Solutions to Exercises 7.2.2 to 7.2.4, and 7.2.6 to 7.2.9.

7.2.2. At stage 0, the Bayes expected loss is $\int_0^1 (\theta - a)^2/(\theta(1-\theta))\,d\theta = \infty$ for all a, so that $\rho_0 = \infty$. At stage j, the posterior distribution of θ given $X_1 = -1, \ldots, X_j = -1$ is still $\mathcal{U}(0,1)$, so $\rho_j(-1,\ldots,-1) = \infty$ for all j. Hence $U_j(-1,\ldots,-1) = \infty$ for all j as well. For any given truncation integer, J, $V_J^{(J)}(-1,\ldots,-1) = \infty$, and by backward induction for any $0 \le j < J$, $V_j^{(J)}(-1,\ldots,-1) = (1/2)V_{j+1}^{(J)}(-1,\ldots,-1) + \cdots = \infty$. Thus, $V_0^{(J)} = \infty$ for all J.

Now consider the sequential decision rule: Stop sampling at the first N such that $X_N \neq -1$; if $X_N = 0$, estimate θ as 0, and if $X_N = 1$, estimate θ as 1. The expected terminal loss is the same if $X_N = 0$ or $X_N = 1$, namely $\int_0^1 (\theta^2/(\theta(1-\theta)))2(1-\theta)\,d\theta = 2\int_0^1 \theta\,d\theta = 1$. The expected sampling cost is $\mathrm{E}_\theta cN = 2c$. Since this rule has finite risk, $V_0^{(\infty)}$ is finite.

7.2.3. **Proof.** Let a_j denote the constant $U_j(x_1,\ldots,x_j)$. Then, $V_J^{(J)} = a_J$, and by backward induction, $V_j^{(J)} = \min\{a_j,\ldots,a_J\}$ for $0 \le j \le J$. Hence the Bayes rule for the problem truncated at J is the fixed sample size rule that stops at $m(J)$, where $a_{m(J)} = \min\{a_0, a_1, \ldots, a_J\}$, and we have $V_0^{(J)} = \min\{a_0, a_1, \ldots, a_J\}$. If in addition, $V_0^{(J)} \to V_0^{(\infty)}$ as $J \to \infty$, then, since our assumption on the cost function implies that $a_j \to \infty$ as $j \to \infty$, we have $V_0^{(\infty)} = \min\{a_0, a_1, \ldots\} = a_m$ for some finite m. The Bayes sequential rule is the fixed sample size rule that stops at stage m whatever the observations. ■

7.2.4. (a) The posterior diatribution of θ given $X_1,\ldots,X_j$ is the gamma, $\mathcal{G}(S_j+1, 1/(j+1))$, where $S_j = \sum_1^j X_j$. Tha Bayes estimate of θ after stopping at stage j is the mean of this distribution, $\hat{\theta}_j = (S_j+1)/(j+1)$. The Bayes terminal risk is K times the variance of this distribution, $\rho_j = K(S_j+1)/(j+1)^2$.

(b) The conditional distribution of X_j given S_{j-1} (with θ integrated out) is the negative binomial, $\mathcal{NB}(S_{j-1}+1, 1/(j+1))$. This has mean $(S_{j-1}+1)/j$, so

$$\begin{aligned}
\mathrm{E}(\rho_j(X_1,\ldots,X_{j-1},X_j|X_1,\ldots,X_{j-1}) &= \frac{K}{(j+1)^2}\mathrm{E}(S_j+1|S_{j-1}) \\
&= \frac{K}{(j+1)^2}(S_{j-1}+1+\frac{S_{j-1}+1}{j}) = \frac{K(S_{j-1}+1)}{j(j+1)}.
\end{aligned}$$

(c) With $K = 12$ and $c = 1$, we have $U_j = \rho_j + j = 12(S_j+1)/(j+1)^2 + j$ and $\mathrm{E}(U_j|S_{j-1}) = 12(S_{j-1}+1)/(j(j+1)) + j$. We have $V_3^{(3)} = U_3 = (3/4)(S_3+1)+3$ and $\mathrm{E}(V_3^{(3)}|S_2) = S_2 + 4$. We compute a table similar to Table 7.1.

