3.6 PROBLEMS AND COMPLEMENTS

Problems for Section 3.2

1. Show that if $X_1, \dots, X_n$ is a $\mathcal{N}(\theta, \sigma^2)$ sample and π is the improper prior $\pi(\theta) = 1$, $\theta \in \Theta = R$, then the improper Bayes rule for squared error loss is $\delta^*(\mathbf{x}) = \bar{x}$.

2. Let $X_1, \dots, X_n$ be the indicators of n Bernoulli trials with success probability θ. Suppose $l(\theta, a)$ is the quadratic loss $(\theta - a)^2$ and that the prior $\pi(\theta)$ is the beta, $\beta(r, s)$, density. Find the Bayes estimate $\widehat{\theta}_B$ of θ and write it as a weighted average $w\theta_0 + (1 - w)\bar{X}$ of the mean θ_0 of the prior and the sample mean $\bar{X} = S/n$. Show that $\widehat{\theta}_B = (S+1)/(n+2)$ for the uniform prior.

3. In Problem 3.2.2 preceeding, give the MLE of the Bernoulli variance $q(\theta) = \theta(1 - \theta)$ and give the Bayes estimate of $q(\theta)$. Check whether $q(\widehat{\theta}_B) = E(q(\boldsymbol{\theta}) \mid \mathbf{x})$, where $\widehat{\theta}_B$ is the Bayes estimate of θ.

4. In the Bernoulli Problem 3.2.2 with uniform prior on the probabilility of success θ, we found that $(S + 1)/(n + 2)$ is the Bayes rule. In some studies (see Section 6.4.3), the parameter $\lambda = \theta/(1 - \theta)$, which is called the *odds ratio* (for success), is preferred to θ. If we put a (improper) uniform prior on λ, under what condition on S does the Bayes rule exist and what is the Bayes rule?

5. Suppose $\boldsymbol{\theta} \sim \pi(\theta)$, $(X \mid \boldsymbol{\theta} = \theta) \sim p(x \mid \theta)$.

(a) Show that the joint density of X and θ is

$$f(x, \theta) = p(x \mid \theta)\pi(\theta) = c(x)\pi(\theta \mid x)$$

where $c(x) = \int \pi(\theta)p(x \mid \theta)d\theta$.

(b) Let $l(\theta, a) = (\theta - a)^2/w(\theta)$ for some weight function $w(\theta) > 0$, $\theta \in \Theta$. Show that the Bayes rule is

$$\delta^* = E_{f_0}(\boldsymbol{\theta} \mid x)$$

where

$$f_0(x, \theta) = p(x \mid \theta)[\pi(\theta)/w(\theta)]/c$$

and

$$c = \int\int p(x \mid \theta)[\pi(\theta)/w(\theta)]d\theta dx$$

is assumed to be finite. That is, if π and l are changed to $a(\theta)\pi(\theta)$ and $l(\theta, a)/a(\theta)$, $a(\theta) > 0$, respectively, the Bayes rule does not change.

Hint: See Problem 1.4.24.

(c) In Example 3.2.3, change the loss function to $l(\theta, a) = (\theta - a)^2/\theta^\alpha(1 - \theta)^\beta$. Give the conditions needed for the posterior Bayes risk to be finite and find the Bayes rule.

6. Find the Bayes risk $r(\pi, \delta)$ of $\delta(\mathbf{x}) = \bar{X}$ in Example 3.2.1. Consider the relative risk $e(\delta, \pi) = R(\pi)/r(\pi, \delta)$, where $R(\pi)$ is the Bayes risk. Compute the limit of $e(\delta, \pi)$ as

(a) $\tau \to \infty$, **(b)** $n \to \infty$, **(c)** $\sigma^2 \to \infty$.

7. For the following problems, compute the posterior risks of the possible actions and give the optimal Bayes decisions when $x = 0$.

(a) Problem 1.3.1(d);

(b) Problem 1.3.2(d)(i) and (ii);

(c) Problem 1.3.19(c).

