Consider estimators of a real-valued $\vartheta = g(\theta)$ based on a sample $X$ from $P_\theta$, $\theta \in \Theta$, under loss $L$ and risk $R_T(\theta) = E[L(T(X), \theta)]$.

### Minimax estimator

A *minimax estimator* minimizes $\sup_{\theta \in \Theta} R_T(\theta)$ over all estimators $T$.

### Discussion

- A minimax estimator can be very conservative and unsatisfactory. It tries to do as well as possible in the worst case.
- A unique minimax estimator is admissible, since any estimator better than a minimax estimator is also minimax.
- We should find an admissible minimax estimator.
- Different for UMVUE: if a UMVUE is inadmissible, it is dominated by a biased estimator.
- If a minimax estimator has some other good properties (e.g., it is a Bayes estimator), then it is often a reasonable estimator.
Consider estimators of a real-valued $\vartheta = g(\theta)$ based on a sample $X$ from $P_\theta$, $\theta \in \Theta$, under loss $L$ and risk $R_T(\theta) = E[L(T(X), \theta)]$.

**Minimax estimator**

A *minimax estimator* minimizes $\sup_{\theta \in \Theta} R_T(\theta)$ over all estimators $T$.

**Discussion**

- A minimax estimator can be very conservative and unsatisfactory. It tries to do as well as possible in the worst case.
- A unique minimax estimator is admissible, since any estimator better than a minimax estimator is also minimax.
- We should find an admissible minimax estimator.
- Different for UMVUE: if a UMVUE is inadmissible, it is dominated by a biased estimator.
- If a minimax estimator has some other good properties (e.g., it is a Bayes estimator), then it is often a reasonable estimator.
Consider estimators of a real-valued $\vartheta = g(\theta)$ based on a sample $X$ from $P_\theta$, $\theta \in \Theta$, under loss $L$ and risk $R_T(\theta) = E[L(T(X), \theta)]$.

**Minimax estimator**

A *minimax estimator* minimizes $\sup_{\theta \in \Theta} R_T(\theta)$ over all estimators $T$.

**Discussion**

- A minimax estimator can be very conservative and unsatisfactory. It tries to do as well as possible in the worst case.
- A unique minimax estimator is admissible, since any estimator better than a minimax estimator is also minimax.
- We should find an admissible minimax estimator.
- Different for UMVUE: if a UMVUE is inadmissible, it is dominated by a biased estimator.
- If a minimax estimator has some other good properties (e.g., it is a Bayes estimator), then it is often a reasonable estimator.
How to find a minimax estimator?

Candidates for minimax: estimators having constant risks.

Theorem 4.11 (minimaxity of a Bayes estimator)

A limit of Bayes estimators

In many cases a constant risk estimator is not a Bayes estimator (e.g., an unbiased estimator under the squared error loss), but a limit of Bayes estimators w.r.t. a sequence of priors. The next result may be used to find a minimax estimator.

Theorem 4.12

Let $\Pi_j, j = 1, 2, \ldots$, be a sequence of priors and $r_j$ be the Bayes risk of a Bayes estimator of $\vartheta$ w.r.t. $\Pi_j$. Let $T$ be a constant risk estimator of $\vartheta$. If $\liminf_j r_j \geq R_T$, then $T$ is minimax.

Although Theorem 4.12 is more general than Theorem 4.11 in finding minimax estimators, it does not provide uniqueness of the minimax estimator even when there is a unique Bayes estimator w.r.t. each $\Pi_j$. 
How to find a minimax estimator?

Candidates for minimax: estimators having constant risks.

Theorem 4.11 (minimaxity of a Bayes estimator)

A limit of Bayes estimators

In many cases a constant risk estimator is not a Bayes estimator (e.g., an unbiased estimator under the squared error loss), but a limit of Bayes estimators w.r.t. a sequence of priors. The next result may be used to find a minimax estimator.

Theorem 4.12

Let $\Pi_j$, $j = 1, 2, \ldots$, be a sequence of priors and $r_j$ be the Bayes risk of a Bayes estimator of $\vartheta$ w.r.t. $\Pi_j$. Let $T$ be a constant risk estimator of $\vartheta$. If $\liminf_j r_j \geq R_T$, then $T$ is minimax.

