists of iid random variables with a gamma distribution with known shape α_0 but unknown *scale* parameter $\theta = \Theta = (0, \infty)$. Consider the decision problem where the decision space is $\mathcal{D} = \Theta$ and loss is $L(d|\theta) = (d - \theta)^2$. Write $T = \sum_{i=1}^n X_i$.

(a) Define the family of estimators $\{d_{k\ell}(\cdot) \colon k, \ell \in \mathbb{R}\}$ according to

$$d_{k\ell}(\mathbf{X}) = \frac{T + k}{n\alpha_0 + \ell} .$$

Determine the risk

$$R(\theta|d_{k\ell}) = E_\theta \left\{ \left[d_{k\ell}(\mathbf{X}) - \theta \right]^2 \right\} .$$

Solution: Firstly, T has a gamma distribution with shape parameter $n\alpha_0$ and *scale* parameter θ. Thus

$$E_\theta (T) = n\alpha_0\theta ;$$
$$Var_\theta (T) = n\alpha_0\theta^2 .$$

Therefore

$$E_\theta \left[d_{k\ell}(\mathbf{X}) \right] = \frac{E_\theta(T) + k}{n\alpha_0 + \ell} = \frac{n\alpha_0\theta + k}{n\alpha_0 + \ell} ;$$

$$Bias_\theta \left[d_{k\ell}(\mathbf{X}) \right] = E_\theta \left[d_{k\ell}(\mathbf{X}) \right] - \theta = \frac{n\alpha_0\theta + k - \theta(n\alpha_0 + \ell)}{n\alpha_0 + \ell} = \frac{k - \ell\theta}{n\alpha_0 + \ell} ;$$

$$Var_\theta \left[d_{k\ell}(\mathbf{X}) \right] = \frac{\text{Var}_\theta(T)}{(n\alpha_0 + \ell)^2} = \frac{n\alpha_0\theta^2}{(n\alpha_0 + \ell)^2} .$$

Thus the risk is

$$R(\theta|d_{k\ell}) = E_\theta\left\{[d_{k\ell}(\mathbf{X}) - \theta]^2\right\} = Var_\theta\left[d_{k\ell}(\mathbf{X})\right] + \left\{Bias_\theta\left[d_{k\ell}(\mathbf{X})\right]\right\}^2 = \frac{n\alpha_0\theta^2 + (k-\ell\theta)^2}{(n\alpha_0+\ell)^2}\,.$$

(b) Determine $d_{\text{flat}}(\mathbf{X})$, the Bayes procedure using the "flat prior" $w(\theta) \equiv 1$.

Solution: The likelihood is

$$f_\theta(\mathbf{X}) = \prod_{i=1}^{n}\left[\frac{X_i^{\alpha_0-1}e^{-X_i/\theta}}{\theta^{\alpha_0}\Gamma(\alpha_0)}\right] = \text{const.}\ \frac{e^{-T/\theta}}{\theta^{n\alpha_0}} = \text{const.}\ \frac{T^{n\alpha_0-1}e^{-T/\theta}}{\theta^{(n\alpha_0-1)+1}\Gamma(n\alpha_0-1)}$$

so the posterior is the Inverse Gamma$(n\alpha_0 - 1, T)$ distribution (this is the distribution of $1/Y$ where Y is gamma with shape $n\alpha_0 - 1$ and *rate* T).

The Bayes procedure is thus the posterior *mean* (since we are using squared-error loss), which is

$$d_{\text{flat}}(\mathbf{X}) = \frac{T}{n\alpha_0 - 2}$$

(c) Show that for any $k, \ell \in \mathbb{R}$, $d_{k\ell}(\mathbf{X})$ is asymptotically minimax. You may assume that for any $0 \le \theta_0 < \theta_1 < \infty$, the Bayes procedure $\widetilde{d}(\mathbf{X})$ based on the $U[\theta_0, \theta_1]$ prior has the same limiting (rescaled) risk as $d_{\text{flat}}(\mathbf{X})$: for all $\theta_0 < \theta < \theta_1$,

$$\lim_{n\to\infty} nR(\theta|\widetilde{d}) = \lim_{n\to\infty} nR(\theta|d_{\text{flat}})\,. \tag{2}$$

