Principles of Statistics	**Part II - Michaelmas 2018**

Example Sheet 3

Lecturer: Quentin Berthet

Throughout, for observations X arising from a parametric model on $\Theta \subseteq \mathbf{R}$ given by $\{f(\cdot, \theta) : \theta \in \Theta\}$,, the quadratic risk of a decision rule $\delta(X)$ is defined as the risk of the squared loss $R(\delta, \theta) = \mathbf{E}_\theta[(\delta(X) - \theta)^2]$.

1. Consider $X|\theta \sim \text{Poisson}(\theta), \theta \in \Theta = [0, \infty)$, and suppose the prior for θ is a Gamma distribution with parameters α, λ. Show that the posterior distribution $\theta|X$ is also a Gamma distribution and find its parameters.

2. For $n \in \mathbb{N}$ fixed, suppose X is binomially $\text{Bin}(n, \theta)$-distributed where $\theta \in [0, 1]$.

a) Consider a prior for θ from a $\text{Beta}(a, b), a, b > 0$, distribution. Show that the posterior distribution Π is $\text{Beta}(a + X, b + n - X)$ and compute the posterior mean given by $\bar{\theta}_n(X) = \mathbf{E}_\Pi[\theta|X]$.

b) Show that the maximum likelihood estimator for θ is *not* identical to the posterior mean with 'ignorant' uniform prior $\theta \sim U[0, 1]$.

c) Assuming that X is sampled from a fixed $\text{Bin}(n, \theta_0), \theta_0 \in (0, 1)$, distribution, derive the asymptotic distribution of $\sqrt{n}(\bar{\theta}_n(X) - \theta_0)$ as $n \to \infty$.

3. Let $X_1, \ldots, X_n$ be i.i.d. copies of a random variable X and consider the Bayesian model $X|\theta \sim \mathcal{N}(\theta, 1)$ with prior π as $\theta \sim \mathcal{N}(\mu, v^2)$. For $0 < \alpha < 1$, consider the credible interval

$$C_n = \{\theta \in \mathbb{R} : |\theta - \mathbf{E}_\Pi[\theta|X_1, \ldots, X_n]| \leq R_n\}$$

where R_n is chosen such that $\Pi(C_n|X_1, \ldots, X_n) = 1 - \alpha$. Now assume $X \sim N(\theta_0, 1)$ for some fixed $\theta_0 \in \mathbb{R}$, and show that, as $n \to \infty$, $P_{\theta_0}^{\mathbb{N}}(\theta_0 \in C_n) \to 1 - \alpha$.

4. In a general decision problem, show that a) a decision rule δ that has constant risk and is admissible is also minimax; b) any unique Bayes rule is admissible.

5. Consider an observation X from a parametric model $\{f(\cdot, \theta) : \theta \in \Theta\}$ with prior π on $\Theta \subseteq \mathbb{R}$ and a general risk function $R(\delta, \theta) = \mathbf{E}_\theta[L(\delta(X), \theta)]$. Assume that there exists some decision rule δ_0 such that $R(\delta_0, \theta) < \infty$ for all $\theta \in \Theta$.

a) For the loss function $L(a, \theta) = |a - \theta|$ show that the Bayes rule associated to π equals any median of the posterior distribution $\Pi(\cdot|X)$.

b) For weight function $w : \Theta \to [0, \infty)$ and loss function $L(a, \theta) = w(\theta)[a - \theta]^2$ show that the Bayes rule δ_π associated to π is unique and equals

$$\delta_\pi(X) = \frac{\mathbf{E}_\Pi[w(\theta)\theta|X]}{\mathbf{E}_\Pi[w(\theta)|X]},$$

assuming that expectations in the last ratio exist, are finite, and that $\mathbf{E}_\Pi[w(\theta)|X] > 0$.

6. a) Considering $X_1, \dots, X_n$ i.i.d. from a $\mathcal{N}(\theta, 1)$-model with $\theta \in \Theta = \mathbb{R}$, show that the maximum likelihood estimator is *not* a Bayes rule for estimating θ in quadratic risk for any prior distribution π.

b) Let $X \sim \text{Bin}(n, \theta)$ where $\theta \in \Theta = [0, 1]$. Find all prior distributions π on Θ for which the maximum likelihood estimator is a Bayes rule for estimating θ in quadratic risk.

7. Consider estimation of $\theta \in \Theta = [0, 1]$ in a $\text{Bin}(n, \theta)$ model under quadratic risk.

a) Find the unique minimax estimator $\tilde{\theta}_n$ of θ and deduce that the maximum likelihood estimator $\hat{\theta}_n$ of θ is *not* minimax for fixed sample size $n \in \mathbb{N}$. [Hint: Find first a Bayes rule of risk constant in $\theta \in \Theta$.]

b) Show, however, that the maximum likelihood estimator dominates $\tilde{\theta}_n$ in the large sample limit by proving that, as $n \to \infty$,

$$\lim_n \frac{\sup_\theta R(\hat{\theta}_n, \theta)}{\sup_\theta R(\tilde{\theta}_n, \theta)} = 1$$

and that

$$\lim_n \frac{R(\hat{\theta}_n, \theta)}{R(\tilde{\theta}_n, \theta)} < 1 \quad \text{for all } \theta \in [0, 1], \theta \neq \frac{1}{2}.$$

8. Consider $X_1, \dots, X_n$ i.i.d. from a $\mathcal{N}(\theta, 1)$-model where $\theta \in \Theta = [0, \infty)$. Show that the sample mean $\bar{X}_n$ is inadmissible for quadratic risk, but that it is still minimax. What happens if $\Theta = [a, b]$ for some $0 < a < b < \infty$?

9. Let X be multivariate normal $\mathcal{N}(\theta, I)$ where $\theta \in \Theta = \mathbb{R}^p, p \geq 3$, and where I is the $p \times p$ identity matrix. Consider estimators

$$\tilde{\theta}^{(c)} = \left(1 - c\frac{p-2}{\|X\|^2}\right) X, \ 0 < c < 2,$$

for θ, under the risk function $R(\delta, \theta) = \mathbf{E}_\theta \|\delta(X) - \theta\|^2$ where $\|\cdot\|$ is the standard Euclidean norm on $\mathbb{R}^p$.

a) Show that the James-Stein estimator $\tilde{\theta}^{(1)}$ dominates all estimators $\tilde{\theta}^{(c)}, c \neq 1$.

b) Let $\hat{\theta}$ be the maximum likelihood estimator of θ. Show that, for any $0 < c < 2$,

$$\sup_{\theta \in \Theta} R(\tilde{\theta}^{(c)}, \theta) = \sup_{\theta \in \Theta} R(\hat{\theta}, \theta).$$

10. Consider $X_1, \dots, X_n$ i.i.d. from a $\mathcal{N}(\theta, 1)$-model with $\theta \in \Theta = \mathbb{R}$ and recall the Hodges' estimator, equal to the maximum likelihood estimator $\bar{X}_n$ of θ whenever $|\bar{X}_n| \geq n^{-1/4}$ and zero otherwise. Recall the asymptotic distribution of $\sqrt{n}(\tilde{\theta}_n - \theta)$ as

$n \to \infty$ under P_θ for every $\theta \in \Theta$, and compare it to the asymptotic distribution of $\sqrt{n}(\bar{X}_n - \theta)$. Now compute the asymptotic maximal risk

$$\lim_{n} \sup_{\theta\in\Theta} \mathbf{E}_\theta\big[\big(\sqrt{n}(T_n - \theta)\big)^2\big]$$

for both $T_n = \bar{X}_n$ and $T_n = \tilde{\theta}_n$.