Although Theorem 4.12 is more general than Theorem 4.11 in finding minimax estimators, it does not provide uniqueness of the minimax estimator even when there is a unique Bayes estimator w.r.t. each $\Pi_j$. 
How to find a minimax estimator?
Candidates for minimax: estimators having constant risks.

Theorem 4.11 (minimaxity of a Bayes estimator)

A limit of Bayes estimators
In many cases a constant risk estimator is not a Bayes estimator (e.g., an unbiased estimator under the squared error loss), but a limit of Bayes estimators w.r.t. a sequence of priors.
The next result may be used to find a minimax estimator.

Theorem 4.12
Let $\Pi_j, j = 1, 2, \ldots$, be a sequence of priors and $r_j$ be the Bayes risk of a Bayes estimator of $\vartheta$ w.r.t. $\Pi_j$.
Let $T$ be a constant risk estimator of $\vartheta$.
If $\liminf_j r_j \geq R_T$, then $T$ is minimax.

Although Theorem 4.12 is more general than Theorem 4.11 in finding minimax estimators, it does not provide uniqueness of the minimax estimator even when there is a unique Bayes estimator w.r.t. each $\Pi_j$. 
Example 2.25

Let $X_1, \ldots, X_n$ be i.i.d. components having the $N(\mu, \sigma^2)$ distribution with an unknown $\mu = \theta \in \mathbb{R}$ and a known $\sigma^2$.

If the prior is $N(\mu_0, \sigma_0^2)$, then the posterior of $\theta$ given $X = x$ is $N(\mu_*(x), \sigma^2)$ with

$$
\mu_*(x) = \frac{\sigma^2}{n\sigma_0^2 + \sigma^2} \mu_0 + \frac{n\sigma_0^2}{n\sigma_0^2 + \sigma^2} \bar{x}
$$

and

$$
c^2 = \frac{\sigma_0^2 \sigma^2}{n\sigma_0^2 + \sigma^2}
$$

We now show that $\bar{X}$ is minimax under the squared error loss.

For any decision rule $T$,

$$
\sup_{\theta \in \mathbb{R}} R_T(\theta) \geq \int_{\mathbb{R}} R_T(\theta) d\Pi(\theta) \geq \int_{\mathbb{R}} R_{\mu_*}(\theta) d\Pi(\theta)
$$

$$
= E \left\{ \left[ \bar{\theta} - \mu_*(X) \right]^2 \right\} = E \left\{ E \left\{ \left[ \bar{\theta} - \mu_*(X) \right]^2 | X \right\} \right\} = E(c^2) = c^2.
$$

Since this result is true for any $\sigma_0^2 > 0$ and $c^2 \to \sigma^2/n$ as $\sigma_0^2 \to \infty$,

$$
\sup_{\theta \in \mathbb{R}} R_T(\theta) \geq \frac{\sigma^2}{n} = \sup_{\theta \in \mathbb{R}} R_{\bar{X}}(\theta),
$$
Example 2.25 (continued)

where the equality holds because the risk of $\bar{X}$ under the squared error loss is $\sigma^2/n$ and independent of $\theta = \mu$. Thus, $\bar{X}$ is minimax.

To discuss the minimaxity of $\bar{X}$ in the case where $\sigma^2$ is unknown, we need the following lemma.

Lemma 4.3

Let $\Theta_0$ be a subset of $\Theta$ and $T$ be a minimax estimator of $\vartheta$ when $\Theta_0$ is the parameter space. Then $T$ is a minimax estimator if

$$\sup_{\theta \in \Theta} R_T(\theta) = \sup_{\theta \in \Theta_0} R_T(\theta).$$

Proof

If there is an estimator $T_0$ with $\sup_{\theta \in \Theta} R_{T_0}(\theta) < \sup_{\theta \in \Theta} R_T(\theta)$, then

$$\sup_{\theta \in \Theta_0} R_{T_0}(\theta) < \sup_{\theta \in \Theta} R_T(\theta),$$

which contradicts the minimaxity of $T$ when $\Theta_0$ is the parameter space. Hence, $T$ is minimax when $\Theta$ is the parameter space.
Example 2.25 (continued)

where the equality holds because the risk of $\bar{X}$ under the squared error loss is $\sigma^2 / n$ and independent of $\theta = \mu$.

