STAT6201: Theoretical Statistics I

Satvik Saha

Midterm exam

1. The proof claims that $\eta(T) = \mathbb{E}[\delta \mid T]$ is an estimator that outperforms δ. For this, we require that η is indeed an estimator, i.e. $\mathbb{E}[\delta \mid T]$ should not be a function of θ. We can guarantee this by demanding that the conditional distribution $\delta \mid T = t$ be independent of θ, which is achieved when T is sufficient for θ.

   Remark: T is sufficient for θ when the distribution $X \mid T = t$ is free of θ for all t.

   For a generic statistic T, we may not have $\mathbb{E}[\delta \mid T]$ calculable without knowledge of θ. For instance, consider $X_1, X_2 \overset{iid}{\sim} N(\theta, 1)$, whence $\mathbb{E}[X_1 \mid X_2] = \mathbb{E}[X_1] = \theta$, which is not a statistic.

2. Let $X_1, \dots, X_n \overset{iid}{\sim} N(\theta, \theta^2)$ for $\theta \in \Theta = (0, \infty)$.

   (a) We have the likelihood

   $$\mathcal{L}(\theta \mid X) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\theta^2}} e^{-(X_i - \theta)^2/2\theta^2} = (2\pi)^{-n/2}\theta^{-n} \exp\left(-\frac{1}{2\theta^2}\sum_{i=1}^{n}(X_i - \theta)^2\right).$$

   Thus, the log likelihood is given by

   $$\ell(\theta \mid X) = -\frac{n}{2}\log(2\pi) - n\log(\theta) - \frac{1}{2\theta^2}\sum_{i=1}^{n}(X_i - \theta)^2.$$

   Maximizing this is equivalent to minimizing

   $$n\log\theta + \frac{1}{2\theta^2}\sum_{i=1}^{n}\left[X_i^2 - 2X_i\theta + \theta^2\right] = n\log\theta - \frac{1}{\theta}\sum_{i=1}^{n} X_i + \frac{1}{2\theta^2}\sum_{i=1}^{n} X_i^2 + \frac{n}{2},$$

   hence we need only minimize

   $$g(\theta) = n\log\theta - \frac{S_1}{\theta} + \frac{S_2}{2\theta^2},$$

   where we abbreviate $S_1 = \sum_i X_i$, $S_2 = \sum_i X_i^2$. Compute

   $$g'(\theta) = \frac{n}{\theta} + \frac{S_1}{\theta^2} - \frac{S_2}{\theta^3}.$$

   This vanishes when $p(\theta) = n\theta^2 + S_1\theta - S_2 = 0$. The positive value of θ for which this is true is

   $$\theta_0 = -\frac{S_1}{2n} + \frac{\sqrt{S_1^2 + 4nS_2}}{2n} = -\frac{S_1}{2n} + \sqrt{\frac{S_1^2}{4n^2} + \frac{S_2}{n^2}}.$$

This is indeed positive since $\sqrt{S_1^2 + 4nS_2} \geq S_1$. To check when this is the minimizer of g, we compute

$$
\begin{aligned}
g''(\theta_0) &= -\frac{n}{\theta_0^2} - \frac{2S_1}{\theta_0^3} + \frac{3S_2}{\theta_0^4} \\
&= -\frac{1}{\theta_0^4}\left[n\theta_0^2 + 2S_1\theta_0 - 3S_2\right] \\
&= -\frac{1}{\theta_0^4}\left[\cancelto{0}{(n\theta_0^2 + S_1\theta_0 - S_2)} + S_1\theta_0 - 2S_2\right] \\
&= -\frac{1}{\theta_0^4}\left[S_1\theta_0 - 2S_2\right] \\
&= \frac{1}{2n\theta_0^4}\left[S_1^2 - S_1\sqrt{S_1^2 + 4nS_2} + 4nS_2\right].
\end{aligned}
$$