Solution: First we need to determine the RHS of (2) above. Note that $d_{\text{flat}}(\mathbf{X})$ is a special case of $d_{k\ell}(\mathbf{X})$ examined in part (a) above, corresponding to $k = 0$ and $\ell = -2$. Therefore we can read the exact risk from that part as

$$R(\theta|d_{k\ell}) = \frac{n\alpha_0\theta^2 + (2\theta)^2}{(n\alpha_0 - 2)^2}$$

so as $n \to \infty$,

$$\begin{aligned}
nR(\theta|d_{\text{flat}}) &= \frac{n^2\alpha_0\theta^2}{(n\alpha_0-2)^2} + \frac{4n\theta^2}{(n\alpha_0-2)^2}\\
&= \frac{n^2\alpha_0\theta^2}{n^2\alpha_0^2\left(1-\frac{2}{n\alpha_0}\right)^2} + \frac{4n\theta^2}{n^2\alpha_0^2\left(1-\frac{2}{n\alpha_0}\right)^2}\\
&= \underbrace{\frac{\theta^2}{\alpha_0\left(1-\frac{2}{n\alpha_0}\right)^2}}_{\to\frac{\theta^2}{\alpha_0}} + \underbrace{\frac{4\theta^2}{n\alpha_0^2\left(1-\frac{2}{n\alpha_0}\right)^2}}_{\to 0}\\
&\to \frac{\theta^2}{\alpha_0} = S(\theta)\,.
\end{aligned}$$

So for any procedure (sequence) $\{d_n(\cdot)\}$, by the AMLB Theorem (and the assumption (2) given above),

$$\lim_{n\to\infty}\max_{a\le\theta\le b} nR(\theta|d_n) \ge \max_{a\le\theta\le b} S(\theta) = \frac{b^2}{\alpha_0}\,. \tag{3}$$

Finally, we need to derive the limiting maximum (rescaled) risk of $d_{k\ell}$:

$$\begin{aligned}
\max_{a\le\theta\le b} nR(\theta|d_{k\ell}) &= \max_{a\le\theta\le b} n\left\{\frac{n\alpha_0\theta^2 + (k-\ell\theta)^2}{(n\alpha_0+\ell)^2}\right\}\\
&\le \max_{a\le\theta\le b}\frac{n^2\alpha_0\theta^2}{(n\alpha_0+\ell)^2} + \max_{a\le\theta\le b}\frac{n(k-\ell\theta)^2}{(n\alpha_0+\ell)^2}\\
&= b^2\underbrace{\frac{\alpha_0}{\left(\alpha_0+\frac{\ell}{n}\right)^2}}_{\to\frac{1}{\alpha_0}} + \underbrace{\frac{1}{n\left(\alpha_0+\frac{\ell}{n}\right)^2}}_{\to 0}\underbrace{\max_{a\le\theta\le b}(k-\ell\theta)^2}_{<\infty}\\
&\to \frac{b^2}{\alpha_0}
\end{aligned}$$

which attains the lower bound (3) above. Thus for each fixed k, ℓ, $d_{k\ell}$ is asymptotically minimax.

(d) Show that

(i) the maximum likelihood estimator;

(ii) $d_{\text{flat}}(\mathbf{X})$;

(iii) any Bayes procedure based on an Inverse Gamma (conjugate) prior

are all asymptotically minimax.

Solution: The derivative of the log-likelihood with respect to θ is

$$\ell'(\theta; \mathbf{X}) = -\frac{n\alpha_0}{\theta} + \frac{T}{\theta^2}\,;$$

setting equal to zero and solving gives

$$\hat{\theta}_{\text{ML}} = \frac{T}{n\alpha_0} = \frac{\bar{X}}{\alpha_0}\,.$$

For any fixed $\gamma_0, \lambda_0 > 0$, taking as prior the Inverse Gamma(γ_0, λ_0) density

$$w(\theta) = \frac{\lambda_0^{\gamma_0} e^{-\lambda_0/\theta}}{\theta^{\gamma_0+1}\Gamma(\gamma_0)}$$

the product of the prior and the likelihood

$$w(\theta) f_\theta(\mathbf{X}) = \text{const.}\ \frac{e^{-(T+\lambda_0)/\theta}}{\theta^{n\alpha_0+\gamma_0+1}} = \text{const.}\ \frac{(T+\lambda_0)^{n\alpha_0+\gamma_0} e^{-(T+\lambda_0)/\theta}}{\theta^{n\alpha_0+\gamma_0+1}\Gamma(n\alpha_0+\gamma_0)}$$

so the posterior density is the Inverse Gamma$(n\alpha_0+\gamma_0, T+\lambda_0)$ density. The corresponding Bayes procedure is the *posterior mean*, which is

$$\frac{T+\lambda_0}{n\alpha_0+\gamma_0-1}\,.$$

Note that this, $d_{\text{flat}}(\mathbf{X})$ and $\hat{\theta}_{\text{ML}}$ are all special cases of $d_{k\ell}(\mathbf{X})$ and thus by the previous part are all asymptotically minimax.

