Solutions to the Exercises of Section 3.7.

3.7.1. The spelling of Dvoretzky should be corrected (twice). The density of X is

$$f_X(x|\theta) = \begin{cases} f_Y(x|\theta) & \text{if } |x| < 1 \\ f_Y(x|\theta) + f_Y(-x|\theta) & \text{if } x > 1 \end{cases} \qquad = \frac{1}{\sqrt{2\pi}} e^{-(x^2+\theta^2)/2} \begin{cases} e^{\theta x} & \text{if } |x| < 1 \\ e^{\theta x} + e^{-\theta x} & \text{if } x > 1. \end{cases}$$

This is not an exponential family so the words "the distribution of X forms an exponential family; hence" should be deleted from the statement of this exercise. For an arbitrary decision rule $\delta(x) = \mathrm{P}(a_1|X = x)$, the Bayes risk with respect to the uniform distribution τ is

$$\begin{aligned} r(\tau, \delta) &= \frac{1}{2} \int_{-2}^{-1} (1 - \mathrm{E}_\theta \delta(X))\, d\theta + \frac{1}{2} \int_1^2 \mathrm{E}_\theta \delta(X)\, d\theta \\ &= \frac{1}{2} + \frac{1}{2} \left[\int_1^2 \mathrm{E}_\theta \delta(X)\, d\theta - \int_{-2}^{-1} \mathrm{E}_\theta \delta(X)\, d\theta \right] \\ &= \frac{1}{2} + \frac{1}{2} \int_{-1}^{\infty} \delta(x) \left[\int_1^2 f_X(x|\theta)\, d\theta - \int_{-2}^{-1} f_X(x|\theta)\, d\theta \right] dx \end{aligned}$$

Any Bayes rule, $\delta^0(x)$, must have value 1 if the term in the square brackets is negative and 0 if it is positive. It is immaterial what value $\delta^0(x)$ assumes if the term in square brackets is zero. From this we may conclude that the rule

$$\delta^0(x) = \begin{cases} 1 & \text{if } -1 < x < 0 \\ 1/2 & \text{if } x > 1 \\ 0 & \text{if } 0 < x < 1 \end{cases}$$

is a Bayes rule with respect to τ. (This corrects the misprint in writing $\delta_x^0(a_1)$ in the text.) Since the distribution of Y forms an exponential family, we have that the risk function of any rule $\delta(x)$ is continuous in θ since $\delta(x)$ may also be considered as a function of y. The admissibility of δ^0 now follows from the proof of Theorem 2.3.3 (but not quite from the statement, where it was assumed that the parameter space was the whole real line).

We now note that X is a complete sufficient statistic for the distribution of X given θ: If $\mathrm{E}_\theta g(X) = 0$ for all θ, then $\mathrm{E}_\theta g(h(Y)) = 0$ for all θ, where h is the function that maps Y into X. But since Y has an exponential family of distributions, Y is a complete sufficient statistic for θ, so that $\mathrm{P}_\theta(g(h(Y)) = 0) = 1$. Thus we have $\mathrm{P}_\theta(g(X) = 0) = 1$, completing the proof.

From this, we may conclude that if δ is any rule that is as good as δ^0, then δ has the same risk function as δ^0 since the latter is admissible. But then from completeness, $\mathrm{E}_\theta \delta(X) - \mathrm{E}\theta\delta^0(X) = 0$ for all θ implies that $\delta(X) - \delta^0(X) = 0$ with probability one. Thus δ must be randomized, and no nonrandomized rule can be as good as δ^0.

3.7.2. The method used for the proof with squared error loss works with minor changes for absolute error loss. Again, we may take $n = 1$ and $T \in \mathcal{N}(\theta, 1)$ without loss of generality. Since absolute error loss, $L(\theta, a) = |\theta - a|$, is convex in a for every θ, we may restrict attention to nonrandomized decision rules, d. Theorem 3.7.2 implies that $R(\theta, d)$ is continuous in θ; condition (a) is satisfied with $B_1 = 1$ and $B_2(\theta_1, \theta_2) = |\theta_1 - \theta_2|$.

Suppose $d'(t) = t$ is not admissible. Then there exists a rule $d''(t)$ such that $R(\theta, d'') \le R(\theta, d')$ for all θ and $R(\theta_0, d'') < R(\theta_0, d')$ for some θ_0. Since $R(\theta, d')$ and $R(\theta, d'')$ are continuous in θ, there exists an $\epsilon > 0$ such that $R(\theta, d'') < R(\theta, d') - \epsilon$ for all θ such that $|\theta - \theta_0| < \epsilon$.

If the prior distribution is $\tau_\sigma = \mathcal{N}(0, \sigma^2)$, then the posterior distribution is $\mathcal{N}(t\sigma^2/(1+\sigma^2), \sigma^2/(1+\sigma^2))$. The Bayes rule is the median of this distribution, namely $d_\sigma(t) = t\sigma^2/(1 + \sigma^2)$, and the Bayes risk is $r(\tau_\sigma, d_\sigma) = c\sqrt{\sigma^2/(1 + \sigma^2)}$, where $c = \sqrt{2/\pi}$. Thus,

$$r(\tau_\sigma, d') - r(\tau_\sigma, d_\sigma) = c \left(1 - \sqrt{\frac{\sigma^2}{1 + \sigma^2}} \right).$$

However,

$$\sigma[r(\tau_\sigma, d'') - r(\tau_\sigma, d')] < -\frac{\epsilon}{2\pi} \int_{\theta_0-\epsilon}^{\theta_0+\epsilon} \exp\{-\frac{\theta^2}{2\sigma^2}\} \, d\theta$$

so that

$$0 \leq -\frac{\epsilon}{\sqrt{2\pi}} \int_{\theta_0-\epsilon}^{\theta_0+\epsilon} \exp\{-\frac{\theta^2}{2\sigma^2}\} \, d\theta + c\sigma \left(1 - \sqrt{\frac{\sigma^2}{1+\sigma^2}}\right).$$

The second term still converges to zero as $\sigma^2 \to \infty$ while the first term converges to $-2\epsilon^2/\sqrt{2\pi} < 0$, yielding a contradiction that completes the proof.