19.5 Exercises

1. Suppose that $x_1, x_2, \ldots$ are conditionally independent Bernoulli experiments, all with success probability θ. You interested in estimating θ and your terminal decision utility function is

$$u(\theta, \boldsymbol{a}) = -\frac{(\theta - \boldsymbol{a})^2}{\theta(1-\theta)}$$

This is the negative of the traditional squared error loss, standardized using the variance of a single observation. With this utility function, errors of the same absolute magnitude are considered worse if the are made near the extremes of the interval $(0, 1)$. You have a Beta(1.5,1.5) prior on θ. Each observation costs you c. What is the optimal fixed sample size? You may start using computers at any stage in the solution of this problem.

2. You are the mayor of Solvantville and you interested in estimating the difference in the concentration of a certain contaminant in the water supply of homes in two different sections of town, each with its own well. If you found that a difference you would further investigate whether there may be contamination due to industrial waste. How big a sample of homes do you need for your purposes? Specify a terminal loss function, a cost of observation, a likelihood and a prior. Keep it very simple, but justify your assumptions and point out limitations if necessary.

3. Can you find the optimal sample size in example 19.1 when σ is not known? You can use a conjugate prior, and standard results such as those of section 3.3 in "Bayesian Data Analysis" by Gelman et al.

4. Proof Theorem (23).

5. Show that the posterior median is the Bayes rule when the utility function is of the form $u(\theta, a) = -|\theta - a|$.

6. In the context of Section 19.4 determine the optimal n by maximizing

$$U(\pi, n) = \int U(\pi_x, \delta_n^\pi(x^n))p(x^n)dx^n$$

$\mathbf{X}_n = \mathbf{x}_n$ is normal with posterior precision $\tau + n\omega$. Therefore,

$$\begin{aligned} U_0(\pi^0, 0) &= -Var_\pi(\theta) = -\frac{1}{\tau} \\ U_0(\pi^n, n) &= -Var_{\pi^n}(\theta) - nc = -\frac{1}{\tau + n\omega} - nc \end{aligned}$$

and, so $\rho_0(\pi^n) = -\frac{1}{\tau+n\omega}$ for $n = 0, 1, 2, \ldots$ This result shows that the expected utility depends only on the number of observations taken and not on their observed values. This implies that the optimal sequential decision procedure is a procedure in which a fixed number of observations is taken. For a formalization of this result, see Exercise 2.

Exploring a bit more this example, we can observe that

$$\begin{aligned} \rho_1(\pi^n) &= \max\{\rho(\pi^n), E^{p^*}[\rho(\pi^n(\theta|X_{n+1}))] - c\} \\ &= \max\left\{-\frac{1}{\tau + n\omega}, -\frac{1}{\tau + (n+1)\omega} - c\right\} \\ &= -\frac{1}{\tau + n\omega} \end{aligned}$$

$$\text{if } \frac{1}{\tau + n\omega} < \frac{1}{\tau + (n+1)\omega} + c \Longleftrightarrow \omega < c(\tau + n\omega)(\tau + (n+1)\omega).$$

In fact, by using an induction argument (Exercise 3) one can show that $U_j(\pi^n) = -\frac{1}{\tau+n\omega} - nc$ for $\beta_n < \omega < \beta_{n+1}$ where β is defined as: $\beta_0 = 0, \beta_n = c(\tau + (n-1)\omega)(\tau + n\omega)$ for $n = 1, \ldots, j$ and $\beta_{j+1} = \infty$. This implies that the optimal sequential decision is to take exactly n^* observations, where n^* is an integer such that $\beta_{n^*} < \omega < \beta_{n^*+1}$. If the decision maker is not allowed to take more than m observations, then if $n^* > m$, he should take m observations. $\star$

19.7 Exercises

1. Sequential problems get very hard very quickly. This is one of the simplest you can try, but it will take some time.

Try to address the same question as worked exercise 19.6, except with Poisson observation and a Gamma$(1, 1)$ prior. Set $c = 1/12$.

2. [From DeG70] Let π be the prior distribution of θ in a sequential decision problem. (a) Prove by using an induction argument that, in a sequential decision problem with fixed cost per observation, if $\rho_0(\pi^n)$ is constant for all observed values $\mathbf{x}^n$, then every function ρ_j for $j = 1, 2, \ldots$ as well as ρ_0

has this property. (b) Conclude that the optimal sequential decision procedure is a procedure in which a fixed number of observations is taken.

3. [From DeG70] (a) In the context of Example 19.6, for $n = 0, 1, \ldots, j$, prove that $U_j(\pi^n) = -\frac{1}{\tau+n\omega} - nc$ for $\beta_n < \omega < \beta_{n+1}$ where β is defined as: $\beta_0 = 0, \beta_n = c(\tau + (n-1)\omega)(\tau + n\omega)$ for $n = 1, \ldots, j$ and $\beta_{j+1} = \infty$. (b) Conclude that if the decision maker is allowed to take no more than m observations, then the optimal procedure is to take $n^* \leq m$ observations where n^* is such that $\beta_{n^*} < \omega < \beta_{n^*+1}$.