3. The beta function is given by

$$\text{beta}(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\,dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

(where $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx$ is the gamma function, satisfying $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, for all $\alpha > 0$), and is the normalising constant in the beta(α, β) density:

$$f_X(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\text{beta}(\alpha,\beta)} \quad \text{for } 0 < x < 1.$$

Suppose X has the density $f_X(\cdot)$ above, and then define $Y = 1/X$.

(a) For $\alpha > 1$, determine $E(Y)$.

Solution:

$$\begin{aligned} E(Y) = E(X^{-1}) = \int_0^1 x^{-1}\frac{x^{\alpha-1}(1-x)^{\beta-1}}{\text{beta}(\alpha,\beta)}\,dx &= \frac{\int_0^1 x^{(\alpha-1)-1}(1-x)^{\beta-1}}{\text{beta}(\alpha,\beta)} \\ &= \frac{\text{beta}(\alpha-1,\beta)}{\text{beta}(\alpha,\beta)} \\ &= \frac{\Gamma(\alpha-1)\Gamma(\beta)\,\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta-1)\,\Gamma(\alpha)\Gamma(\beta)} \\ &= \frac{\Gamma(\alpha-1)}{\Gamma(\alpha)}\ \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta-1)} \\ &= \frac{\alpha+\beta-1}{\alpha-1}\,, \end{aligned}$$

using the property $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ twice.

(b) Determine the density of Y.

Solution: Note that since $0 < X < 1$, $1 < Y < \infty$. Using the "CDF method", the CDF of Y is, for $y > 1$,

$$F_Y(y) = P(Y \leq y) = P(X^{-1} \leq y) = P(X \geq y^{-1}) = 1 - P(X < y^{-1}) = 1 - F_X\left(y^{-1}\right)$$

since X has a continuous distribution. Therefore the density of Y is

$$\begin{aligned} f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}\left[1 - F_X\left(y^{-1}\right)\right] &= -f_X\left(y^{-1}\right)\frac{d}{dy}\left(y^{-1}\right) \\ &= \frac{1}{y^2}f_X\left(\frac{1}{y}\right) \\ &= \frac{1}{\text{beta}(\alpha,\beta)}\left(\frac{1}{y}\right)^{\alpha-1}\left(1-\frac{1}{y}\right)^{\beta-1}\frac{1}{y^2} \\ &= \frac{1}{\text{beta}(\alpha,\beta)}\frac{(y-1)^{\beta-1}}{y^{\alpha+\beta}}. \end{aligned}$$

4. If Y has a geometric distribution with

$$P(Y = y) = (1-p)^{y-1}p \text{ for } y = 1, 2, \cdots$$

then $E(Y) = 1/p$ and $Var(Y) = (1-p)/p^2$. Suppose $\mathbf{X} = (X_1, \ldots, X_n)$ consists of iid geometric random variables with unknown *mean* $\theta \in \Theta = (1, \infty)$. Consider the decision problem with decision space $\mathcal{D} = \Theta$ and loss $L(d|\theta) = (d-\theta)^2$. Assume that $n \geq 3$.

(a) Determine $E_\theta(T)$ and $Var_\theta(T)$ where $T = \sum_{i=1}^n X_i$ as functions of θ.

Solution: Writing $p = 1/\theta$ we have

$$E_\theta(X_1) = \frac{1}{p} = \theta \text{ and } Var_\theta(X_1) = \left(\frac{1}{p^2} - \frac{1}{p}\right) = \theta^2 - \theta = \theta(\theta-1).$$

Therefore

$$E_\theta(T) = n\theta \text{ and } Var_\theta(T) = n\theta(\theta-1).$$

(b) Define the family of estimators $\{d_{k\ell}(\cdot)\colon k, \ell \in \mathbb{R}\}$ according to

$$d_{k\ell}(\mathbf{X}) = \frac{T+k}{n+\ell}.$$

Determine the risk

$$R(\theta|d_{k\ell}) = E_\theta\left\{[d_{k\ell}(\mathbf{X}) - \theta]^2\right\}.$$

Solution: We have

$$\begin{aligned} E_\theta\left[d_{k\ell}(\mathbf{X})\right] &= \frac{E_\theta(T)+k}{n+\ell} = \frac{n\theta+k}{n+\ell}; \\ Bias_\theta\left[d_{k\ell}(\mathbf{X})\right] &= E_\theta\left[d_{k\ell}(\mathbf{X})\right] - \theta = \frac{n\theta+k-\theta(n+\ell)}{n+\ell} = \frac{k-\ell\theta}{n+\ell}; \\ Var_\theta\left[d_{k\ell}(\mathbf{X})\right] &= \frac{Var_\theta(T)}{(n+\ell)^2} = \frac{n\theta(\theta-1)}{(n+\ell)^2}. \end{aligned}$$