8. Suppose that $N_1, \ldots, N_r$ given $\boldsymbol{\theta} = \theta$ are multinomial $\mathcal{M}(n, \theta)$, $\theta = (\theta_1, \ldots, \theta_r)^T$, and that $\boldsymbol{\theta}$ has the Dirichlet distribution $\mathcal{D}(\alpha)$, $\alpha = (\alpha_1, \ldots, \alpha_r)^T$, defined in Problem 1.2.15. Let $q(\theta) = \sum_{j=1}^r c_j\theta_j$, where $c_1, \ldots, c_r$ are given constants.

(a) If $l(\theta, a) = [q(\theta) - a]^2$, find the Bayes decision rule δ^* and the minimum conditional Bayes risk $r(\delta^*(x) \mid x)$.
Hint: If $\boldsymbol{\theta} \sim \mathcal{D}(\alpha)$, then $E(\boldsymbol{\theta}_j) = \alpha_j/\alpha_0$, $\text{Var}(\boldsymbol{\theta}_j) = \alpha_j(\alpha_0 - \alpha_j)/\alpha_0^2(\alpha_0 + 1)$, and $\text{Cov}(\boldsymbol{\theta}_j, \boldsymbol{\theta}_j) = -\alpha_i\alpha_j/\alpha_0^2(\alpha_0 + 1)$, where $\alpha_0 = \sum_{j=1}^r \alpha_j$. (Use these results, do not derive them.)

(b) When the loss function is $l(\theta, a) = (q(\theta) - a)^2/\prod_{j=1}^r \theta_j$, find necessary and sufficient conditions under which the Bayes risk is finite and under these conditions find the Bayes rule.

(c) We want to estimate the vector $(\theta_1, \ldots, \theta_r)$ with loss function $l(\theta, a) = \sum_{j=1}^r (\theta_j - a_j)^2$. Find the Bayes decision rule.

9. *Bioequivalence trials* are used to test whether a generic drug is, to a close approximation, equivalent to a name-brand drug. Let $\theta = \mu_G - \mu_B$ be the difference in mean effect of the generic and name-brand drugs. Suppose we have a sample $X_1, \ldots, X_n$ of differences in the effect of generic and name-brand effects for a certain drug, where $E(X) = \theta$. A regulatory agency specifies a number $\epsilon > 0$ such that if $\theta \in (-\epsilon, \epsilon)$, then the generic and brand-name drugs are, by definition, bioequivalent. On the basis of $\mathbf{X} = (X_1, \ldots, X_n)$ we want to decide whether or not $\theta \in (-\epsilon, \epsilon)$. Assume that given θ, $X_1, \ldots, X_n$ are i.i.d. $\mathcal{N}(\theta, \sigma_0^2)$, where σ_0^2 is known, and that $\boldsymbol{\theta}$ is random with a $\mathcal{N}(\eta_0, \tau_0^2)$ distribution.

There are two possible actions:

$$\begin{aligned} a &= 0 \Leftrightarrow \text{Bioequivalent} \\ a &= 1 \Leftrightarrow \text{Not Bioequivalent} \end{aligned}$$

with losses $l(\theta, 0)$ and $l(\theta, 1)$. Set

$$\lambda(\theta) = l(\theta, 0) - l(\theta, 1)$$

= difference in loss of acceptance and rejection of bioequivalence. Note that $\lambda(\theta)$ should be negative when $\theta \in (-\epsilon, \epsilon)$ and positive when $\theta \notin (-\epsilon, \epsilon)$. One such function (Lindley, 1998) is

$$\lambda(\theta) = r - \exp\left\{-\frac{1}{2c^2}\theta^2\right\}, \; c^2 > 0$$

where $0 < r < 1$. Note that $\lambda(\pm\epsilon) = 0$ implies that r satisfies

$$\log r = -\frac{1}{2c^2}\epsilon^2.$$

This is an example with two possible actions 0 and 1 where $l(\theta, 0)$ and $l(\theta, 1)$ are not constant. Any two functions with difference $\lambda(\theta)$ are possible loss functions at $a = 0$ and 1.

