Solutions to Exercises 4.6.1 to 4.6.7, and 4.6.9 to 4.6.12.

4.6.1. Sufficiency reduces the problem to choosing rules based on $T = \min\{X_1, \ldots, X_n\}$. This has distribution with density, $f(t|\theta) = n\exp\{-nt\}I(t > \theta)$ for which θ is a location parameter.

(a) $\mathrm{E}_0(T-b)^2$ is minimized by $b = \mathrm{E}_0 T = 1/n$. So the best invariant estimate is $d(T) = T - 1/n$.

(b) $\mathrm{E}_0|T-b|$ is minimized if $b =$ the median of the distribution of T when $\theta = 0$. Since $1/2 = \int_b^\infty n\exp\{-nt\}\,dt = \exp\{-bn\}$ gives $b = \log(2)/n$ as the median. Hence, $d(T) = T - \log(2)/n$ is the best invariant estimate.

(c) The best invariant estimate of θ is $d(T) = T - b$ where b is the midpoint of the modal interval of length $2c$ of the distribution of T when $\theta = 0$. Since this density is decreasing from zero to infinity, the modal interval of length $2c$ is the interval $(0, 2c)$. This gives $d(T) = T - c$ as the best invariant estimate.

4.6.2. $Y = \min_{1\le i\le n} X_i$ is a sufficient statistic for θ, and $\mathrm{P}(Y \ge \theta + k) = \mathrm{P}(X_1 \ge \theta + k)^n = p^{kn}$ so that $Y - \theta$ has a geometric distribution with parameter p^n.

(a) The best invariant estimate of θ is $Y - b$ where b is chosen to minimize $\mathrm{E}_0(Y-b)^2$. This occurs if $b = \mathrm{E}_0 Y = p^n/(1-p^n)$.

(b) In this case, the best invariant estimate of θ is $Y - b$ where b minimizes $\mathrm{P}_0(Y \ne b)$ or, equivalently, maximizes $\mathrm{P}_0(Y = b)$; thus, b is the mode of the distribution of Y when $\theta = 0$. This clearly gives $b = 0$. Here, the best invariant estimate of θ is the maximum likelihood estimate.

4.6.3. We may reduce the problem by sufficiency to consideration of estimates that are functions of $S_n = X_1 + \cdots + X_n$ which has the gamma distribution $\mathcal{G}(n\alpha, \beta)$. Thus, we solve the problem for a single observation, X; the solution for the general problem is found by replacing X by S_n and α by $n\alpha$.

(a) The problem is invariant under change of scale and the best invariant estimate is $d(X) = X/b$ where b minimizes $\mathrm{E}_{\beta=1}(X/b - 1)^2$. This leads to $b = \mathrm{E}X^2/\mathrm{E}X = (\alpha + \alpha^2)/\alpha = \alpha + 1$.

(b) To find b that minimizes $\mathrm{E}((X/b) - 1 - \log(X/b)|\beta = 1)$, we take a derivative with respect to b, set to zero and find $b = \mathrm{E}X = \alpha$. The best invariant estimate is $d(X) = X/\alpha$.

4.6.4. The minimum, $Z = \min X_i$, is a sufficient statistic for θ and has density

$$f_Z(z|\theta) = \frac{n\theta^n}{z^{n+1}}\mathrm{I}_{(\theta,\infty)}(z),$$

with θ as a scale parameter. In (a), (b), and (c), the loss is of the form $L(\theta, a) = L(a/\theta)$, so these problems are invariant under change of scale and the best invariant rule is $d(z) = z/b_0$, where b_0 minimizes $\mathrm{E}_{\theta=1}L(Z/b)$.

(a) For $L(\theta, a) = ((a/\theta) - 1)^2$, we have $b_0 = \mathrm{E}_1 Z^2/\mathrm{E}_1 Z = (n/(n-2))/(n/(n-1) = (n-1)/(n-2)$. So the best invariant rule is $d(z) = (n-2)Z/(n-1)$.

(b) $\mathrm{E}_1(|\log Z - \log b|)$ is minimized by taking $\log b$ to be the median of the distribution of $\log Z$ when $\theta = 1$, which is the same as taking b to be the median of Z when $\theta = 1$. This is easily computed to be $b = 2^{1/n}$. The best invariant rule is $d(z) = z/2^{1/n}$.

(c) Minimizing $\mathrm{E}_1|(Z/b) - 1| = \mathrm{E}_1 Z|(1/b) - (1/Z)| = \int_1^\infty |(1/b) - (1/z)|zf(z|1)\,dz$ is equivalent to minimizing $\mathrm{E}|(1/b) - (1/Z)|$ where Z has density proportional to $zf(z|1)$, which is the original density of Z with n replaced by $n-1$. This is done by choosing $1/b$ to be the median of the distribution of $1/Z$ with this density, or equivalently, choosing b as the median of the distribution of Z with this density. This median is $2^{1/(n-1)}$ so the best invariant estimate is $d(z) = z/2^{1/(n-1)}$.