Thus, $\bar{X}$ is minimax.

To discuss the minimaxity of $\bar{X}$ in the case where $\sigma^2$ is unknown, we need the following lemma.

Lemma 4.3

Let $\Theta_0$ be a subset of $\Theta$ and $T$ be a minimax estimator of $\varphi$ when $\Theta_0$ is the parameter space. Then $T$ is a minimax estimator if

$$\sup_{\theta \in \Theta} R_T(\theta) = \sup_{\theta \in \Theta_0} R_T(\theta).$$

Proof

If there is an estimator $T_0$ with $\sup_{\theta \in \Theta} R_{T_0}(\theta) < \sup_{\theta \in \Theta} R_T(\theta)$, then

$$\sup_{\theta \in \Theta_0} R_{T_0}(\theta) \leq \sup_{\theta \in \Theta} R_{T_0}(\theta) < \sup_{\theta \in \Theta} R_T(\theta) = \sup_{\theta \in \Theta_0} R_T(\theta),$$

which contradicts the minimaxity of $T$ when $\Theta_0$ is the parameter space. Hence, $T$ is minimax when $\Theta$ is the parameter space.
Example 2.25 (continued)

where the equality holds because the risk of $\bar{X}$ under the squared error loss is $\sigma^2/n$ and independent of $\theta = \mu$.
Thus, $\bar{X}$ is minimax.

To discuss the minimaxity of $\bar{X}$ in the case where $\sigma^2$ is unknown, we need the following lemma.

Lemma 4.3

Let $\Theta_0$ be a subset of $\Theta$ and $T$ be a minimax estimator of $\vartheta$ when $\Theta_0$ is the parameter space. Then $T$ is a minimax estimator if

$$\sup_{\theta \in \Theta} R_T(\theta) = \sup_{\theta \in \Theta_0} R_T(\theta).$$

Proof

If there is an estimator $T_0$ with $\sup_{\theta \in \Theta} R_{T_0}(\theta) < \sup_{\theta \in \Theta} R_T(\theta)$, then

$$\sup_{\theta \in \Theta_0} R_{T_0}(\theta) \leq \sup_{\theta \in \Theta} R_{T_0}(\theta) < \sup_{\theta \in \Theta} R_T(\theta) = \sup_{\theta \in \Theta_0} R_T(\theta),$$

which contradicts the minimaxity of $T$ when $\Theta_0$ is the parameter space. Hence, $T$ is minimax when $\Theta$ is the parameter space.
Example 2.25 (continued)

where the equality holds because the risk of $\tilde{X}$ under the squared error loss is $\sigma^2/n$ and independent of $\theta = \mu$.

Thus, $\tilde{X}$ is minimax.

To discuss the minimaxity of $\tilde{X}$ in the case where $\sigma^2$ is unknown, we need the following lemma.

Lemma 4.3

Let $\Theta_0$ be a subset of $\Theta$ and $T$ be a minimax estimator of $\vartheta$ when $\Theta_0$ is the parameter space. Then $T$ is a minimax estimator if

$$\sup_{\theta \in \Theta} R_T(\theta) = \sup_{\theta \in \Theta_0} R_T(\theta).$$

Proof

If there is an estimator $T_0$ with $\sup_{\theta \in \Theta} R_{T_0}(\theta) < \sup_{\theta \in \Theta} R_T(\theta)$, then

$$\sup_{\theta \in \Theta_0} R_{T_0}(\theta) \leq \sup_{\theta \in \Theta} R_{T_0}(\theta) < \sup_{\theta \in \Theta} R_T(\theta) = \sup_{\theta \in \Theta_0} R_T(\theta),$$

which contradicts the minimaxity of $T$ when $\Theta_0$ is the parameter space. Hence, $T$ is minimax when $\Theta$ is the parameter space.
Example 4.19

Let $X_1, \ldots, X_n$ be i.i.d. from $N(\mu, \sigma^2)$ with unknown $\theta = (\mu, \sigma^2)$. Consider the estimation of $\mu$ under the squared error loss.