(a) Show that the Bayes rule is equivalent to

$$\text{"Accept bioequivalence if } E(\lambda(\boldsymbol{\theta}) \mid \mathbf{X} = \mathbf{x}) < 0\text{"} \tag{3.6.1}$$

and show that (3.6.1) is equivalent to

$$\text{"Accept bioequivalence if } [E(\boldsymbol{\theta} \mid \mathbf{x})]^2 < (\tau_0^2(n) + c^2)\{\log(\tfrac{c^2}{\tau_0^2(n)+c^2}) + \tfrac{\epsilon^2}{c^2}\}\text{"}$$

where

$$E(\boldsymbol{\theta} \mid \mathbf{x}) = w\eta_0 + (1-w)\bar{x},\ w = \tau_0^2(n)/\tau_0^2\sigma_0^2,\ \tau_0^2(n) = \left(\frac{1}{\tau_0^2} + \frac{n}{\sigma_0^2}\right)^{-1}.$$

Hint: See Example 3.2.1.

(b) It is proposed that the preceding prior is "uninformative" if it has $\eta_0 = 0$ and τ_0^2 large ("$\tau_0^2 \to \infty$"). Discuss the preceding decision rule for this "prior."

(c) Discuss the behavior of the preceding decision rule for large n ("$n \to \infty$"). Consider the general case (a) and the specific case (b).

10. For the model defined by (3.2.16) and (3.2.17), find

(a) the linear Bayes estimate of Δ_1.

(b) the linear Bayes estimate of μ.

(c) Is the assumption that the Δ's are normal needed in (a) and (b)?

Problems for Section 3.3

1. In Example 3.3.2 show that $L(\mathbf{x}, 0, v) \geq \pi/(1-\pi)$ is equivalent to $T \geq t$.

2. Suppose $g : S \times T \to R$. A point (x_0, y_0) is a *saddle point* of g if

$$g(x_0, y_0) = \sup_S g(x, y_0) = \inf_T g(x_0, y).$$

Suppose S and T are subsets of R^m, R^p, respectively, $(\mathbf{x}_0, \mathbf{y}_0)$ is in the interior of $S \times T$, and g is twice differentiable.

(a) Show that a necessary condition for $(\mathbf{x}_0, \mathbf{y}_0)$ to be a saddle point is that, representing $\mathbf{x} = (x_1, \ldots, x_m)$, $\mathbf{y} = (y_1, \ldots, y_p)$,

$$\frac{\partial g}{\partial x_i}(\mathbf{x}_0, \mathbf{y}_0) = \frac{\partial g}{\partial y_j}(\mathbf{x}_0, \mathbf{y}_0) = 0,$$

and

$$\frac{\partial^2 g}{\partial x_a \partial x_b}(\mathbf{x}_0, \mathbf{y}_0) \le 0, \quad \frac{\partial^2 g(\mathbf{x}_0, \mathbf{y}_0)}{\partial y_c \partial y_d} \ge 0$$

for all $1 \le i, a, b \le m$, $1 \le j, c, d \le p$.

(b) Suppose $S_m = \{\mathbf{x} : x_i \ge 0,\ 1 \le i \le m,\ \sum_{i=1}^m x_i = 1\}$, the simplex, and $g(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^m \sum_{j=1}^p c_{ij} x_i y_j$ with $\mathbf{x} \in S_m$, $\mathbf{y} \in S_p$. Show that the von Neumann minimax theorem is equivalent to the existence of a saddle point for any twice differentiable g.