4.6.5. $Z = \max\{X_1, \ldots, X_n\}$ is sufficient for θ and has density $f(z|\theta) = nz^{n-1}/\theta^n$ for $0 \le z \le \theta$. The problems are scale invariant and invariant rules are of the form $d(z) = z/b$ for some $b > 0$.

(a) $\mathrm{E}_1((Z/b) - 1)^2$ is minimized for $b = \mathrm{E}_1 Z^2/\mathrm{E}_1 Z = (n+1)/(n+2)$. So the best invariant rule is $d(z) = (n+2)z/(n+1)$.

(b) We must find b to minimize $\mathrm{E}_1 L(1, Z/b) = 1 - \mathrm{P}_1\{(1/c) \le (Z/b) \le c\} = (b/c)^n + 1 - (bc)^n$ for $b < 1/c$ and$= (b/c)^n$ for $b \ge 1/c$. The minimum clearly occurs at $b = 1/c$, so $d(z) = cz$ is the best invariant rule.

(c) As in problem 4.6.4(c), the optimal b is the median of the distribution with density proportional to $zf(z|1)$. This gives $b = 2^{-1/(n+1)}$ and $d(z) = 2^{1/(n+1)}z$ as the best invariant estimate.

4.6.6. Sufficiency reduces the problem to considering $T^2 = \sum_1^n X_i^2$. Since the distribution of T^2/σ^2 is χ_n^2, independent of σ^2, the problem is scale invariant under the loss $L(\sigma, a) = (\sigma - a)^2/\sigma^2$, and invariant rules are of the form $d(T) = T/b$. As in problem 4.6.4(a), the best invariant rule is $d(T) = T/b_0$, where $b_0 = \mathrm{E}_1 T^2/\mathrm{E}_1 T$. We have $\mathrm{E}_1 T^2 = n$ and

$$\mathrm{E}_1 T = \int_0^\infty z^{1/2} \frac{1}{\Gamma(n/2)2^{n/2}} e^{-z/2} z^{(n/2)-1}\, dz = \frac{\Gamma((n+1)/2)\sqrt{2}}{\Gamma(n/2)}.$$

So $b_0 = n\Gamma(n/2)/(\sqrt{2}\Gamma((n+1)/2))$. For estimating σ^2 with loss $(a/\sigma^2 - 1)^2$, the best invariant estimate is $T^2/(n+2)$. For estimating σ^2 with loss $(\log a - \log\sigma^2)^2$, the best invariant estimate is $T^2/(2\exp[\Psi(n/2)]) \sim T^2/(n-1)$. The estimate of σ^2 corresponding to the above is $2T^2\Gamma((n+1)/2)^2/(n\Gamma(n/2))^2 \sim T^2/(n+.5)$.

4.6.7. (a) The sufficient statistics are independent, with $\overline{X}$ having a normal distribution, $\mathcal{N}(\mu, \sigma^2/n)$ and ns^2/σ^2 having a χ_{n-1}^2 distribution. If $(U, V) = g_{b,c}(\overline{X}, s) = (b\overline{X} + c, bs)$, then U and V are still independent with U having the distribution $\mathcal{N}(b\mu + c, b^2\sigma^2)$ and $nV^2/(b^2\sigma^2)$ having a χ_{n-1}^2 distribution. Thus the distributions are invariant with $\bar{g}_{b,c}(\mu, \sigma) = (b\mu + c, b\sigma)$. Moreover, for the given loss function,

$$L(\bar{g}_{b,c}(\mu,\sigma), \tilde{g}_{b,c}a) = \frac{(v(b\mu + c) + wb\sigma - \tilde{g}_{b,c}a)^2}{b^2\sigma^2} = \frac{(v\mu + w\sigma - a)^2}{\sigma^2}$$

provided $\tilde{g}_{b,c}a = ba + vc$, proving the loss and hence the problem is invariant.

(b) A nonrandomized invariant rule satisfies $d(b\overline{X}+c, bs) = bd(\overline{X}, s) + vs$ for all $\overline{X}, s, b, c$. Set $\overline{X} = 0$ and $s = 1$ and find $d(c, b) = bd(0,1) + v$ for all b, c, so that nonrandomized rules have the form $d_k(\overline{X}, s) = v\overline{X} + ks$ for some constant k.