Suppose first that $\Theta = \mathbb{R} \times (0, c]$ with a constant $c > 0$.

Let $\Theta_0 = \mathbb{R} \times \{c\}$.

From Example 2.25, $\bar{X}$ is a minimax estimator of $\mu$ when the parameter space is $\Theta_0$.

By Lemma 4.3, $\bar{X}$ is also minimax when the parameter space is $\Theta$.

Although $\sigma^2$ is assumed to be bounded by $c$, the minimax estimator $\bar{X}$ does not depend on $c$.

Consider next the case where $\Theta = \mathbb{R} \times (0, \infty)$, i.e., $\sigma^2$ is unbounded.

Let $T$ be any estimator of $\mu$. For any fixed $\sigma^2$,

$$\frac{\sigma^2}{n} \leq \sup_{\mu \in \mathbb{R}} R_T(\theta),$$

since $\sigma^2/n$ is the risk of $\bar{X}$ that is minimax when $\sigma^2$ is known.

Letting $\sigma^2 \to \infty$, we obtain that $\sup_{\theta} R_T(\theta) = \infty$ for any estimator $T$.

Thus, minimaxity is meaningless (any estimator is minimax).
Theorem 4.14 (Admissibility in one-parameter exponential families)

Suppose that $X$ has the p.d.f. $c(\theta)e^{\theta T(x)}$ w.r.t. a $\sigma$-finite measure $\nu$, where $T(x)$ is real-valued and $\theta \in (\theta_-, \theta_+ \subset \mathbb{R}$.

Consider the estimation of $\varphi = E[T(X)]$ under the squared error loss. Let $\lambda \geq 0$ and $\gamma$ be known constants and let

$$T_{\lambda, \gamma}(X) = (T + \gamma \lambda)/(1 + \lambda).$$

Then a sufficient condition for the admissibility of $T_{\lambda, \gamma}$ is that

$$\int_{\theta_0}^{\theta_+} \frac{e^{-\gamma \lambda \theta}}{[c(\theta)]^\lambda} d\theta = \int_{\theta_-}^{\theta_0} \frac{e^{-\gamma \lambda \theta}}{[c(\theta)]^\lambda} d\theta = \infty,$$

where $\theta_0 \in (\theta_-, \theta_+)$. 
Remarks

- Theorem 4.14 provides a class of admissible estimators.

- The reason why \( T_{\lambda, \gamma} \) is considered is that it is often a Bayes estimator w.r.t. some prior; see Examples 2.25, 4.1, 4.7, and 4.8.

- Using this theorem and Theorem 4.13, we can obtain a class of minimax estimators.

- Although the proof of this theorem is more complicated than that of Theorem 4.3, the application of Theorem 4.14 is typically easier.

- To find minimax estimators, we may use the following result.

Corollary 4.3

Assume that \( X \) has the p.d.f. as described in Theorem 4.14 with \( \theta_- = -\infty \) and \( \theta_+ = \infty \).

(i) As an estimator of \( \vartheta = E(T) \), \( T(X) \) is admissible under the squared error loss and the loss \( (a - \vartheta)^2 / \text{Var}(T) \).

(ii) \( T \) is the unique minimax estimator of \( \vartheta \) under the loss \( (a - \vartheta)^2 / \text{Var}(T) \).
Remarks

- Theorem 4.14 provides a class of admissible estimators.
- The reason why $T_{\lambda,\gamma}$ is considered is that it is often a Bayes estimator w.r.t. some prior; see Examples 2.25, 4.1, 4.7, and 4.8.
- Using this theorem and Theorem 4.13, we can obtain a class of minimax estimators.
- Although the proof of this theorem is more complicated than that of Theorem 4.3, the application of Theorem 4.14 is typically easier.
- To find minimax estimators, we may use the following result.

Corollary 4.3

Assume that $X$ has the p.d.f. as described in Theorem 4.14 with $\theta_- = -\infty$ and $\theta_+ = \infty$.

