1. Suppose that random variables $Y_1, ..., Y_n$ satisfy

\[ Y_i = \beta x_i + \varepsilon_i, \quad i = 1, ..., n, \]

where $x_1, ..., x_n$ are fixed constants and $\varepsilon_i$ are iid $N(0, \sigma^2)$, $\sigma^2$ unknown. (i) Find a two-dimensional sufficient statistic for $(\beta, \sigma^2)$. (ii) Find the ML estimator of $\beta$ and show that it is unbiased. (iii) Find the distribution of the MLE of $\beta$.

Since $Y_i \sim N(\beta x_i, \sigma^2)$, we have

\[
f(y; \beta, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi \sigma}} \exp \left\{ -\frac{(y_i - \beta x_i)^2}{2\sigma^2} \right\} = \left( \frac{1}{\sqrt{2\pi \sigma}} \right)^n \exp \left\{ -\frac{\sum_{i=1}^{n} y_i^2 - 2\beta \sum_{i=1}^{n} x_i y_i + \beta^2 \sum_{i=1}^{n} x_i^2}{2\sigma^2} \right\},
\]

and hence, by the Factorization Theorem, $(\sum_{i=1}^{n} Y_i^2, \sum_{i=1}^{n} X_i Y_i)$ is a sufficient statistic for $(\beta, \sigma^2)$.

The ML estimator $\hat{\beta}$ is the minimizer of the function

\[
-\log f(y; \beta, \sigma^2) = c_1 + c_2 \sum_{i=1}^{n} (Y_i - \beta x_i)^2,
\]

where $c_2 > 0$ and $c_1$ are constants independent of $\beta$. By minimizing the function $\sum_{i=1}^{n} (Y_i - \beta x_i)^2$ we obtain

\[
\hat{\beta} = \sum_{i=1}^{n} \alpha_i Y_i,
\]

where $\alpha_i = x_i / (\sum_{i=1}^{n} x_i^2)$.

It follows that $\hat{\beta}$ has normal distribution with mean

\[
E[\hat{\beta}] = \sum_{i=1}^{n} \alpha_i E[Y_i] = \beta \sum_{i=1}^{n} \alpha_i = \beta,
\]

and variance

\[
Var[\hat{\beta}] = \sum_{i=1}^{n} \alpha_i^2 Var[Y_i] = \sigma^2 \sum_{i=1}^{n} \alpha_i^2 = \frac{\sigma^2}{\sum_{i=1}^{n} x_i^2}.
\]

2. Let $X_1, ..., X_n$ be an iid random sample from a population with mean $\theta$ and finite variance $\sigma^2 > 0$. Consider estimating $\theta$ using squared error loss.

(i) Show that any estimator of the form $a \bar{X} + b$, where $a > 1$ and $b$ are constants, is inadmissible. (ii) Show that if $a = 1$ and $b \neq 0$, then the estimator $a \bar{X} + b$ is inadmissible.
Because of the squared error loss function we have here that

\[ R(\theta, \delta) = Var(\delta) + (\mathbb{E}[\delta] - \theta)^2. \]

For \( \delta(X) = aX + b \) this becomes:

\[ R(\theta, \delta) = a^2 \sigma^2 / n + [\theta(a - 1) + b]^2. \]

(1)

It follows that the estimator \( \bar{X} \) is better than \( a\bar{X} + b \) in both cases if \( a > 1 \) and if \( a = 1 \) and \( b \neq 0 \).

(iii) Show that if \( X_i \sim N(\theta, \sigma^2) \), \( \sigma^2 \) is known, then \( a\bar{X} + b \) is admissible if \( 1 > a \geq 0 \).

