Solutions to Exercises 4.5.1 through 4.5.7, and 4.5.10.

4.5.1. The median, θ, of the Cauchy distribution is a location parameter, and the loss, $L(\theta, a) = L(a - \theta)$, is a function of $a - \theta$. The risk function of an invariant rule, $d_b(x) = x - b$, is the constant, $R(\theta, d_b) = R(0, d_b) = \mathrm{P}_0(|X - b| > c)$. By Theorem 4.5.1, the best invariant rule is d_b where b minimizes this quantity, or equivalently, maximizes, $\mathrm{P}_0(|X - b| \le c)$. But since the Cauchy distribution is unimodal and symmetric about 0, $\mathrm{P}_0(|X - b| \le c)$ is maximized by $b = 0$, i.e. $d(x) = x$ is the best invariant rule. It does not depend on β or c (or the distribution of $X - \theta$ provided it is unimodal and symmetric about 0).

4.5.2. By Theorem 1, the best invariant estimate of θ is $d(X) = X - b_0$, where b_0 is that value of b that minimizes $\mathrm{E}_0(X - b)^2$, namely $b_0 = \mathrm{E}_0 X$. We now note that this estimate is unbiased, $\mathrm{E}_\theta d(X) = \mathrm{E}_\theta X - b_0 = \theta + \mathrm{E}_\theta(X - \theta) - b_0 = \theta + \mathrm{E}_0 X - \mathrm{E}_0 X = \theta$. But any function of a complete sufficient statistic is a best unbiased estimate of its expectation (if it has finite risk). Thus $d(X)$ is a best unbiased estimate of θ as well.

4.5.3. We are given that $\theta > 0$ is a scale parameter, $f(x|\theta) = (1/\theta) f(x/\theta)$, and that the loss is a function of $a - \theta$, say $L(\theta, a) = L(a/\theta)$. Such a problem is invariant under the group of scale changes, $g_c(x) = cx$ for $c > 0$, with $\bar{g}_c(\theta) = c\theta$ and $\tilde{g}_c(a) = ca$. An invariant nonrandomized estimate of θ then satisfies $d(g_c(x)) = \tilde{g}_c(d(x))$ for all $c > 0$ and all x. For x=1, this implies that $d(c) = c\,d(1)$. Hence the class of invariant rules are the rules of the form $d_b(x) = x/b$ for some b (here equal to $1/d(1)$). From Theorem 4.2.1, the risk function of any invariant rule $\delta \in \mathcal{D}$ satisfies $R(\theta, \delta) = R(c\theta, \delta)$ for all c and θ, so that the risk of an invariant rule is a constant independent of θ. By Lemma 4.5.1 and by the fact that the orbits of $\mathcal{X}$ have multiplicity one, we may restrict attention to nonrandomized rules. The risk of a nonrandomized invariant rule is $R(\theta, d_b) = \mathrm{E}(L(X/b)|\theta = 1)$ since it is independent of θ, and the best invariant rule is the one that minimize this. So, we have

In the problem of estimating a scale parameter with loss $L(a, \theta) = L(a/\theta)$, if $\mathrm{E}(L(X/b)|\theta = 1)$ exists and is finite for some b, and if there exists a b_0 such that

$$\mathrm{E}(L(X/b_0)|\theta = 1) = \inf_b \mathrm{E}(L(X/b)|\theta = 1),$$

where the infimum is taken over all b for which $\mathrm{E}(L(X/b)|\theta = 1)$ exists, then $d(x) = X/b_0$ is a best invariant rule.

4.5.4. If δ is invariant, then by Theorem 4.2.1, $R(\theta, \delta)$ is constant on orbits of $\bar{\mathcal{G}}$. If $\bar{\mathcal{G}}$ is transitive, there is only one orbit because every point of Θ can be carried into every other point by some $\bar{g} \in \bar{\mathcal{G}}$. Thus $R(\theta, \delta)$ is constant which shows that δ is an equalizer rule.