(i) As an estimator of $\vartheta = E(T)$, $T(X)$ is admissible under the squared error loss and the loss $(a - \vartheta)^2 / \text{Var}(T)$.

(ii) $T$ is the unique minimax estimator of $\vartheta$ under the loss $(a - \vartheta)^2 / \text{Var}(T)$. 
Example 4.20

Let $X_1, \ldots, X_n$ be i.i.d. from $N(0, \sigma^2)$ with an unknown $\sigma^2 > 0$ and let $Y = \sum_{i=1}^n X_i^2$.

Consider the estimation of $\sigma^2$.

The risk of $Y/(n+2)$ is a constant under the loss $(a - \sigma^2)^2/\sigma^4$.

We now apply Theorem 4.14 to show that $Y/(n+2)$ is admissible.

Note that the joint p.d.f. of $X_i$’s is of the form $c(\theta)e^{\theta T(X)}$ with $\theta = -n/(4\sigma^2)$, $c(\theta) = (-2\theta/n)^{n/2}$, $T(X) = 2Y/n$, $\theta_- = -\infty$, and $\theta_+ = 0$.

By Theorem 4.14, $T_{\lambda,\gamma} = (T + \gamma\lambda)/(1 + \lambda)$ is admissible under the squared error loss if, for some $c > 0$,

$$\int_{-\infty}^{-c} e^{-\gamma\lambda\theta} \left(-\frac{2\theta}{n}\right)^{-n\lambda/2} d\theta = \int_{0}^{c} e^{\gamma\lambda\theta} \theta^{-n\lambda/2} d\theta = \infty$$

This means that $T_{\lambda,\gamma}$ is admissible if $\gamma = 0$ and $\lambda = 2/n$, or if $\gamma > 0$ and $\lambda \geq 2/n$.

In particular, $2Y/(n+2)$ is admissible for estimating $E(T) = 2E(Y)/n = 2\sigma^2$, under the squared error loss.
Example 4.20 (continued)

It is easy to see that \( Y/(n+2) \) is then an admissible estimator of \( \sigma^2 \) under the squared error loss and the loss \( (a - \sigma^2)^2/\sigma^4 \).
Hence \( Y/(n+2) \) is minimax under the loss \( (a - \sigma^2)^2/\sigma^4 \).
Note that we cannot apply Corollary 4.3 directly since \( \theta_+ = 0 \).

Example 4.21

Let \( X_1, \ldots, X_n \) be i.i.d. from the Poisson distribution \( P(\theta) \) with an unknown \( \theta > 0 \).
The joint p.d.f. of \( X_i \)'s w.r.t. the counting measure is

\[
(x_1! \cdots x_n!)^{-1} e^{-n\theta} e^{n\bar{x}\log\theta}
\]

For \( \eta = n\log\theta \), the conditions of Corollary 4.3 are satisfied with \( T(X) = \bar{X} \).
Since \( E(T) = \theta \) and \( \text{Var}(T) = \theta/n \), by Corollary 4.3, \( \bar{X} \) is the unique minimax estimator of \( \theta \) under the loss function \( (a - \theta)^2/\theta \).
Example 4.20 (continued)

It is easy to see that \( Y/(n+2) \) is then an admissible estimator of \( \sigma^2 \) under the squared error loss and the loss \( (a - \sigma^2)^2/\sigma^4 \).
Hence \( Y/(n+2) \) is minimax under the loss \( (a - \sigma^2)^2/\sigma^4 \).
Note that we cannot apply Corollary 4.3 directly since \( \theta_+ = 0 \).

Example 4.21

Let \( X_1, \ldots, X_n \) be i.i.d. from the Poisson distribution \( P(\theta) \) with an unknown \( \theta > 0 \).
The joint p.d.f. of \( X_i \)'s w.r.t. the counting measure is

\[
(\prod_{i=1}^{n} x_i!)^{-1} e^{-n\theta} e^{n\bar{x}\log\theta}
\]

For \( \eta = n\log\theta \), the conditions of Corollary 4.3 are satisfied with \( T(X) = \bar{X} \).
Since \( E(T) = \theta \) and \( \text{Var}(T) = \theta/n \), by Corollary 4.3, \( \bar{X} \) is the unique
minimax estimator of \( \theta \) under the loss function \( (a - \theta)^2/\theta \).
Exercise 37 (#4.83)

Let $X$ be an observation from the distribution with Lebesgue density
\[
\frac{1}{2} c(\theta) e^{\theta x - |x|}, \quad |\theta| < 1.
\]
(i) Show that $c(\theta) = 1 - \theta^2$.
(ii) Show that if $0 \leq \alpha \leq \frac{1}{2}$, then $\alpha X + \beta$ is admissible for estimating $E(X)$ under the squared error loss.