This is non-negative when $\sqrt{S_1^2 + 4nS_2} \geq S_1$, which always holds. Also, note that as $\theta \to \infty$ or as $\theta \to 0$, we have $g(\theta) \to \infty$. Thus, we have

$$\hat{\theta}_{\text{MLE}} = -\frac{S_1}{2n} + \sqrt{\frac{S_1^2}{4n^2} + \frac{S_2}{n^2}}.$$

(b) From our earlier computation of the likelihood, note that

$$\mathcal{L}(\theta \mid X) = (2\pi)^{-n/2}\theta^{-n}\exp\left(-\frac{S_2}{2\theta^2} + \frac{S_1}{\theta}\right)\exp\left(-\frac{n}{2}\right).$$

By the Neyman-Fisher factorization theorem, we see that (S_1, S_2) is sufficient for $\theta \in \Theta$. Furthermore, this is minimal sufficient for θ. To see this, note that the natural parameter space $\{(-1/2\theta^2, 1/\theta)\}_{\theta \in (0,\infty)}$ contains three affinely independent vectors; say for $\theta = 1, 2, 3$, we have natural parameters $(-1/2, 1), (-1/8, 1/2), (-1/18, 1/3)$ which are not collinear.

With this, for $\hat{\theta}_{\text{MLE}}$ to be sufficient for θ, we would have to be able to write (S_1, S_2) as a (measurable) function of $\hat{\theta}_{\text{MLE}}$, which is not possible. Thus, $\hat{\theta}_{\text{MLE}}$ is not sufficient for θ.

3. Let $X_1, \ldots, X_n \overset{iid}{\sim} p_\theta$, with $\theta \sim \pi$. Further assume that the variables and parameters are real valued.

(a) Set

$$R_n = \inf_{\delta \in \mathscr{D}_n} \mathbb{E}\left|\delta(X_1, \ldots, X_n) - \theta\right|,$$

where $\mathscr{D}_n$ is the class of estimators of the form $\delta \colon \mathscr{X}^n \to \Theta$. Then, for each $\delta \in \mathscr{D}_n$, we may define $\delta_0 \in \mathscr{D}_{n+1}$ where $\delta_0(X_1, \ldots, X_n, X_{n+1}) = \delta(X_1, \ldots, X_n)$. Thus, we have

$$\mathbb{E}\left|\delta(X_1, \ldots, X_n) - \theta\right| = \mathbb{E}\left|\delta_0(X_1, \ldots, X_n, X_{n+1}) - \theta\right| \geq R_{n+1}.$$

Taking infimums,

$$R_n = \inf_{\delta \in \mathscr{D}_n} \mathbb{E}\left|\delta(X_1, \ldots, X_n) - \theta\right| \geq R_{n+1}.$$

This shows that $\{R_n\}_{n \in \mathbb{N}}$ is a non-increasing sequence.

(b) Set

$$B_n = \mathbb{E}\left|\mathbb{E}(\theta \mid X_1, \dots, X_n) - \theta\right|.$$

Also denote $\delta_n(X) = \mathbb{E}[\theta \mid X_1, \dots, X_n]$.

Remark: Note that

$$\mathbb{E}\left[(\delta_n(X) - \theta)^2\right] = \inf_{\delta \in \mathscr{D}_n} \mathbb{E}\left[(\delta(X) - \theta)^2\right].$$

The right minimizer for the absolute error loss is the posterior *median*, not the posterior *mean*.

Remark: Note that $\{B_n\}$ is indeed decreasing in settings such as $X_i \overset{iid}{\sim} N(\theta, 1)$, $\pi(\theta) = 1$.

Remark: We have

$$R_n^2 \leq B_n^2 \leq \mathbb{E}\left[(\delta_n(X) - \theta)^2\right] = \inf_{\delta \in \mathscr{D}_n} \mathbb{E}\left[(\delta(X) - \theta)^2\right] := S_n^2,$$

with $\{S_n\}$ non-increasing via the same argument as before.