(c) There is just one orbit in Θ, so nonrandomized rules have constant risk

$$\begin{aligned} R((\mu,\sigma), d_k) &= \mathrm{E}_{(0,1)} L((0,1), v\overline{X} + ks) = \mathrm{E}(0,1)(w - v\overline{X} - ks)^2 \\ &= w^2 - 2wv\mathrm{E}_{(0,1)}\overline{X} + v^2\mathrm{E}_{(0,1)}\overline{X}^2 - 2wk\mathrm{E}_{(0,1)}s + 2vk\mathrm{E}_{(0,1)}\overline{X}s + k^2\mathrm{E}_{(0,1)}s^2 \\ &= w^2 + v^2/n - 2wk\mathrm{E}_{(0,1)}s + k^2(n-1)/n \end{aligned}$$

This is quadratic in k with a minimum at $k = nw\mathrm{E}_{(0,1)}s/(n-1)$, and the minimum risk is $v^2/n + w^2(1 - n(\mathrm{E}_{(0,1)}s)^2/(n-1))$. It is easy to compute that $\mathrm{E}_{(0,1)}s = \sqrt{2}\Gamma(n/2)/\Gamma((n-1)/2)$.

4.6.9 (a) If $g_c(x) = x + c$, then the distributions are invariant with $\bar{g}_c(\theta) = \theta + c$. If $a = (y, z)$ with $y < z$, then the loss satisfies $L(\bar{g}_c(\theta), \tilde{g}_c(y,z)) = L(\theta, (y,z))$ provided $\tilde{g}_c(y,z) = (y+c, z+c)$. This shows the problem is invariant under location changes.

(b) Write a decision rule d as $d(x) = (d_1(x), d_2(x))$, with $d_1(x) \le d_2(x)$ for all x. Then d is invariant if $(d_1(x+c), d_2(x+c)) = (d_1(x), d_2(x)) + c$ for all x_1, x_2, c. Setting $x = 0$ gives $(d_1(c), d_2(c)) = (d_1(0), d_2(0)) + c$, and then replacing c by x shows that all nonrandomized rules have the form $(d_1(x), d_2(x)) = (x - b_1, x - b_2)$ for some $b_1 \ge b_2$.

(c) By Theorem 4.2.1, all invariant rules have have constant risk, and by the argument of Lemma 4.5.1, we may restrict attention to the nonrandomized invariant rules. The risk of an invariant rule is $\mathrm{E}_0 L(0, (X - b_1, X - b_2)) = k(b_1 - b_2) - \mathrm{P}_0(X - b_1 < 0 < X - b_2) = k(b_1 - b_2) - \mathrm{P}_0(b_2 < X < b_1)$. The derivative of this with respect to b_1 is $k - (2\pi)^{-1/2}\exp\{-b_1^2/2\}$. This is always positive if $k > 1/\sqrt{2\pi}$ and the best invariant rule in this case is to take $b_1 = b_2$, i.e. an empty interval. If $k < 1/\sqrt{2\pi}$, then the derivative is zero at $b_1 = \pm(-\log 2\pi k^2)^{1/2}$, with the plus sign producing the minimum risk. A symmetric analysis for b_2 leads to the following best invariant rule.

$$(d_1(x), d_2(x)) = \begin{cases} (0,0) & \text{if } k > 1/\sqrt{2\pi} \\ (x - b, x + b) & \text{if } k \le 1/\sqrt{2\pi} \end{cases}$$

where $b = (\log(1/2\pi k^2))^{1/2}$.

4.6.10. (a) The problem is invariant under change of location and scale as in Exercise 4.6.7 with $g_{b,c}(\overline{X}, s) = (b\overline{X} + c, bs)$ and $\bar{g}_{b,c}(\mu,\sigma) = (b\mu + c, \sigma)$, but here $\tilde{g}_{b,c}(y,z) = (by + c, bz + c)$.

(b) A rule $d(\overline{X}, s) = (d_1(\overline{X}, s), d_2(\overline{X}, s))$ is invariant if $d_i(b\overline{X} + c, bs) = bd_i(\overline{X}, s) + c$ for $i = 1, 2$. Put $\overline{X} = 0$ and $s = 1$ to find $d_i(c, b) = bd_i(1, 0) + c$ for $i = 1, 2$. Thus we find that invariant rules are of the form $d_{(a_1,a_2)}(\overline{X}, s) = (\overline{X} - a_1 s, \overline{X} - a_2 s)$ for some $a_1 \geq a_2$.