3. Suppose $\Theta = \{\theta_0, \theta_1\}$, $\mathcal{A} = \{0, 1\}$, and that the model is regular. Suppose

$$l(\theta_i, i) = 0,\ l(\theta_i, j) = w_{ij} > 0,\ i, j = 0, 1,\ i \ne j.$$

Let $L_X(\theta_0, \theta_1) = p(X, \theta_1)/p(X, \theta_0)$ and suppose that $L_X(\theta_0, \theta_1)$ has a continuous distribution under both P_{θ_0} and P_{θ_1}. Show that

(a) For every $0 < \pi < 1$, the test rule δ_π given by

$$\begin{aligned} \delta_\pi(X) &= 1 \text{ if } L_X(\theta_0, \theta_1) \ge \tfrac{(1-\pi)w_{01}}{\pi w_{10}} \\ &= 0 \text{ otherwise} \end{aligned}$$

is Bayes against a prior such that $P[\theta = \theta_1] = \pi = 1 - P[\theta = \theta_0]$, and

(b) There exists $0 < \pi^* < 1$ such that the prior π^* is least favorable against δ_{π^*}, that is, the conclusion of von Neumann's theorem holds.

Hint: Show that there exists (a unique) π^* so that

$$R(\theta_0, \delta_{\pi^*}) = R(\theta_1, \delta_{\pi^*}).$$

4. Let $S \sim \mathcal{B}(n, \theta)$, $l(\theta, a) = (\theta - a)^2$, $\delta(S) = \bar{X} = S/n$, and

$$\delta^*(S) = (S + \frac{1}{2}\sqrt{n})/(n + \sqrt{n}).$$

(a) Show that δ^* has constant risk and is Bayes for the beta, $\beta(\sqrt{n}/2, \sqrt{n}/2)$, prior. Thus, δ^* is minimax.

Hint: See Problem 3.2.2.

(b) Show that $\lim_{n\to\infty}[R(\theta, \delta^*)/R(\theta, \delta)] > 1$ for $\theta \ne \frac{1}{2}$; and show that this limit equals 1 when $\theta = \frac{1}{2}$.

5. Let $X_1, \dots, X_n$ be i.i.d. $\mathcal{N}(\mu, \sigma^2)$ and $l(\sigma^2, d) = \left(\frac{d}{\sigma^2} - 1\right)^2$.

(a) Show that if μ is known to be 0

$$\delta^*(X_1, \dots, X_n) = \frac{1}{n+2} \sum X_i^2$$

is minimax.

(b) If $\mu = 0$, show that δ^* is uniformly best among all rules of the form $\delta_c(\mathbf{X}) = c\sum X_i^2$. Conclude that the MLE is inadmissible.

(c) Show that if μ is unknown, $\delta(\mathbf{X}) = \frac{1}{n+1}\sum(X_i - \bar{X})^2$ is best among all rules of the form $\delta_c(\mathbf{X}) = c\sum(X_i - \bar{X})^2$ and, hence, that both the MLE and the estimate $S^2 = (n-1)^{-1}\sum(X_i - \bar{X})^2$ are inadmissible.

Hint: (a) Consider a gamma prior on $\theta = 1/\sigma^2$. See Problem 1.2.12. (c) Use (B.3.29).

6. Let $X_1, \dots, X_k$ be independent with means $\mu_1, \dots, \mu_k$, respectively, where

$$(\mu_1, \dots, \mu_k) = (\mu_{i_1}^0, \dots, \mu_{i_k}^0),\ \mu_1^0 < \cdots < \mu_k^0$$

is a known set of values, and $i_1, \dots, i_k$ is an arbitrary unknown permutation of $1, \dots, k$. Let $\mathcal{A} = \{(j_1, \dots, j_k) : \text{Permutations of } 1, \dots, k\}$

$$l((i_1, \dots, i_k), (j_1, \dots, j_k)) = \sum_{l,m} 1(i_l < i_m, j_l > j_m).$$

Show that the minimax rule is to take

$$\delta(X_1, \dots, X_k) = (R_1, \dots, R_k)$$

where R_j is the rank of X_j, that is, $R_j = \sum_{l=1}^k 1(X_l \le X_j)$.

Hint: Consider the uniform prior on permutations and compute the Bayes rule by showing that the posterior risk of a permutation $(i_1, \dots, i_k)$ is smaller than that of $(i'_1, \dots, i'_k)$, where $i'_j = i_j$, $j \ne a, b$, $a < b$, $i'_a = i_b$, $i'_b = i_a$, and $R_a < R_b$.