S_2	d_2^0	U_2	$\mathrm{E}(V_3^{(3)}\|S_2)$	$V_2^{(3)}$	ϕ_2^0
0	1/3	10/3	4	10/3	1
1	2/3	14/3	5	14/3	1
2	3/3	18/3	6	6	any
3	4/3	22/3	7	7	0
⋮	⋮	⋮	⋮	⋮	⋮

S_1	d_1^0	U_1	$\mathrm{E}(V_2^{(3)}\|S_1)$	$V_1^{(3)}$	ϕ_1^0
0	1/2	4	$4-(1/54)$	$4-(1/54)$	0
1	2/2	7	$6-(4/27)$	$6-(4/27)$	0
2	3/2	10	$7+(1/2)$	$7+(1/2)$	0
3	4/2	13	9	9	0
⋮	⋮	⋮	⋮	⋮	⋮

S_0	d_0^0	U_0	$\mathrm{E}(V_1^{(3)}\|S_0)$	$V_0^{(3)}$	ϕ_0^0
0	1	12	$6-(8/27)$	$6-(8/27)$	0

An optimal stopping rule for the problem truncated at 3 is to stop at stage 2 if $S_2 \le 2$ and stop at stage 3 otherwise. The value is $5.7037\cdots$.

5. The posterior distribution of θ given $X_1,\ldots,X_j$ is $\mathcal{B}e(S_j+1, j-S_j+1)$, where $S_j = X_1+\cdots+X_j$. The mean of the posterior is $(S_j+1)/(j+2)$. The Bayes terminal decision rule, d_j, at stage j is to take action a_1 if $S_j \le j/2$ and to take action a_2 otherwise. The Bayes terminal loss, ρ_j, is $K(S_j+1)/(j+2)$ if $S_j \le j/2$ and $K(j-S_j+1)/(j+2)$ otherwise. We compute a table similar to Table 7.1 for the problem truncated at $J=5$.

S_5	d_5^0	Cost	ρ_5	$U_5 = V_5^{(5)}$
0	a_1	5	8	13
1	a_1	5	16	21
2	a_1	5	24	29
3	a_2	5	24	29
4	a_2	5	16	21
5	a_2	5	8	13

S_4	d_4^0	Cost	ρ_4	U_4	$E(V_5^{(5)}\|S_4)$	$V_4^{(5)}$	$\phi_4^{(0)}$
0	a_1	4	9.333	13.333	14.333	13.333	1
1	a_1	4	18.667	22.667	23.667	22.667	1
2	a_1	4	28.000	32.000	29.000	29.000	0
3	a_2	4	18.667	22.667	23.667	22.667	1
4	a_2	4	9.333	13.333	14.333	13.333	1

S_3	d_3^0	Cost	ρ_3	U_3	$E(V_4^{(5)}\|S_3)$	$V_3^{(5)}$	$\phi_3^{(0)}$
0	a_1	3	11.200	14.200	15.200	14.200	1
1	a_1	3	22.400	25.400	25.200	25.200	0
2	a_2	3	22.400	25.400	25.200	25.200	0
3	a_2	3	11.200	14.200	15.200	14.200	1

S_2	d_2^0	Cost	ρ_2	U_2	$E(V_3^{(5)}\|S_2)$	$V_2^{(5)}$	$\phi_2^{(0)}$
0	a_1	2	14.000	16.000	16.950	16.000	1
1	a_1	2	28.000	30.000	25.200	25.200	0
2	a_2	2	14.000	16.000	16.950	16.000	1

S_1	d_1^0	Cost	ρ_1	U_1	$E(V_2^{(5)}\|S_1)$	$V_1^{(5)}$	$\phi_1^{(0)}$
0	a_1	1	18.667	19.667	19.067	19.067	0
1	a_2	1	18.667	19.667	19.067	19.067	0

S_0	d_0^0	Cost	ρ_0	U_0	$E(V_1^{(5)}\|S_0)$	$V_0^{(5)}$	$\phi_0^{(0)}$
0	a_1	0	28.000	28.000	19.067	19.067	0

7.2.6. (a) The joint density of $X_1,\ldots,X_n$ and θ is proportional to $\exp\{-(1/2)(\sum_1^n X_i^2 - 2\theta S_n + n\theta^2) - (1/2\sigma^2)\theta^2\}$ where $S_n = \sum_1^n X_i$, so the posterior density of θ given $X_1,\ldots,X_n$ is proportional to this also. This gives

$$g(\theta|X_1,\ldots,X_n) = c\exp\{-\frac{1}{2}\frac{1+n\sigma^2}{\sigma^2}(\theta - \frac{\sigma^2 S_n}{1+n\sigma^2})^2\}.$$

This is the normal distribution, $\mathcal{N}(\sigma^2 S_n/(1+n\sigma^2), \sigma^2/(1+n\sigma^2))$.

(b) With squared error loss, the Bayes terminal decision rule is $d_n^0 = \mathrm{E}(\theta|X_1,\ldots,X_n) = \sigma^2 S_n/(1+n\sigma^2)$, with expected loss $\rho_n = \sigma^2/(1+n\sigma^2)$. Thus, the $U_n = nc + (\sigma^2/(1+n\sigma^2))$ is constant, and since $\rho_n \to 0$, Theorems 3 and 4 imply that the Bayes sequential decision rule is the fixed sample size rule: Take a sample of size n, where n is that value of j that minimizes U_j, and estimate θ to be $\sigma^2 S_n/(1+n\sigma^2$. It has Bayes risk $\min_j(jc + (\sigma^2/(1+n\sigma^2))$.