Remark: For $\{B_n\}$ to be non-increasing, we need each

$$\mathbb{E}\left[|\delta_n(X) - \theta| - |\delta_{n+1}(X) - \theta|\right] \geq 0.$$

4. Let $X_1, \dots, X_n \overset{iid}{\sim} \frac{1}{2}N(\theta, 1) + \frac{1}{2}N(-\theta, 1)$, for $\theta \in [0, \infty)$. Then, the likelihood is of the form

$$\begin{aligned}
\mathcal{L}(\theta \mid X) &= \prod_{i=1}^n \frac{1}{2\sqrt{2\pi}}\left[e^{-(X_i - \theta)^2/2} + e^{-(X_i + \theta)^2/2}\right] \\
&= \prod_{i=1}^n \frac{1}{2\sqrt{2\pi}}\left[2e^{-X_i^2/2}e^{-\theta^2/2}\cosh(X_i\theta)\right].
\end{aligned}$$

Note that this is unchanged by replacing $(X_1, \dots, X_n)$ with $T(X) = (|X|_{(1)}, \dots, |X|_{(n)})$, the ordered absolute values of the sample. It is clear that this is sufficient for $\theta \in [0, \infty)$. We further claim that this is minimal sufficient for θ.

Examine

$$\frac{\mathcal{L}(\theta \mid X)}{\mathcal{L}(\theta \mid Y)} = \prod_{i=1}^n \frac{e^{-X_i^2/2}\cosh(X_i\theta)}{e^{-Y_i^2/2}\cosh(Y_i\theta)}.$$

It is clear that $T(X) = T(Y)$ makes this ratio identically 1, hence independent of θ.

On the other hand, suppose that this ratio is free of θ for all $\theta \in [0, \infty)$. Then, so is

$$r(X, Y) = \prod_{i=1}^n \frac{\cosh(X_i\theta)}{\cosh(Y_i\theta)} > 0.$$

Putting $\theta = 0$ tells us that $r(X, Y) = 1$. Write

$$\prod_{i=1}^n \cosh(X_i\theta) = \prod_{i=1}^n \cosh(Y_i\theta), \qquad \text{for all } \theta \in [0, \infty). \tag{$\star$}$$

We will turn these products into sums, by brute force. Note that $2\cosh(W_1)\cosh(W_2) = \cosh(W_1 + W_2) + \cosh(W_1 - W_2)$. Applying this repeatedly, for $0 \leq W_1 \leq \cdots \leq W_n$, we may write

$$2^{n-1}\prod_{i=1}^n \cosh(W_i) = \sum_{J \subseteq \{0,1\}^{n-1}} \cosh\left(\sum_{j=1}^{n-1}(-1)^{J_j}W_j + W_n\right) := \sum_{J \subseteq \{0,1\}^{n-1}} \cosh(W^J).$$

The sum on the right has 2^{n-1} terms. Importantly, the two largest terms in the sum are those corresponding to $W^0 := W_1 + (W_2 + \cdots + W_n)$ and $W^1 := -W_1 + (W_2 + \cdots + W_n)$ inside the hyperbolic cosines.

Next, observe that for $W, Z \geq 0$,

$$2e^{-Z\theta}\cosh(W\theta) = e^{-Z\theta}\left(e^{W\theta} + e^{-W\theta}\right) = e^{(-Z+W)\theta} + e^{(-Z-W)\theta}.$$

Taking limits as $\theta \to \infty$,

$$\lim_{\theta\to\infty} 2e^{-Z\theta}\cosh(W\theta) = \begin{cases} 1, & \text{if } W = Z, \\ 0, & \text{if } W < Z, \\ \infty, & \text{if } W > Z. \end{cases}$$