7. Show that X has a Poisson (λ) distribution and $l(\lambda, a) = (\lambda - a)^2/\lambda$. Then X is minimax.

Hint: Consider the gamma, $\Gamma(k^{-1}, 1)$, prior. Let $k \to \infty$.

8. Let X_i be independent $\mathcal{N}(\mu_i, 1)$, $1 \le i \le k$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_k)^T$. Write $\mathbf{X} = (X_1, \dots, X_k)^T$, $\mathbf{d} = (d_1, \dots, d_k)^T$. Show that if

$$l(\boldsymbol{\mu}, \mathbf{d}) = \sum_{i=1}^k (d_i - \mu_i)^2$$

then $\boldsymbol{\delta}(\mathbf{X}) = \mathbf{X}$ is minimax.

Remark: Stein (1956) has shown that if $k \ge 3$, $\mathbf{X}$ is no longer unique minimax. For instance,

$$\delta^*(\mathbf{X}) = \left(1 - \frac{k-2}{|\mathbf{X}|^2}\right)\mathbf{X}$$

is also minimax and $R(\boldsymbol{\mu}, \delta^*) < R(\boldsymbol{\mu}, \boldsymbol{\delta})$ for all $\boldsymbol{\mu}$. See Volume II.

9. Show that if $(N_1, \dots, N_k)$ has a multinomial, $\mathcal{M}(n, p_1, \dots, p_k)$, distribution, $0 < p_j < 1$, $1 \le j \le k$, then $\frac{\mathbf{N}}{n}$ is minimax for the loss function

$$l(\mathbf{p}, \mathbf{d}) = \sum_{j=1}^k \frac{(d_j - p_j)^2}{p_j q_j}$$

where $q_j = 1 - p_j$, $1 \le j \le k$.

Hint: Consider Dirichlet priors on $(p_1, \dots, p_{k-1})$ with density defined in Problem 1.2.15. See also Problem 3.2.8.

10. Let $X_i (i = 1, \dots, n)$ be i.i.d. with unknown distribution F. For a given x we want to estimate the proportion $F(x)$ of the population to the left of x. Show that

$$\delta = \frac{\text{No. of } X_i \leq x}{\sqrt{n}} \cdot \frac{1}{1 + \sqrt{n}} + \frac{1}{2(1 + \sqrt{n})}$$

is minimax for estimating $F(x) = P(X_i \leq x)$ with squared error loss.

Hint: Consider the risk function of δ. See Problem 3.3.4.

11. Let $X_1, \dots, X_n$ be independent $\mathcal{N}(\mu, 1)$. Define

$$\begin{aligned} \delta(\bar{X}) &= \bar{X} + \frac{d}{\sqrt{n}} \text{ if } \bar{X} < -\frac{d}{\sqrt{n}} \\ &= 0 \text{ if } |\bar{X}| \leq \frac{d}{\sqrt{n}} \\ &= \bar{X} - \frac{d}{\sqrt{n}} \text{ if } \bar{X} > \frac{d}{\sqrt{n}}. \end{aligned}$$

(a) Show that the risk (for squared error loss) $E(\sqrt{n}(\delta(\bar{X}) - \mu))^2$ of these estimates is bounded for all n and μ.

(b) How does the risk of these estimates compare to that of $\bar{X}$?

12. Suppose that given $\boldsymbol{\theta} = \theta$, X has a binomial, $\mathcal{B}(n, \theta)$, distribution. Show that the Bayes estimate of θ for the Kullback–Leibler loss function $l_p(\theta, a)$ is the posterior mean $E(\boldsymbol{\theta} \mid X)$.