4.5.5. If $\mathbf{X}_1, \ldots, \mathbf{X}_n$ is a sample from the k-dimensional normal distribution, $\mathcal{N}(\theta, \mathbf{I})$, then $\overline{\mathbf{X}}_n$ is sufficient for θ and has a $\mathcal{N}(\theta, (1/n)\mathbf{I})$ distribution. So if we can show $\overline{\mathbf{X}}_n$ is minimax for quadratic loss, $L(\theta, \mathbf{a}) = \sum_{i=1}^k (\theta_i - a_i)^2$, for $n = 1$ it will be minimax for arbitrary n as well. So take $n = 1$ and let $d(\mathbf{X}) = \mathbf{X}$. The risk function of d is constant, $R(\theta, d) = \mathrm{E}_\theta L(\theta, \mathbf{X}) = \mathrm{E}_\theta \sum_{i=1}^k (X_i - \theta_i)^2 = k$, so it will be minimax if it is extended Bayes. Take as the prior distribution τ_σ of θ the $\mathcal{N}(0, \sigma^2\mathbf{I})$ distribution. The posterior distribution of θ given $\mathbf{X}$ is then $\mathcal{N}(\mathbf{X}\sigma^2/(1+\sigma^2), (\sigma^2/(1+\sigma^2))\mathbf{I})$, so that the Bayes rule with respect to τ_σ is $d_\sigma(\mathbf{X}) = \mathbf{X}\sigma^2/(1+\sigma^2)$ and it has Bayes risk

$$r(\tau_\sigma, d_\sigma) = \mathrm{E}(\mathrm{E}\{\|\theta - d_\sigma(X)\|\,|\mathbf{X}\}) = \frac{k\sigma^2}{1+\sigma^2}$$

Now, since $r(\tau_\sigma, d_\sigma) \to k = R(\theta, d)$ as $\sigma \to \infty$, we have that d is extended Bayes and therefore minimax.

4.5.6. There is a misprint in the definition of the rule d_b. It should read

$$d_b(x) = (x+1)\mathrm{I}_{(-\infty,b)}(x) + (x-1)\mathrm{I}_{[b,\infty)}(x).$$

By the methods of the Example 1, we have

$$R(\theta, d_b) = \begin{cases} 0 & \text{if } b-1 \le \theta < b+1 \\ 1/2 & \text{otherwise.} \end{cases} \quad = \frac{1}{2} - \frac{1}{2}\mathrm{I}(\theta - 1 < b \le \theta + 1).$$

The risk of the randomized rule, δ, that chooses b according to a strictly increasing distribution function, $F(b)$, is

$$\begin{aligned} R(\theta,\delta) &= \int_{-\infty}^{\infty} R(\theta,d_b)\,dF(b) \\ &= \frac{1}{2} - \frac{1}{2}\mathrm{P}(\theta-1 < b \le \theta+1) < \frac{1}{2} \end{aligned}$$

A best invariant rule has constant risk $1/2$, so this is better than a best invariant rule at all θ.

4.5.7. (a) Invariant rules are of the form $d_b(x) = x - b$ for some number b. The risk function of d_b is

$$\begin{aligned} R(\theta,d_b) &= \mathrm{E}_\theta L(\theta, X-b) = \mathrm{E}_0 L(0,X-b) = \mathrm{E}_0(X-b)^+ \\ &= \sum_{\substack{x\ge 1\\ x>b}} (x-b)\frac{1}{x(x+1)} \ge \sum_{\substack{x\ge 1\\ x>2b}} \frac{1}{2}\,\frac{1}{x+1} = +\infty \end{aligned}$$

for all θ and b.