With this, define for multi-indices $J \subseteq \{0,1\}^{n-1}$ the terms

$$X^J = \sum_{j=1}^{n}(-1)^{J_j}|X|_{(j)} + |X|_{(n)}, \qquad Y^J = \sum_{j=1}^{n}(-1)^{J_j}|Y|_{(j)} + |Y|_{(n)},$$

These are merely sums involving the elements of $T(X)$ and $T(Y)$, with signs inserted according to J. Specifically denote $X^0 := X^{(0,\dots,0)}$, $X^1 := X^{(1,0,\dots,0)}$, and similarly for Y^0, Y^1. Then, $(\star)$ says that for all $\theta \in [0,\infty)$,

$$\sum_{J\subseteq\{0,1\}^{n-1}} \cosh(X^J\theta) = \sum_{J\subseteq\{0,1\}^{n-1}} \cosh(Y^J\theta). \tag{$\star\star$}$$

Suppose without loss of generality[1] that $X^0 \geq Y^0$. Then,

$$\begin{aligned} 2e^{-X^0\theta}\cosh(X^0\theta) + \sum_{\substack{J\subseteq\{0,1\}^{n-1} \\ J\neq(0,\dots,0)}} 2e^{-X^0\theta}\cosh(X^J\theta) \\ = 2e^{-X^0\theta}\cosh(Y^0\theta) + \sum_{\substack{J\subseteq\{0,1\}^{n-1} \\ J\neq(0,\dots,0)}} 2e^{-X^0\theta}\cosh(Y^J\theta). \end{aligned}$$

If we had $X^0 > Y^0$, taking limits as $\theta \to \infty$, the left hand side gives 1 while the right hand side gives 0, a contradiction. Thus, we must have $X^0 = Y^0$. We can now cancel the first term from $(\star\star)$. Again, without loss of generality, suppose that $X^1 \geq Y^1$. Multiplying by $2e^{-X^1\theta}$, we obtain

$$\begin{aligned} 2e^{-X^1\theta}\cosh(X^1\theta) + \sum_{\substack{J\subseteq\{0,1\}^{n-1} \\ J\neq(0,\dots,0) \\ J\neq(1,0,\dots,0)}} 2e^{-X^1\theta}\cosh(X^J\theta) \\ = 2e^{-X^1\theta}\cosh(Y^1\theta) + \sum_{\substack{J\subseteq\{0,1\}^{n-1} \\ J\neq(0,\dots,0) \\ J\neq(1,0,\dots,0)}} 2e^{-X^1\theta}\cosh(Y^J\theta). \end{aligned}$$

Like before, if we had $X^1 > Y^1$, taking limits as $\theta \to \infty$ would give 1 on the left side, 0 on the right, a contradiction. Thus, we must have $X^1 = Y^1$.

Subtracting $X^0 = Y^0$ and $X^1 = Y^1$ yields $|X|_{(1)} = |Y|_{(1)}$. Canceling these terms in $(\star)$, we have one fewer term; repeating the same argument as above will successively yield each $|X|_{(i)} = |Y|_{(i)}$, hence $T(X) = T(Y)$.

[1] If not, interchange the roles of X and Y.

This proves that T is indeed minimal sufficient for $\theta \in [0, \infty)$, via the result in HW1, Problem 2(a).

Remark: Since zero sample values, ties, etc in the sample occur with probability zero, we ignore such cases.

Remark: Alternatively, we can check that the left and right hand sides of $(\star)$ are analytic functions of θ, hence must be equal for all $\theta \in \mathbb{C}$. Since $\cosh(z) = 0$ precisely when $z = (n + \frac{1}{2})\pi i$, $n \in \mathbb{N}$, we can compare roots on both sides of $(\star)$, perhaps in order of magnitude of X_i, Y_i, and reach the same conclusion.

5. Let $X_1, \ldots, X_n \overset{iid}{\sim} U(0, \theta)$ for $\theta \in (0, \infty)$. Set $\delta_1(X) = X_1$, and $\delta_2(X) = \mathbb{E}_\theta[X_1 \mid X_{(n)}]$.