13. Suppose that given $\boldsymbol{\theta} = \theta = (\theta_1, \dots, \theta_k)^T$, $X = (X_1, \dots, X_k)^T$ has a multinomial, $\mathcal{M}(n, \theta)$, distribution. Let the loss function be the Kullback–Leibler divergence $l_p(\theta, a)$ and let the prior be the uniform prior

$$\pi(\theta_1, \dots, \theta_{k-1}) = (k-1)!, \ \theta_j \geq 0, \ \sum_{j=1}^{k-1} \theta_j = 1.$$

Show that the Bayes estimate is $(X_i + 1)/(n + k)$.

14. Let $K(p_\theta, q)$ denote the KLD (Kullback–Leibler divergence) between the densities p_θ and q and define the *Bayes KLD* between $\mathcal{P} = \{p_\theta : \theta \in \Theta\}$ and q as

$$k(q, \pi) = \int K(p_\theta, q)\pi(\theta)d\theta.$$

Show that the marginal density of X,

$$p(x) = \int p_\theta(x)\pi(\theta)d\theta,$$

minimizes $k(q, \pi)$ and that the minimum is

$$I_{\theta,X} \equiv \int \left[E_\theta \left\{ \log \frac{p_\theta(X)}{p(X)} \right\} \right] \pi(\theta) d\theta.$$

$I_{\theta,X}$ is called the *mutual information* between θ and X.

Hint: $k(q, \pi) - k(p, \pi) = \int \left[E_\theta \left\{ \log \frac{p(X)}{q(X)} \right\} \right] \pi(\theta) d\theta \geq 0$ by Jensen's inequality.

15. *Jeffrey's "Prior."* A density proportional to $\sqrt{I_p(\theta)}$ is called Jeffrey's prior. It is often improper. Show that in the $\mathcal{N}(\theta, \sigma_0^2)$, $\mathcal{N}(\mu_0, \theta)$ and $\mathcal{B}(n, \theta)$ cases, Jeffrey's priors are proportional to 1, θ^{-1}, and $\theta^{-\frac{1}{2}}(1-\theta)^{-\frac{1}{2}}$, respectively. Give the Bayes rules for squared error in these three cases.

Problems for Section 3.4

1. Let $X_1, \dots, X_n$ be the indicators of n Bernoulli trials with success probability θ. Show that $\bar{X}$ is an UMVU estimate of θ.

2. Let $\mathcal{A} = R$. We shall say a loss function is *convex*, if $l(\theta, \alpha a_0 + (1-\alpha)a_1) \leq \alpha l(\theta, a_0) + (1-\alpha) l(\theta, a_1)$, for any $a_0, a_1, \theta, 0 < \alpha < 1$. Suppose that there is an unbiased estimate δ of $q(\theta)$ and that $T(\mathbf{X})$ is sufficient. Show that if $l(\theta, a)$ is convex and $\delta^*(\mathbf{X}) = E(\delta(\mathbf{X}) \mid t(\mathbf{X}))$, then $R(\theta, \delta^*) \leq R(\theta, \delta)$.

Hint: Use *Jensen's inequality:* If g is a convex function and X is a random variable, then $E(g(X)) \geq g(E(X))$.

3. *Equivariance.* Let $X \sim p(x, \theta)$ with $\theta \in \Theta \subset R$, suppose that assumptions I and II hold and that h is a monotone increasing differentiable function from Θ onto $h(\Theta)$. Reparametrize the model by setting $\eta = h(\theta)$ and let $q(x, \eta) = p(x, h^{-1}(\eta))$ denote the model in the new parametrization.

(a) Show that if $I_p(\theta)$ and $I_q(\eta)$ denote the Fisher information in the two parametrizations, then

$$I_q(\eta) = I_p(h^{-1}(\eta))/[h'(h^{-1}(\eta))]^2.$$

That is, Fisher information is not equivariant under increasing transformations of the parameter.

(b) *Equivariance of the Fisher Information Bound.* Let $B_p(\theta)$ and $B_q(\eta)$ denote the information inequality lower bound $(\psi')^2/I$ as in (3.4.12) for the two parametrizations $p(x, \theta)$ and $q(x, \eta)$. Show that $B_q(\eta) = B_p(h^{-1}(\eta))$; that is, the Fisher information lower bound is equivariant.