(c) The risk function is the constant $R((0, 1), d_{(a_1,a_2)}) = k(a_1 - a_2)\mathrm{E}_{(0,1)}s - \mathrm{P}_{(0,1)}(\overline{X} - a_1 s < 0 < \overline{X} - a_2 s) = k(a_1 - a_2)\mathrm{E}_{(0,1)}s - \mathrm{P}_{(0,1)}(\sqrt{n-1}\, a_2 < t < \sqrt{n-1} a_1)$, where $t = \sqrt{n-1}\overline{X}/s$ has a t-distribution with $n-1$ degrees of freedom. The derivative with respect to a_1 is $k\mathrm{E}_{(0,1)}s - f_{n-1}(\sqrt{n-1}a_1)\sqrt{n-1}$, where f_{n-1} is the density of the t_{n-1}-distribution. $\mathrm{E}_{(0,1)}s$ is found in Exercise 4.6.7, and the density of t_{n-1} in Section 3.1. This leads to

$$\frac{\partial R((0, 1), d_{(a_1,a_2)})}{\partial a_1} = \frac{k\sqrt{2}\Gamma(n/2)}{\Gamma((n-1)/2)} - \frac{\Gamma(n/2)\sqrt{n-1}}{\sqrt{(n-1)\pi}\Gamma((n-1)/2)(a_1^2 + 1)^{n/2}}$$

If $2\pi k^2 > 1$, the risk is minimized by choosing $a_1 = a_2$ i.e. the interval estimate is the empty set. If $2\pi k^2 \leq 1$, the optimal value of a_1 can be found by setting the above derivative to zero and solving for a_1. The corresponding point for a_2 is symmetric, and we find that the best invariant interval estimate is $(\overline{X} - as, \overline{X} + as)$, where (unlike the formula in the book) $a = ((2\pi k^2)^{-1/n} - 1)^{1/2}$.

4.6.11. (a) The distributions are invariant under the transformations $g_{b,c}(x_1, x_2) = (bx_1 + c, bx_2 + c)$ with $\bar{g}_{b,c}(\mu, \sigma) = (b\mu + c, b\sigma)$. The loss is invariant with $\tilde{g}_{b,c}a = ba + vc$.

(b) A nonrandomized rule d is invariant if $d(bx_1 + c, bx_2 + c) = bd(x_1, x_2) + vc$ for all $x_1, x_2, b > 0$ and c. Putting $x_1 = 0$ and $x_2 = 1$, we find $d(c, b + c) = bd(0, 1) + vc$ for all $b > 0$ and c. Replacing c by x_1 and b by $x_2 - x_1$, we find that every nonrandomized rule is of the form $d_\alpha(x_1, x_2) = \alpha(x_2 - x_1) + vx_1$ for some number α.

(c) The best invariant rule then is d_α where α minimizes $\mathrm{E}_{(0,1)}(W(w - vX_1 - \alpha(X_2 - X_1))^2$.

(d) If $W(z) = z^2$, then the α that minimizes the expectation in (c) is

$$\alpha_0 = \frac{\mathrm{E}_{(0,1)}((w - vX_1)(X_2 - X_1))}{\mathrm{E}_{(0,1)}(X_2 - X_1)^2}.$$

The best invariant rule is d_{α_0}.

4.6.12. The joint distribution of X_1 and X_2 is

$$f_{X_1,X_2}(x_1, x_2|\mu, \sigma) = n(n-1)\frac{1}{\sigma^2}[\frac{1}{\sigma}(x_2 - x_1)]^{n-2} \quad \text{for} \quad \mu < x_1 < x_2 < \mu + \sigma.$$

Thus, this density has the form required in Problem 4.6.11. We calculate the three expectations,

$$\mathrm{E}_{(0,1)}(X_2 - X_1) = \int_0^1 \int_0^{x_2} n(n-1)(x_2 - x_1)^{n-1}\, dx_1\, dx_2 = \frac{n-1}{n+1}.$$

$$\mathrm{E}_{(0,1)}(X_2 - X_1)^2 = \int_0^1 \int_0^{x_2} n(n-1)(x_2 - x_1)^{n}\, dx_1\, dx_2 = \frac{n(n-1)}{(n+1)(n+2)}.$$

$$\begin{aligned}\mathrm{E}_{(0,1)}(X_2 - X_1)(x_1X_2 - x_2X_1) &= x_1\mathrm{E}_{(0,1)}[X_2(X_2 - X_1)] - x_2\mathrm{E}_{(0,1)}[X_1(X_2 - X_1)] \\ &= x_1\frac{n-1}{n+2} - x_2\frac{n-1}{(n+1)(n+2)} = \frac{n-1}{(n+1)(n+2)}[(n+1)x_1 - x_2].\end{aligned}$$

Therefore, by Problem 11(d), the best invariant decision rule is

$$d(x_1, x_2) = w\frac{n+2}{n}(x_2 - x_1) + v[x_1 - \frac{1}{n}(x_2 - x_1)].$$