   (a) Note that $X_{(n)}$ is a complete sufficient statistic for $\theta > 0$. Indeed, the likelihood

$$\mathcal{L}(\theta \mid X) = \prod_{i=1}^{n} \frac{1}{\theta} \mathbf{1}_{(0,\theta)}(X_i) = \frac{1}{\theta^n} \mathbf{1}_{(0,\theta)}(X_{(n)}),$$

shows that $X_{(n)}$ is sufficient for θ via Neyman-Fisher factorization. Furthermore, we have for $0 \leq x \leq 1$,

$$P(X_{(n)} \leq x) = P(X_1 \leq x, \ldots, X_n \leq x) = \frac{x^n}{\theta^n},$$

so $f_{X_{(n)}}(x) = nx^{n-1}\theta^{-n}\mathbf{1}_{(0,\theta)}(x)$. It follows that

$$\mathbb{E}[X_{(n)}] = \int_0^\theta \frac{nx^n}{\theta^n}\, dx = \frac{n\theta}{n+1},$$

so $(n+1)X_{(n)}/n$ is unbiased for θ. Now suppose that for measurable $h\colon \mathbb{R}_+ \to \mathbb{R}$, we have $\mathbb{E}_\theta[h(X_{(n)})] = 0$ for all $\theta > 0$. This would imply that

$$\int_0^\theta h(x)\, x^{n-1}\, dx = 0, \qquad \text{for all } \theta > 0.$$

By a similar argument as in HW2, Problem 3(a), we can conclude that $h = 0$ almost everywhere on $(0, \infty)$, whence $X_{(n)}$ is complete.
Remark: Alternatively, note that the map

$$\theta \mapsto \int_0^\theta h(x)\, x^{n-1}\, dx$$

is differentiable by the Fundamental Theorem of Calculus, and has derivative $h(\theta)\,\theta^{n-1}$ at each $\theta \in (0, \infty)$. Since this is a zero map, its derivative is also zero, and the claim follows.

Now, note that

$$\mathbb{E}[X_1] = \int_0^\theta \frac{x}{\theta}\, dx = \frac{\theta}{2},$$

so $2X_1$ is unbiased for θ. The Lehmann-Scheffe Theorem[2] now guarantees that $2\delta_2(X) = \mathbb{E}[2X_1 \mid X_{(n)}]$ is the *unique* UMVUE for θ. On the other hand, $(n+1)X_{(n)}/n$ is an unbiased function of the complete sufficient statistic $X_{(n)}$, and hence must be UMVUE too. This means that we must have equality, hence

$$\delta_2(X) = \mathbb{E}_\theta[X_1 \mid X_{(n)}] = \left(\frac{n+1}{2n}\right) X_{(n)}.$$

[2] See HW2, Problem 5.

(b) We have $L(\theta,\delta) = \sqrt{|\theta-\delta|}$. First we claim that L is concave in δ for each θ (note that we need only look at δ such that $\delta \le \theta$ almost surely, since our estimators fall within this category), from which the result in Problem 1 immediately tells us that

$$R(\theta,\delta_1) \le R(\theta,\delta_2), \qquad \text{for all } \theta \in (0,\infty).$$

This is because $T = X_{(n)}$ is sufficient for $\theta \in (0,\infty)$, and $\ell = -L$ is convex; the proof supplied in Problem 1 proceeds exactly as given, with $\delta = \delta_1, \eta = \delta_2$. We will then show that for $\theta = 1$, we have $R(1,\delta_1) < R(1,\delta_2)$, rendering δ_2 inadmissible.
Remark: Since $L(\theta,\delta) = \sqrt{\theta}L(1,\delta/\theta)$, we see that δ_1 has uniformly *strictly* lower risk than δ_2.
Indeed, for fixed $\theta > 0$, we need only show that the map $x \mapsto \sqrt{x}$ is concave for $x \ge 0$, which is clear by the differentiability criterion; we have the second derivative $-x^{-3/2}/4 < 0$. This further tells us that our loss is *strictly* concave in δ, so Jensen's inequality gives

$$\mathbb{E}[L(1,\delta_1) \mid X_{(n)}] < L(1,\delta_2).$$

Thus, taking expectations gives us a strict inequality, whence $R(1,\delta_1) < R(1,\delta_2)$.