Solution

(i) Note that
\[
\frac{1}{c(\theta)} = \frac{1}{2} \int_{-\infty}^{\infty} e^{\theta x - |x|} \, dx
\]
\[
= \frac{1}{2} \left( \int_{-\infty}^{0} e^{\theta x + x} \, dx + \int_{0}^{\infty} e^{\theta x - x} \, dx \right)
\]
\[
= \frac{1}{2} \left( \int_{-\infty}^{\infty} e^{-(1+\theta)x} \, dx + \int_{0}^{\infty} e^{-(1-\theta)x} \, dx \right)
\]
\[
= \frac{1}{2} \left( \frac{1}{1 + \theta} + \frac{1}{1 - \theta} \right) = \frac{1}{1 - \theta^2}.
\]
Exercise 37 (#4.83)

Let $X$ be an observation from the distribution with Lebesgue density
\[
\frac{1}{2} c(\theta) e^{\theta x - |x|}, \quad |\theta| < 1.
\]

(i) Show that $c(\theta) = 1 - \theta^2$.

(ii) Show that if $0 \leq \alpha \leq \frac{1}{2}$, then $\alpha X + \beta$ is admissible for estimating $E(X)$ under the squared error loss.

Solution

(i) Note that
\[
\frac{1}{c(\theta)} = \frac{1}{2} \int_{-\infty}^{\infty} e^{\theta x - |x|} \, dx
\]
\[
= \frac{1}{2} \left( \int_{-\infty}^{0} e^{\theta x + x} \, dx + \int_{0}^{\infty} e^{\theta x - x} \, dx \right)
\]
\[
= \frac{1}{2} \left( \int_{0}^{\infty} e^{-(1+\theta)x} \, dx + \int_{0}^{\infty} e^{-(1-\theta)x} \, dx \right)
\]
\[
= \frac{1}{2} \left( \frac{1}{1+\theta} + \frac{1}{1-\theta} \right) = \frac{1}{1-\theta^2}.
\]
(ii) Consider first $\alpha > 0$. Let $\alpha = (1 + \lambda)^{-1}$ and $\beta = \gamma \lambda / (1 + \lambda)$.

$$
\int_{-1}^{0} \frac{e^{-\gamma \lambda \theta}}{(1 - \theta^2)^{\lambda}} d\theta = \int_{0}^{1} \frac{e^{-\gamma \lambda \theta}}{(1 - \theta^2)^{\lambda}} d\theta = \infty
$$

if and only if $\lambda \geq 1$, i.e., $\alpha \leq \frac{1}{2}$.

Hence, $\alpha \lambda X + \beta$ is an admissible estimator of $E(X)$ when $0 < \alpha \leq \frac{1}{2}$.

Consider next $\alpha = 0$.

$$
E(X) = \frac{1 - \theta^2}{2} \left( \int_{-\infty}^{0} x e^{\theta x + x} dx + \int_{0}^{\infty} x e^{\theta x - x} dx \right)
$$

$$
= \frac{1 - \theta^2}{2} \left( - \int_{0}^{\infty} xe^{(1+\theta)x} dx + \int_{0}^{\infty} xe^{(1-\theta)x} dx \right)
$$

$$
= \frac{1 - \theta^2}{2} \left( \frac{1 + \theta}{1 - \theta} - \frac{1 - \theta}{1 + \theta} \right) = \frac{2\theta}{1 - \theta^2},
$$

which takes any value in $(-\infty, \infty)$.

Hence, the constant estimator $\beta$ is an admissible estimator of $E(X)$. 