Thus the risk is

$$R(\theta|d_{k\ell}) = E_\theta\left\{[d_{k\ell}(\mathbf{X})-\theta]^2\right\} = Var_\theta\left[d_{k\ell}(\mathbf{X})\right] + \left\{Bias_\theta\left[d_{k\ell}(\mathbf{X})\right]\right\}^2 = \frac{n\theta(\theta-1)+(k-\ell\theta)^2}{(n+\ell)^2}.$$

(c) Write down the probability mass function of X_1 as a function of θ.

Solution:

$$P_\theta(X_1 = x) = \left(1 - \frac{1}{\theta}\right)^{x-1}\frac{1}{\theta} = \frac{(\theta-1)^{x-1}}{\theta^x}, \quad \text{for } x = 1, 2, \ldots$$

(d) Write out the likelihood.

Solution:

$$f_\theta(\mathbf{X}) = \prod_{i=1}^{n}\left[\frac{(\theta-1)^{X_i-1}}{\theta^{X_i}}\right] = \frac{(\theta-1)^{T-n}}{\theta^T}\,.$$

(e) Determine the Bayes procedure $d_{\text{flat}}(\mathbf{X})$ using a flat prior $w(\theta)\equiv 1$ (question 3 may prove useful here).

Solution: The product of the likelihood and the prior is of the form

$$w(\theta)f_\theta(\mathbf{X}) = \frac{(\theta-1)^{T-n}}{\theta^T}\text{.const.}\ \frac{1}{\text{beta}(n-1,T-n+1)}\ \frac{(\theta-1)^{(T-n+1)-1}}{\theta^{(n-1)+(T-n+1)}}$$

so according to part (b) of question 3 the posterior density is that of $1/Y$ where $Y\sim\text{beta}(n-1,T-n+1)$.

Since we are using squared-error loss, the Bayes procedure is the posterior *mean*, which, according to part (a) of question 3 is

$$d_{\text{flat}}(\mathbf{X}) = \frac{(n-1)+(T-n+1)-1}{(n-1)-1} = \frac{T-1}{n-2}\,;$$

note that since $n>2$ this is finite (this corresponds to $\alpha>1$ in the previous question).

(f) Show that

(i) the maximum likelihood estimator;

(ii) $d_{\text{flat}}(\mathbf{X})$;

(iii) any Bayes procedure based on a (conjugate) prior of the form

$$w(\theta) = \frac{1}{\text{beta}(\alpha_0,\beta_0)}\ \frac{(\theta-1)^{\beta_0-1}}{\theta^{\alpha_0+\beta_0}}\,,\quad \text{for } \theta>1 \tag{4}$$

are all asymptotically minimax. You may assume that for any $1<\theta_0<\theta_1<\infty$, the Bayes procedure $\widetilde{d}(\mathbf{X})$ based on the $U[\theta_0,\theta_1]$ prior has the same limiting (rescaled) risk as $d_{\text{flat}}(\mathbf{X})$: for all $\theta_0<\theta<\theta_1$,

$$\lim_{n\to\infty} nR(\theta|\widetilde{d}) = \lim_{n\to\infty} nR(\theta|d_{\text{flat}})\,.$$

Hint: determine the forms of all the estimators first.

Solution: The *form* of the mle may be obtained by differentiating the log-likelihood $\ell(\theta;\mathbf{X}) = (T-n)\log(\theta-1) - T\log(\theta)$, setting to zero and solving:

$$\begin{aligned}\ell'(\theta;\mathbf{X}) = \frac{T-n}{\theta-1} - \frac{T}{\theta} = T\left(\frac{1}{\theta-1}-\frac{1}{\theta}\right) - \frac{n}{\theta-1} &= T\left(\frac{\theta-(\theta-1)}{\theta(\theta-1)}\right) - \frac{n}{\theta-1}\\ &= \frac{T}{\theta(\theta-1)} - \frac{n}{\theta-1}\\ &= \frac{n}{\theta(\theta-1)}\left(\frac{T}{n}-\theta\right)\end{aligned}$$

yielding $\hat{\theta}_{\text{ML}} = T/n = \bar{X}$ the sample mean; this also shows that the score function is in the "nice form" indicating that $\bar{X}$ is MVU in this case.