6. Let $\mathcal{S}_2$ be the unit ball in $\mathbb{R}^2$, and let $(x_0,y_0) \in \mathcal{S}_2$. The process described is equivalent to the following: generate for $n \in \mathbb{N}$,

   (i) $y_n \sim U(-\sqrt{1-x_{n-1}^2}, \sqrt{1-x_{n-1}^2})$.

   (ii) $x_n \sim U(-\sqrt{1-y_n^2}, \sqrt{1-y_n^2})$.

   Remark: Drawing $x \sim U(-a,a)$ is equivalent to independently drawing $x' \sim U(0,a)$, $r \sim \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$ and setting $x = rx'$. Also, $x^2 = x'^2$.

   We claim that the stationary distribution π of this chain is uniform on $\mathcal{S}_2$, with density

$$f(x,y) = \frac{1}{\pi}\mathbf{1}_{\mathcal{S}_2}(x,y) = \frac{1}{\pi}\mathbf{1}\left(x_0^2+y_0^2 \le 1\right).$$

   We can describe the transition kernels (conditional densities)

$$\begin{aligned}
K((x_0,y_0),(x_0,y_1)) &= \frac{1}{2\sqrt{1-x_0^2}}\mathbf{1}\left(-\sqrt{1-x_0^2} \le y_1 \le \sqrt{1-x_0^2}\right) \\
&= \frac{1}{2\sqrt{1-x_0^2}}\mathbf{1}\left(x_0^2+y_1^2 \le 1\right), \\
K((x_0,y_1),(x_1,y_1)) &= \frac{1}{2\sqrt{1-y_1^2}}\mathbf{1}\left(-\sqrt{1-y_1^2} \le x_1 \le \sqrt{1-y_1^2}\right) \\
&= \frac{1}{2\sqrt{1-y_1^2}}\mathbf{1}\left(x_1^2+y_1^2 \le 1\right).
\end{aligned}$$

   With this, we have the transition kernel for the chain

$$K((x_0,y_0),(x_1,y_1)) = K((x_0,y_0),(x_0,y_1)) \cdot K((x_0,y_1),(x_1,y_1)).$$

   Compute

$$\begin{aligned}
&\int f(x_0,y_0)K((x_0,y_0),(x_0,y_1))\,dy_0 \\
&\qquad = \frac{1}{\pi}\int_{-\sqrt{1-x_0^2}}^{\sqrt{1-x_0^2}} \frac{1}{2\sqrt{1-x_0^2}}\mathbf{1}\left(x_0^2+y_1^2 \le 1\right)\,dy_0 \\
&\qquad = \frac{1}{\pi}\mathbf{1}\left(x_0^2+y_1^2 \le 1\right),
\end{aligned}$$

so the density of (x_1, y_1) is given by

$$
\begin{aligned}
\tilde{f}(x_1, y_1) &= \int f(x_0, y_0) K((x_0, y_0), (x_1, y_1)) \, dy_0 \, dx_0 \\
&= \int \left[\int f(x_0, y_0) K((x_0, y_0), (x_0, y_1)) \, dy_0 \right] K((x_0, y_1), (x_1, y_1)) \, dx_0 \\
&= \frac{1}{\pi} \int_{-\sqrt{1-y_1^2}}^{\sqrt{1-y_1^2}} \frac{1}{2\sqrt{1-y_1^2}} \mathbf{1}\left(x_1^2 + y_1^2 \le 1\right) \, dx_0 \\
&= \frac{1}{\pi} \mathbf{1}\left(x_1^2 + y_1^2 \le 1\right).
\end{aligned}
$$

Thus, $\tilde{f} = f$, whence f is the stationary distribution of this chain.