(c) Let n be the integer j that minimizes $jc + 1/j$. The rule that observes a fixed sample of size n and estimates θ to be S_n/n is minimax since it has constant risk, $nc + \mathrm{E}_\theta(\theta - S_n/n)^2 = nc + 1/n$, and is extended Bayes (since the minimum Bayes risk of part (b) converges to $nc + 1/n$ as $\sigma^2 \to \infty$).

7.2.7. (a) The posterior density of θ given $X_1,\ldots,X_j$ is proportional to $\theta^{S_j+\alpha-1}\exp\{-\theta(\lambda+j)\}$, which makes it the gamma distribution, $\mathcal{G}(S_j+\alpha,(\lambda+j)^{-1})$.

(b) $\mathrm{E}(\theta-a)^2/\theta$ is minimized by $a=1/\mathrm{E}\theta^{-1}$ and the minimum value is $\mathrm{E}\theta-1/\mathrm{E}\theta^{-1}$. For the $\mathcal{G}(S_j+\alpha,(\lambda+j)^{-1})$ distribution, we have $\mathrm{E}\theta=(S_j+\alpha)/(\lambda+j)$ and $\mathrm{E}\theta^{-1}=(\lambda+j)/(S_j+\alpha-1)$. So the Bayes estimate is $d_0=(S_j+\alpha-1)/(\lambda+j)$ and the Bayes expected loss is $\rho_j=(\lambda+j)^{-1}$.

(c) Since the $U_j=jc+(\lambda+j)^{-1}$ are constant and since $\rho_j\to 0$, we have from Theorems 3 and 4 that the Bayes sequential decision rule is the fixed sample size rule: Take a sample of size n, where n is that value of j that minimizes U_j, stop and estimate θ to be $(S_n+\alpha)/(\lambda+n)$. It has Bayes risk $\min_j(jc+(\lambda+j)^{-1})$. Since n is the first j such that $(j+1)c+(j+1)^{-1}\geq jc+j^{-1}$, it is also the first j such that $j(j+1)\geq 1/c$, and (adding 1/4 to both sides), the first j such that $(j-1/2)^2\geq(1/c)+(1/4)$. From this we see that $n=[\sqrt{1/c+1/4}]$, where $[x]$ denotes the integer nearest to x.

(d) Consider the rule, call it (d,ϕ): Find n that minimizes $jc+j^{-1}$, take a sample of size n and estimate θ to be S_n/n. This corresponds to the rule of part (c) with $\alpha=0$ and $\lambda=0$. We will show it is minimax by showing it has constant risk, and is extended Bayes. The risk function is $R(\theta,(d,\phi))=jc+\mathrm{E}_\theta(\theta-S_j/j)^2/\theta=jc+\mathrm{Var}(S_j/j)/\theta=jc+1/j$, independentof θ. Since the Bayes risks of (c), namely $\min_j(jc+(\lambda+j)^{-1})$, converge to $\min_j(jc+j^{-1})$ as $\lambda\to 0$, we see that the rule (d,ϕ) is extended Bayes and hence minimax.

7.2.8. (a) The joint density of $X_1,\ldots,X_n$ and θ is proportional to $\theta^{n+\alpha-1}\exp\{\lambda+S_n\}$ where $S_n=\sum_1^n X_i$, so the posterior density of θ given $X_1,\ldots,X_n$ must be proportional to this also. This is the gamma distribution, $\mathcal{G}(n+\alpha,(\lambda+S_n)^{-1})$.

(b) $\mathrm{E}(\theta-a)^2/\theta^2$ is minimized by $a=\mathrm{E}(\theta^{-1})/\mathrm{E}(\theta^{-2})$ and the minimum value is $1-(\mathrm{E}\theta^{-1})^2/\mathrm{E}\theta^{-2}$. For the $\mathcal{G}(n+\alpha,(\lambda+S_n)^{-1})$ distribution, we have $\mathrm{E}\theta^{-1}=(\lambda+S_n)/(n+\alpha-1)$ and $\mathrm{E}\theta^{-2}=(\lambda+S_n)^2/((n+\alpha-1)(n+\alpha-2))$. Therefore, the Bayes estimate is $d_n^0=(n+\alpha-2)/(\lambda+S_n)$ and the Bayes terminal risk is $\rho_n=1/(n+\alpha-1)$. (Note the misprint in the text.)

(c) Since the $U_n=nc+1/(n+\alpha-1)$ are constant and since $\rho_n\to 0$, we have from Theorems 3 and 4 that the Bayes sequential decision rule is the fixed sample size rule: Take a sample of size n, where n minimizes U_j, stop and estimate θ to be $(n+\alpha-2)/(\lambda+S_n)$. It has Bayes risk $\min_j U_j$.