Using the conjugate prior $w(\theta)$ given at (4) above, the product of the likelihood and the prior is of the form

$$f_\theta(\mathbf{X})w(\theta) = \text{const.}\ \frac{(\theta-1)^{T-n+\beta_0-1}}{\theta^{T+\alpha_0+\beta_0}} = \text{const.}\ \frac{1}{\text{beta}(n+\alpha_0,T-n+\beta_0)}\ \frac{(\theta-1)^{(T-n+\beta_0)-1}}{\theta^{(T-n+\beta_0)+(n+\alpha_0)}}\,,$$

so the posterior density is that of $1/Y$ where Y has a $\text{beta}(n+\alpha_0,T-n+\beta_0)$ distribution. According to question 3 part (a), the mean of this distribution (which is also the posterior mean, i.e. the Bayes procedure) is

$$\frac{T+\alpha_0+\beta_0-1}{n+\alpha_0-1}\,.$$

Note that all estimators of this form, the mle $\hat{\theta}_{\mathrm{ML}}$ and $d_{\mathrm{flat}}(\mathbf{X})$ are all special cases of $d_{k\ell}(\mathbf{X})$ defined in part (b) above. It thus suffices to show that $d_{k\ell}(\mathbf{X})$ is asymptotically minimax for each k, ℓ.
Firstly, for any $1 < \theta_0 < \theta_1 < \infty$, the Bayes procedure $\widetilde{d}(\mathbf{X})$ based on the $U[\theta_0, \theta_1]$ prior has limiting risk equal to

$$\lim_{n\to\infty} nR(\theta|\widetilde{d}) = \lim_{n\to\infty} nR(\theta|d_{\mathrm{flat}})\,.$$

But since $d_{\mathrm{flat}}(\mathbf{X})$ is a special case of $d_{k\ell}(\mathbf{X})$ (with $k = -1$, $\ell = -2$), we can use the results of part (b) above to determine the limiting (rescaled) risk:

$$\begin{aligned}\lim_{n\to\infty} nR(\theta|d_{\mathrm{flat}}) &= \lim_{n\to\infty} n\left\{\frac{n\theta(\theta-1) + (2\theta-1)^2}{(n-2)^2}\right\}\\ &= \theta(\theta-1)\lim_{n\to\infty}\underbrace{\left(\frac{n}{n-2}\right)^2}_{\to 1} + (2\theta-1)^2\lim_{n\to\infty}\underbrace{\left(\frac{n}{(n-2)^2}\right)}_{\to 0}\\ &= \theta(\theta-1)\,.\end{aligned}$$

Note also that the contribution from the bias (the $\frac{(2\theta-1)^2}{(n-2)^2}$ term) is "asymptotically negligible" compared to the variance contribution.
Thus we have

$$\lim_{n\to\infty} nR(\theta|\widetilde{d}) = \theta(\theta-1) = S(\theta)$$

for all $\theta_0 < \theta < \theta_1$. Thus according to the Asymptotic Minimax Lower Bound Theorem, for any other procedure (sequence) $\{d_n(\cdot)\}$, for any $1 < a < b$,

$$\lim_{n\to\infty}\max_{a\le\theta\le b} nR(\theta|d_n) \ge \max_{a\le\theta\le b} S(\theta) = \max_{a\le\theta\le b}\theta(\theta-1) = b(b-1)$$

since $S(\theta)$ is an increasing function for $\theta > 1$.
Now, the maximum risk of $d_{k\ell}(\mathbf{X})$ is

$$\max_{a\le\theta\le b} R(\theta|d_{k\ell}) \le \frac{n}{(n+\ell)^2}\max_{a\le\theta\le b}\theta(\theta-1) + \frac{1}{(n+\ell)^2}\max_{a\le\theta\le b}(k-\ell\theta)^2\,.$$

The term $(k-\ell\theta)^2$ is a parabola in θ with a positive coefficient of θ^2, so takes its maximum value over $a \le \theta \le b$ at one of the endpoints. So

$$\begin{aligned}\lim_{n\to\infty} n\max_{a\le\theta\le b} R(\theta|d_{k\ell}) &\le \max_{a\le\theta\le b}\theta(\theta-1)\lim_{n\to\infty}\overbrace{\left(\frac{n}{n+\ell}\right)^2}^{\to 1}\\ &\quad + \max\left[(k-a\ell)^2, (k-b\ell)^2\right]\lim_{n\to\infty}\underbrace{\frac{n}{(n+\ell)^2}}_{\to 0}\\ &= \max_{a\le\theta\le b}\theta(\theta-1) = \max_{a\le\theta\le b} S(\theta)\,.\end{aligned}$$

However this upper bound is also the lower bound for any estimator. Therefore for each k, ℓ, $d_{k\ell}(\mathbf{X})$ is asymptotically minimax. Thus each of the estimators above is too.