Remark: This is a Gibbs sampling procedure!

7. Let $\Theta \subseteq \mathbb{R}^k$ be open, and let $R(\theta, \delta)$ be continuous in θ for every estimator $\delta(X)$. Suppose that π_n is a sequence of priors on Θ, and let δ^* satisfy

   (i) $r(\pi_n, \delta^*) < \infty$ for all $n \in \mathbb{N}$.

   (ii) For any open $\Theta_0 \subseteq \Theta$,
   $$\liminf_{n \to \infty} \int_{\Theta_0} \pi_n(\theta) \, d\theta > 0.$$

   (iii) If δ_{π_n} are Bayes estimators with respect to each prior π_n, then
   $$\lim_{n \to \infty} r(\pi_n, \delta^*) - r(\pi_n, \delta_{\pi_n}) = 0.$$

   Suppose that δ is an estimator such that
   $$R(\theta, \delta) \le R(\theta, \delta^*), \qquad \text{for all } \theta \in \Theta.$$
   Further suppose that there exists $\theta_0 \in \Theta$ such that $R(\theta_0, \delta) < R(\theta_0, \delta^*)$. Set $2\epsilon = R(\theta_0, \delta^*) - R(\theta_0, \delta) > 0$. Then, by continuity of $R(\,\cdot\,, \delta^*) - R(\,\cdot\,, \delta)$, there exists an open neighborhood of $\Theta_0 \subseteq \Theta$ of θ_0 such that
   $$R(\theta, \delta^*) - R(\theta, \delta) \ge \epsilon, \qquad \text{for all } \theta \in \Theta_0.$$
   Using (ii), we can descend to a subsequence of priors π_{n_k} for which $\lim_{n\to\infty} \int_{\Theta_0} \pi_{n_k} = 2C > 0$, hence to a further sub-subsequence $\pi_{m_{n_k}}$ where each $\int_{\Theta_0} \pi_{m_{n_k}} > C > 0$. Without loss of generality, let this sub-subsequence of priors be simply π_n (by relabeling). Then, for all $n \in \mathbb{N}$, we have
   $$
   \begin{aligned}
   r(\pi_n, \delta^*) - r(\pi_n, \delta) &= \mathbb{E}_{\theta \sim \pi_n}\left[R(\theta, \delta^*) - R(\theta, \delta)\right] \\
   &\ge \mathbb{E}_{\theta \sim \pi_n}\left[\left(R(\theta, \delta^*) - R(\theta, \delta)\right) \mathbf{1}_{\Theta_0}(\theta)\right] \\
   &\ge \mathbb{E}_{\theta \sim \pi_n}\left[\epsilon \mathbf{1}_{\Theta_0}(\theta)\right] \\
   &\ge C\epsilon \\
   &> 0.
   \end{aligned}
   $$
   But each $r(\pi_n, \delta_{\pi_n}) \le r(\pi_n, \delta)$ by construction of the (potentially generalized) Bayes estimators[3]. Thus, for all $n \in \mathbb{N}$, we have
   $$r(\pi_n, \delta^*) - r(\pi_n, \delta_{\pi_n}) \ge r(\pi_n, \delta^*) - r(\pi_n, \delta) \ge C\epsilon > 0.$$
   Taking limits,
   $$\lim_{n \to \infty} r(\pi_n, \delta^*) - r(\pi_n, \delta_{\pi_n}) \ge C\epsilon > 0.$$
   This contradicts (iii). Thus, no such δ can exist, whence δ^* is admissible.

[3] These exist with finite Bayes risk using (i) and the construction in HW3, Problem 6.