Foundations of Statistical Inference, BS2a, Exercises 3

1. Let θ be a real-valued parameter and $f(x \mid \theta)$ be the probability density function of an observation x, given θ. Let H_0 be the hypothesis that $\theta = \theta_0$ and H_1 be the hypothesis that $\theta \sim g(\theta)$ where $g(\theta)$ is the prior for θ under H_1. The prior probability for H_0 is β.

   (a) Write down expressions for the (joint) distribution $P(H, \theta)$ (observe that H can take only two values), the marginal distribution $P(x)$ and $P(H_0, \theta|x)$ and $P(H_1, \theta|x)$.

   (b) Derive an expression for $\pi(H_1 \mid x)$, the posterior probability of H_1.

   Suppose that $x_1, \ldots, x_n$ is a sample from a normal distribution with mean θ and variance v. Let $\beta = 1/2$ and let

   $$g(\theta) = (2\pi w^2)^{-1/2} \exp\left\{-\theta^2/(2w^2)\right\}$$

   for $-\infty < \theta < \infty$.

   Show that, if $\theta_0 = 0$ and the sample mean is observed to be $10(v/n)^{1/2}$ then

   (c) the most powerful frequentist test of size $\alpha = 0.05$ will reject H_0 for any value of n;

   (d) the posterior probability of H_0 converges to 1, as $n \to \infty$.

   (e) Comment on the apparent contradiction between (c) and (d).

---

2. . The risks for five decision processes $\delta_1, \ldots, \delta_5$ depend on the value of a positive-valued parameter θ. The risks are given in the table below

	δ_1	δ_2	δ_3	δ_4	δ_5
$0 \le \theta < 1$	10	10	7	6	8
$1 \le \theta < 2$	8	11	8	5	10
$2 \le \theta$	15	11	12	14	14

   (a) Which decision procedures are at least as good as δ_1 for all θ ?

(b) Which decision procedures are admissible?

(c) Which is the minimax procedure?

(d) Suppose θ has a uniform distribution on $[0, 5]$. Which is the Bayes procedure and what is the Bayes risk for that procedure?

---

3. Suppose X is an observation from the distribution

$$f(x; \theta) = (x-1)\theta^2(1-\theta)^{x-2},\ x = 2, 3, \ldots,\ 0 < \theta < 1.$$

The prior distribution for θ is a Beta distribution

$$\pi(\theta) = \frac{\Gamma(7)}{\Gamma(4)\Gamma(3)}\theta^3(1-\theta)^2,\ 0 \le \theta \le 1,$$

and the loss from estimating θ by $\hat{\theta}$ is $(\theta - \hat{\theta})^2$.

(a) Find the posterior distribution, $\pi(\theta|x)$.

(b) Show that the Bayes estimator of θ is $6/(7+x)$.

(c) Show that the Bayes risk associated with this estimator is $43200 \sum_{x=2}^{\infty} \frac{x-1}{x+7} \times \frac{(x+1)!}{(8+x)!}$.

---

4. . Let $X_1, \ldots, X_n$ be a random sample from $N(\theta, 1)$. There are two hypotheses, $H_0 : \theta = 1$ and H_1, where $P(H_0) = p$, and $P(H_1) = 1 - p$.

(a) Given that H_1 specifies $\theta = -1$, show that

$$P(H_0 \mid \mathbf{x}) = \frac{pe^{\sum x_i}}{pe^{\sum x_i} + (1-p)e^{-\sum x_i}}$$

(b) Given that H_1 specifies $\theta \neq 1$ and gives θ the prior distribution

$$p(\theta \mid H_1) = \frac{1}{\sqrt{2\pi}} \exp(-\theta^2/2),\ \theta \neq 1,$$

determine $P(H_0 \mid \mathbf{x})$ when $\sum x_i = n$.

---

5. Let $X \sim \text{Binomial}(n, \theta)$ and suppose we have three priors $H_0 : \theta = 1/2$, $H_0' : \theta \sim \text{Beta}(\alpha, \alpha)$ and $H_1 : \theta \sim U(0, 1)$.

   (a) Write down expressions for the Bayes factors (i) B for H_0 *v.* H_1 and (ii) B' for H_0' *v.* H_1

(b) Describe what happens to B' as α varies from 1 to ∞.

---

6. Suppose we have a single observation, x_1, which comes from a distribution with density function $f(x)$ and we want to test

$$H_0 : f(x) = f_0(x) = 2(1-x),\ 0 \le x \le 1,$$

against

$$H_1 : f(x) = f_1(x) = 2x,\ 0 \le x \le 1.$$

(a) Using Neyman-Pearson, show that the best critical region for the likelihood ratio test of H_0 vs H_1 is given by $x_1 \ge B$ for some constant B.