(b) For the noninvariant rule, $d(x) = x - c|x|$, with $c > 1$, the risk is $R(\theta,d) = \mathrm{E}_\theta(X - c|X| - \theta)^+ = \mathrm{E}(Y - c|Y+\theta|)^+$ where $Y = X - \theta$ has mass function independent of θ, $f(y) = 1/(y(y+1))$ for $y = 1,2,\ldots$. If $\theta \ge 0$, then $Y - c|Y+\theta| \le 0$ w.p. 1 since $c > 1$. Thus $R(\theta,d) = 0$ for $\theta \ge 0$. If $\theta < 0$, then $y - c|y+\theta| > 0$ if and only if $a < y < b$ where a and b are the roots of $y = c|y+\theta|$, namely, $a = |\theta|/(c+1)$ and $b = |\theta|/(c-1)$. Hence, for $\theta < 0$,

$$\begin{aligned} R(\theta,d) &= \sum_{a<y<b} \frac{(y - c|y+\theta|)}{y+1} \le \sum_{a<y<b} \frac{1}{y+1} \\ &\le \frac{1}{2} + \sum_{a<y<b-1} \frac{1}{y+1} = \frac{1}{2} + \sum_{a+1<y<b} \frac{1}{y} \\ &\le \frac{1}{2} + \int_a^b \frac{1}{y} = \frac{1}{2} + \log\frac{b}{a} \\ &= \frac{1}{2} + \log\frac{c+1}{c-1} = \log e^{1/2}\frac{c+1}{c-1} \\ &\le \log 2\frac{c+1}{c-1} \end{aligned}$$

for all θ.

(c) If $L(x) = 1$ when x is an integer and $L(x) = \max(x,0)$ otherwise, then the invariant rule $d(x) = x$ has constant risk equal to one; yet if c is irrational, the risk of the rule $d(x) = x - c|x|$ has risk bounded by $\log 2(c+1)/(c-1)$, which if c is sufficiently large is less than one.

4.5.10. Suppose X,Y has joint density $f_{X,Y}(x,y|\theta_1,\theta_2) = f(x-\theta_1, y-\theta_2)$ with finite second moments, where the parameter space, Θ is the whole Euclidean plane.

(a) If we put $g_{b,c}(X,Y) = (X+b, Y+c) = (S,T)$, then the joint density of S,T is $f_{S,T}(s,t|\theta_1,\theta_2) = f_{X,Y}(s-b, t-c|\theta_1,\theta_2) = f(s-(b+\theta_1), t-(c+\theta_2))$, so that the distributions are invariant with $\bar{g}_{b,c}(\theta_1,\theta_2) = (b+\theta_1, c+\theta_2)$. Moreover, for the given loss function we have

$$L(\bar{g}_{b,c}(\theta_1,\theta_2), \tilde{g}_{b,c}a) = \left(\frac{b+\theta_1+c+\theta_2}{2} - \tilde{g}_{b,c}a\right)^2 = \left(\frac{\theta_1+\theta_2}{2} - a\right)^2$$

provided $\tilde{g}_{b,c}a = a + ((b+c)/2)$. This shows that the loss, and hence the problem, is invariant.

(b) The group $\overline{\mathcal{G}} = \{\bar{g}_{b,c}\}$ is transitive on Θ, (so that the risk function of an invariant rule is constant) and the single orbit in $\mathcal{X}$ has multiplicity one (so that from Lemma 4.5.1 and the discussion, we may restrict attention to nonrandomized rules in our search for a best invariant rule.

(c) An invariant rule satisfies $d(g_{b,c}(x,y)) = \tilde{g}_{b,c}(d(x,y))$ which means $d(x+b, y+c) = d(x,y) + (b+c)/2$ for all x,y,b,c. Put $x = y = 0$ and find $d(b,c) = (b+c)/2 - d(0,0)$ for all b,c, so that the nonrandomized invariant rules have the form $d_\alpha(x,y) = (x+y)/2 - \alpha$ for some α (here $\alpha = -d_\alpha(0,0)$).

(d) The risk function of d_α is the constant $R((\theta_1,\theta_2), d_\alpha) = \mathrm{E}_{(0,0)}((1/2)(X+Y) - \alpha)^2$ which is minimized by $\alpha_0 = \mathrm{E}_{(0,0)}((X+Y)/2)$. Therefore, the best invariant rule is $(X+Y)/2 - \mathrm{E}_{(0,0)}((X+Y)/2)$.