Suppose that \( 1 > a > 0 \). Consider the prior distribution \( \pi(\theta) \sim N(\mu, \tau^2) \). The Bayes rule, with respect to \( \pi \), is

\[ \delta^\pi(x) = \mathbb{E}(\theta|x) = \frac{\tau^2}{\tau^2 + \sigma^2/n} \bar{x} + \frac{\sigma^2/n}{\tau^2 + \sigma^2/n} \mu. \]

Now choose \( \tau \) and \( \mu \) in such a way that

\[ a = \frac{\tau^2}{\tau^2 + \sigma^2/n} \quad \text{and} \quad b = \frac{\sigma^2/n}{\tau^2 + \sigma^2/n} \mu. \]

Note that because \( a \in (0, 1) \), it is possible to solve the above equations. For such prior we have that \( \delta^\pi(x) = a\bar{x} + b \). Also here \( R(\theta, \delta) \) is a continuous function of \( \theta \). It follows that \( \delta^\pi(x) \) is admissible.

Now for \( a = 0 \) the estimator \( \delta(X) = a\bar{X} + b \) becomes \( \delta(X) = b \), and hence \( R(b, \delta) = 0 \). It follows that if an estimator \( \delta'(X) \) is as good as \( \delta(X) \), then \( R(b, \delta') = 0 \). That is,

\[ Var(\delta') + (\mathbb{E}[\delta'] - b)^2 = 0. \]

It follows that \( Var(\delta') = 0 \) and \( \mathbb{E}[\delta'] = b \). Consequently \( R(\theta, \delta') = (b - \theta)^2 = R(\theta, \delta) \) for all \( \theta \), and hence \( \delta' \) is not better than \( \delta \). This shows that \( \delta \) is admissible.

3. Let \( \pi_n, n = 1, 2, ..., \) be a sequence of prior distributions. Let \( \delta_n \) denote a Bayes rule with respect to \( \pi_n \). Show that if the sequence

\[ B(\pi_n, \delta_n) = \int R(\theta, \delta_n) \pi_n(\theta) d\theta \]

converges to a number \( c \) and if \( \delta \) is a decision rule with \( R(\theta, \delta) \leq c \) for all \( \theta \in \Theta \), then \( \delta \) is minimax.

Consider a decision rule \( \delta' \). We have

\[ \sup_{\theta' \in \Theta} R(\theta', \delta') = [\sup_{\theta' \in \Theta} R(\theta', \delta')] \int \pi_n(\theta) d\theta = \int [\sup_{\theta' \in \Theta} R(\theta', \delta')] \pi_n(\theta) d\theta \geq \int R(\theta, \delta') \pi_n(\theta) d\theta \geq \int R(\theta, \delta_n) \pi_n(\theta) d\theta, \]

2
where the last inequality holds since $\delta_n$ is a Bayes rule with respect to $\pi_n$. By passing to the limit we obtain

$$\sup_{\theta \in \Theta} R(\theta, \delta') \geq c \geq R(\theta, \delta) \quad \text{for all } \theta \in \Theta,$$

and hence

$$\sup_{\theta \in \Theta} R(\theta, \delta') \geq \sup_{\theta \in \Theta} R(\theta, \delta).$$

This shows that $\delta$ is minimax.

4. Let $X_1, \ldots, X_n$ be an iid random sample from a $N(\theta, \sigma^2)$ population, $\sigma^2$ known. Consider estimating $\theta$ using squared error loss. Show that $\bar{X}$ is a minimax estimator (hint: use the result of the previous problem).

Consider a sequence of prior distributions $\pi_n(\theta) \sim N(0, \tau_n^2)$, with $\tau_n \to \infty$. Let $\delta_n$ be a Bayes rule with respect to $\pi_n$. Then we know that

$$B(\pi_n, \delta_n) = \frac{\sigma^2 \tau_n^2}{\sigma^2 + n\tau_n^2},$$

and hence $B(\pi_n, \delta_n) \to \sigma^2 / n$ as $n \to \infty$. We also have that $R(\theta, \bar{X}) = \text{Var}(\bar{X}) = \sigma^2 / n$. By the result of the previous question it follows that $\bar{X}$ is a minimax estimator.