4. Prove Proposition 3.4.4.

5. Suppose $X_1, \dots, X_n$ are i.i.d. $\mathcal{N}(\mu, \sigma^2)$ with $\mu - \mu_0$ known. Show that

(a) $\hat{\sigma}_0^2 = n^{-1} \sum_{i=1}^n (X_i - \mu_0)^2$ is a UMVU estimate of σ^2.

(b) $\hat{\sigma}_0^2$ is inadmissible.

Hint: See Problem 3.3.5(b).

(c) if μ_0 is not known and the true distribution of X_i is $\mathcal{N}(\mu, \sigma^2)$, $\mu \neq \mu_0$, find the bias of $\widehat{\sigma}_0^2$.

6. Show that assumption I implies that if $A \equiv \{x : p(\mathbf{x}, \theta) > 0\}$ doesn't depend on θ, then for any set B, $P_\theta(B) = 1$ for some θ if and only if $P_\theta(B) = 1$ for all θ.

7. In Example 3.4.4, compute $\text{Var}(\widehat{\theta})$ using each of the three methods indicated.

8. Establish the claims of Example 3.4.8.

9. Show that $S^2 = (\mathbf{Y} - \mathbf{Z}_D\widehat{\boldsymbol{\beta}})^T(\mathbf{Y} - \mathbf{Z}_D\widehat{\boldsymbol{\beta}})/(n-p)$ is an unbiased estimate of σ^2 in the linear regression model of Section 2.2.

10. Suppose $\widehat{\theta}$ is UMVU for estimating θ. Let a and b be constants. Show that $\widehat{\lambda} = a + b\widehat{\theta}$ is UMVU for estimating $\lambda = a + b\theta$.

11. Suppose $Y_1, \dots, Y_n$ are independent Poisson random variables with $E(Y_i) = \mu_i$ where $\mu_i = \exp\{\alpha + \beta z_i\}$ depends on the levels z_i of a covariate; $\alpha, \beta \in R$. For instance, z_i could be the level of a drug given to the ith patient with an infectious disease and Y_i could denote the number of infectious agents in a given unit of blood from the ith patient 24 hours after the drug was administered.

(a) Write the model for $Y_1, \dots, Y_n$ in two-parameter canonical exponential form and give the sufficient statistic.

(b) Let $\boldsymbol{\theta} = (\alpha, \beta)^T$. Compute $I(\boldsymbol{\theta})$ for the model in (a) and then find the lower bound on the variances of unbiased estimators $\widehat{\alpha}$ and $\widehat{\beta}$ of α and β.

(c) Suppose that $z_i = \log[i/(n+1)]$, $i = 1, \dots, n$. Find $\lim n^{-1}I(\boldsymbol{\theta})$ as $n \to \infty$, and give the limit of n times the lower bound on the variances of $\widehat{\alpha}$ and $\widehat{\beta}$.

Hint: Use the integral approximation to sums.

12. Let $X_1, \dots, X_n$ be a sample from the beta, $\mathcal{B}(\theta, 1)$, distribution.

(a) Find the MLE of $1/\theta$. Is it unbiased? Does it achieve the information inequality lower bound?

(b) Show that $\bar{X}$ is an unbiased estimate of $\theta/(\theta+1)$. Does $\bar{X}$ achieve the information inequality lower bound?

13. Let $\mathcal{F}$ denote the class of densities with mean θ^{-1} and variance $\theta^{-2}(\theta > 0)$ that satisfy the conditions of the information inequality. Show that a density that minimizes the Fisher information over $\mathcal{F}$ is $f(x, \theta) = \theta e^{-\theta x} 1(x > 0)$.

Hint: Consider $T(X) = X$ in Theorem 3.4.1.

14. Show that if $(X_1, \dots, X_n)$ is a sample drawn without replacement from an unknown finite population $\{x_1, \dots, x_N\}$, then

(a) $\bar{X}$ is an unbiased estimate of $\bar{x} = \frac{1}{N}\sum_{i=1}^N x_i$.