(b) Consider now choosing B using decision theory. Suppose the losses incurred by Type I and Type II errors are equal. A decision $\delta_B(x)$ chooses H_1 if $x_1 \ge B$.

i. Calculate the risks $R(H_0, \delta_B(x))$ and $R(H_1, \delta_B(x))$ as functions of B. Use this to find the value of B which gives the minimax test procedure directly, without finding the Bayes procedure.

ii. Calculate the Bayes risk, when the prior probabilities are $1/4, 3/4$ for H_0 and H_1, and find the value of B which gives the Bayes test.

iii. Suppose the prior probabilities are $\nu, 1-\nu$ for H_0 and H_1. Find the value of B which gives the Bayes test procedure $\delta_{B(\nu)}(x)$), and the value of ν which gives $R(H_0, \delta_{B(\nu)}(x)) = R(H_1, \delta_{B(\nu)}(x))$ (so the Bayes procedure is a minimax procedure).

(c) Suppose the losses incurred by Type I and Type II errors are a and b respectively and the prior probabilities are $\nu, 1-\nu$ for H_0 and H_1. Show (by direct calculation) that the critical region for the Bayes test is

$$x : \frac{f_1(x)}{f_0(x)} \ge \frac{\nu a}{(1-\nu)b}.$$

---

7. Events occur in a Poisson process with intensity λ. Let n_t be the number of events which have been observed by time t. The prior density of λ is given by

$$\pi(\lambda) = \frac{(\beta\lambda)^{\alpha-1} e^{-\beta\lambda}\beta}{\Gamma(\alpha)}$$

with α, β and $\lambda > 0$ where α and β are known parameters. The cost associated with estimating λ by d is $(d - \lambda)^2$.

(a) i. Derive the posterior distribution of λ given n_t.

ii. Find the Bayes estimator of λ.

(b) Suppose now that there is a cost ct associated with collecting observations for a period of time t. The combined cost of estimating λ by d is then $(d - \lambda)^2 + ct$.

i. What is the prior expectation of this cost, anticipating the use of the Bayes estimator of λ? (Hint: Start by computing $R(\lambda, d)$ the expected cost of the estimator d obtained in the previous question for a given value λ).

ii. Show that, when $c > \alpha/\beta^3$, it is not worth collecting any observations.

iii. What is the optimal value of t when $c < \alpha/\beta^3$?

---

8. Suppose $X_i \sim f(x;\theta), i = 1, ..., n$ for $\theta \in \Theta$. Consider a pair of hypotheses of the form $H_0 : \theta \in \Theta_0$ and $H_1 : \theta \in \Theta \setminus \Theta_0$

(a) How is the posterior probability for H_0 related to the Bayes factor (in favor of H_0) comparing H_0 and H_1?

The cost of accepting H_1 when H_0 is true is a_0 and the cost of accepting H_0 when H_1 is true is a_1. There is no cost for accepting the correct hypothesis.

(b) Show that the Bayes decision rule accepts H_0 when the posterior probability of H_0 is larger than $a_1/(a_0 + a_1)$ and accepts H_1 otherwise.

Suppose $X_i \sim N(\theta, \sigma^2), i = 1, ..., n$ iid and $\theta \sim N(0, \tau^2)$, where σ and τ are known constants. Let H_0 be the hypothesis that $\theta < 0$.

(c) Show that the posterior distribution for θ given data $X_i = x_i, i = 1, 2, ..., n$ is normal, and give its mean and variance.

(d) Find the Bayes factor B (in favor of H_0) comparing H_0 and H_1.

(e) You decide to reject H_0 when $B < 1/10$. Show that this is equivalent to rejecting H_0 when $\bar{x} > c$ and find a formula for the value of c.

---

9. **(Optional)** Let $X_1, \ldots, X_n$ be a random sample from $N(\theta, \sigma^2)$, σ^2 known, and let the prior distribution be $\theta \sim N(\phi, \tau^2)$. Suppose θ is to be estimated under a quadratic loss function.

(a) State the Bayes estimator $\hat{\theta}_N$ of θ and show that this estimator has constant risk as $\tau \to \infty$.

(b) What is the Bayes estimator $\hat{\theta}_U$ for the Uniform prior? Compare with $\lim \hat{\theta}_N$ when $\tau \to \infty$. Show that $R(\theta, \hat{\theta}_U) = \lim_{\tau \to \infty} R(\theta, \hat{\theta}_N)$.

(c) Deduce the minimax estimator of θ for a uniform prior and state its risk.