0.1 Notation

- (model) $X \sim P_{\theta}$ with values in $X$, where $\theta \in \Omega$ is the parameter
- (loss and risk) $L(\theta, d)$ and $R(\theta, \delta) = E_{\theta}[L(\theta, \delta(X))]$, where $\delta(X)$ is a statistic
- (average risk) For a prior $\Lambda$ and estimator $\delta$, we denote by $r(\Lambda, \delta) = \int R(\theta, \delta)\lambda(d\theta)$, which is the average risk with respect to $\Lambda$. We also let $\delta^A$ denote a Bayes estimator (when one exists) and $r_\Lambda = r(\Lambda, \delta^A)$ the average risk of that estimator, also called the Bayes risk.
- (maximum risk) For an estimator $\delta$ we let $R(\delta) = \sup_{\theta \in \Omega} R(\theta, \delta)$, which is its maximum risk.

0.2 Equivariance setting

- We say that a transformation $g : X \rightarrow X$ leaves the model invariant if (i) it is one-to-one; (ii) if for any $\theta \in \Omega$, $X \sim P_{\theta}$ implies $gX \sim P_{\theta'}$ for some $\theta' \in \Omega$, which allows one to define $\tilde{g} : \Omega \rightarrow \Omega$ that associates $\theta'$ to $\theta$; (iii) $\tilde{g}$ is one-to-one (here we assume the model is identifiable).
- We work with a group $G$ of transformations, each leaving the model invariant, and define $G = \{\tilde{g} : g \in G\}$.
- We work with an estimand $h(\theta)$ that is equivariant with respect to $G$. This allows one to define $g^* \forall$ for each $g \in G$ such that $h(\tilde{g}\theta) = g^*h(\theta)$ for all $\theta \in \Omega$.
- We work with a loss function $L$ that is invariant in that $L(\tilde{g}\theta, g^*d) = L(\theta, d)$ for all $\theta$, all $d$, and all $g$.
- We then focus on estimators that are equivariant in the sense that $\delta(gx) = g^*\delta(x)$ for all $x$ and all $g$.

0.3 Location model

This model is of the form $X = (X_1, \ldots, X_n) \sim f(x - \xi)$ where $f$ is a given density on $\mathbb{R}^n$ and $\xi \in \mathbb{R}$ is unknown. (As usual, $x - \xi$ is understood coordinate-wise.) The transformations of interest are of the form $x \mapsto a + x$ where $a \in \mathbb{R}$. We want to estimate $\xi$ and work with a loss of the form $L(\xi, d) = \rho(d - \xi)$. Let $Y = (Y_1, \ldots, Y_{n-1})$ with $Y_i = X_i - X_n$.

**Theorem 1.** In the present setting, let $\delta_0$ be any equivariant statistic with finite risk, and suppose we may define $v^*(y) = \arg \min_{v \in \mathbb{R}} E_0[\rho(\delta_0(X) - v) | Y = y]$. Then $\delta^*(x) = \delta_0(x) - v^*(y)$ is MRE.

**Corollary 1.** Under squared error loss, the MRE may be expressed as $\delta^*(x) = \int_{-\infty}^{\infty} v f(x - v) dv / \int_{-\infty}^{\infty} f(x - v) dv$.

**Proposition 1.** Under squared error loss, if an UMVUE exists and is equivariant, then it is MRE.

0.4 Scale model

This model is of the form $X = (X_1, \ldots, X_n) \sim \tau^{-1}f(x/\tau)$ where $f$ is a given density on $\mathbb{R}^n$ and $\tau > 0$ is unknown. The transformations of interest are of the form $x \mapsto b x$ where $b > 0$. We want to estimate $\tau$ and work with a loss of the form $L(\tau, d) = \gamma(d/\tau)$. Let $Z = (Z_1, \ldots, Z_n)$ with $Z_i = X_i / X_n$ for $i \neq n$ and $Z_n = X_n / |X_n|$.

**Theorem 2.** In the present setting, let $\delta_0$ be any equivariant statistic with finite risk, and suppose we may define $w^*(z) = \arg \min_{w > 0} E_1[\gamma(\delta_0(X)/w) | Z = z]$. Then $\delta^*(x) = \delta_0(x)/w^*(z)$ is MRE.

**Corollary 2.** When $L(\tau, d) = (d/\tau - 1)^2$, the MRE may be expressed as $\delta^*(x) = \int_0^\infty w^* f(wx) dw / \int_0^\infty w^* f(wx) dw$.

0.5 Location-scale model

This model is of the form $X = (X_1, \ldots, X_n) \sim \tau^{-n}f((x - \xi)/\tau)$ where $f$ is a given density on $\mathbb{R}^n$ and $\xi \in \mathbb{R}$ and $\tau > 0$ are both unknown. The transformations of interest are of the form $x \mapsto a + bx$ where $a \in \mathbb{R}$ and $b > 0$. When our goal is to estimate $\xi$, we work with a loss of the form $L(\xi, \tau; d) = \rho((d - \xi)/\tau)$. When our goal is to estimate $\tau$, we work with a loss of the form $L(\xi, \tau; d) = \gamma(d/\tau)$.
0.6 Bayes estimation

B1 Under loss $L(\theta, d) = w(\theta)(d - h(\theta))^2$, the Bayes estimator is $\delta^\Lambda(x) = \mathbb{E}[w(\theta)h(\theta) | X = x] / \mathbb{E}[w(\theta) | X = x].$

B2 In a Bayesian setting, suppose the loss is strictly convex (in $d$) and that $Q$ denotes the marginal of $X$. Then the Bayes estimator is unique if the Bayes risk is finite and, for any measurable set $A$, $Q(A) = 0$ implies $P_{\theta}(A) = 0$ for all $\theta$.

0.7 Minimax estimation

M1 Suppose that $\Lambda$ is a prior such that $r_{\Lambda} = \bar{R}(\delta^\Lambda)$. Then $\delta^\Lambda$ is minimax, and uniquely so if it is unique Bayes.

M2 If an estimator is Bayes for some prior and has constant risk, it is minimax.

M3 If, for an estimator $\delta$, we can find a sequence of priors $(\Lambda_k)$ such that $\lim \inf_k r_{\Lambda_k} \geq \bar{R}(\delta)$, then $\delta$ is minimax.

M4 Consider $\Omega_0 \subset \Omega$. If an estimator is minimax over $\Omega_0$ and achieves its maximum risk at some $\theta \in \Omega_0$, then this estimator is also minimax over $\Omega$.

0.8 Admissibility

A1 A unique Bayes estimator is admissible.

A2 (Karlin’s theorem) Suppose $X \sim f_\theta$, where $f_\theta(x) = \beta(\theta)e^{\theta T(x)}$ with respect to some underlying measure. Let $\Omega = [\theta_*, \theta^*]$ denote the natural parameter space. Suppose $L(\theta, d) = (d - h(\theta))^2$, where $h(\theta) = \mathbb{E}_\theta(T)$. Then, for $a \geq 0$ and $b \in \mathbb{R}$, a sufficient condition for $\frac{1}{1+a}T + \frac{a}{1+a}b$ to be admissible is that $\int_{\theta_*}^{\theta^*} K(\theta)d\theta = \infty$ and $\int_{0}^{\theta^*} K(\theta)d\theta = \infty$, where $K(\theta) = e^{-bad(\theta)}\beta(\theta)^{-a}$.

A3 If an estimator has constant risk and is admissible, it is minimax.

A4 If an estimator is unique minimax, it is admissible.