(d) Consider the rule, call it (d,ϕ): Find j that minimizes $jc+(j-1)^{-1}$, take a sample of size j and estimate θ to be $(j-2)/S_j$. This corresponds to the rule of part (c) with $\alpha=0$ and $\lambda=0$. We will show it is minimax by showing it has constant risk, and is extended Bayes. The risk function is

$$\begin{aligned} R(\theta,(d,\phi)) &= jc+\mathrm{E}_\theta(\theta-(j-2)/S_j)^2/\theta^2 \\ &= jc+1-2(j-2)\mathrm{E}_\theta(\theta S_j)^{-1}+(j-2)^2\mathrm{E}_\theta(\theta S_j)^{-2}. \end{aligned}$$

The distribution of θS_j is $\mathcal{G}(j,1)$ independent of θ so the risk is independent of θ. We find $\mathrm{E}(\theta S_j)^{-1}=1/(j-1)$ and $\mathrm{E}(\theta S_j)^{-2}=1/((j-1)(j-2))$, provided $j>2$. Substitution into the formula for R gives $R(\theta,(d,\phi))=jc+(j-1)^{-1}$. (This formula also holds for $j=2$ as is checked by direct computation.) Since the Bayes risks of (c), namely $\min_n\{nc+1/(n+\alpha-1)\}$, converge to $jc+(j-1)^{-1}=\min_n\{nc+1/(n-1)\}$ as $\alpha\to 0$, we see that the rule (d,ϕ) is extended Bayes and minimax. The optimal $j=1+[\sqrt{(1/c)+(1/4)}]$, where $[x]$ denotes the integer nearest to x, is always at least 2.

7.2.9. (a) The joint density of $X_1,\ldots,X_j$ and θ is proportional to $\theta^{-(j+\alpha+1)}\exp\{(\lambda+S_j)/\theta\}$ so the posterior distribution of θ given $X_1,\ldots,X_j$ is proportional to this also. This is the reciprocal gamma $\mathcal{G}^{-1}(j+\alpha,(\lambda+S_j)^{-1})$.

(b) For the $\mathcal{G}^{-1}(j+\alpha,(\lambda+S_j)^{-1})$ distribution, we have $\mathrm{E}\theta^{-1}=(j+\alpha)/(\lambda+S_j)$ and $\mathrm{E}\theta^{-2}=(j+\alpha)(j+\alpha-1)/(\lambda+S_n)^2$. Therefore, the Bayes estimate is $d_j^0=\mathrm{E}(\theta^{-1})/\mathrm{E}(\theta^{-2})=(\lambda+S_j)/(j+\alpha+1)$ and the Bayes terminal risk is $\rho_j=1-(\mathrm{E}\theta^{-1})^2/\mathrm{E}\theta^{-2}=1/(j+\alpha+1)$.

(c) Since the $U_j=jc+(j+\alpha+1)^{-1}$ are constant and since $\rho_j\to 0$, we have from Theorems 3 and 4 that the Bayes sequential decision rule is the fixed sample size rule: Take a sample of size n, where n minimizes U_j, stop and estimate θ to be $(\lambda+S_n)/(n+\alpha+1)$. It has Bayes risk $nc+(n+\alpha+1)^{-1}$.

(d) Consider the rule, call it (d,ϕ): Find j that minimizes $jc+(j+1)^{-1}$, take a sample of size j and estimate θ to be $S_j/(j+1)$. This corresponds to the rule of part (c) with $\alpha=0$ and $\lambda=0$. We will show it is minimax by showing it has constant risk, and is extended Bayes. The risk function is

$$\begin{aligned} R(\theta,(d,\phi)) &= jc + \mathrm{E}_\theta(\theta - S_j/(j+1))^2/\theta^2 \\ &= jc + 1 - 2(j+1)^{-1}\mathrm{E}_\theta(S_j/\theta) + (j+1)^{-2}\mathrm{E}_\theta(S_j/\theta)^2. \end{aligned}$$

The distribution of S_j/θ is $\mathcal{G}(j,1)$ independent of θ so the risk is independent of θ. We find $\mathrm{E}(S_j/\theta) = j$ and $\mathrm{E}(S_j/\theta)^2 = j(j+1)$. Substitution into the formula for R gives $R(\theta,(d,\phi)) = jc + (j+1)^{-1}$. Since the Bayes risks of (c), namely $\min_j\{jc + (n+\alpha-1)^{-1}\}$, converge to $jc + (j-1)^{-1} = \min_j\{jc + (n-1)^{-1}\}$ as $\alpha \to 0$, we see that the rule (d,ϕ) is extended Bayes and minimax. The optimal j is $[\sqrt{(1/c)+(1/4)}]-1$, where $[x]$ denotes the integer nearest to x.