

Problem 1.27: Suppose that $X \sim f(x|\theta)$, with prior distribution $\pi(\theta)$, and interest is in the estimation of the parameter $h(\theta)$.

a) Using the loss function $L(\delta, h(\theta))$, show that the estimator that minimizes the Bayes risk

$\iint L(\delta, h(\theta)) f(x|\theta) \pi(\theta) dx d\theta$ is given by the estimator that minimizes (for each x)

$$\int L(\delta, h(\theta)) \pi(\theta|x) d\theta.$$

Solution: We know that $f(x|\theta)\pi(\theta) = \pi(\theta|x)M(x)$. Substituting this in the formula for the Bayes risk we get

$$\iint L(\delta, h(\theta)) f(x|\theta) \pi(\theta) dx d\theta = \iint L(\delta, h(\theta)) \pi(\theta|x) M(x) dx d\theta = \int \left(\int L(\delta, h(\theta)) \pi(\theta|x) d\theta \right) M(x) dx.$$

Thus δ minimizing the Bayes risk is given by the estimator that, for each x , minimizes the inner integral, i.e. δ minimizing $\int L(\delta, h(\theta)) \pi(\theta|x) d\theta$.

b) For $L(\delta, h(\theta)) = \|h(\theta) - \delta\|^2$, show that the Bayes estimator of $h(\theta)$ is

$$\delta^{\pi}(x) = E^{\pi}[h(\theta)|x].$$

Solution: According to part a), the Bayes estimator of $h(\theta)$ is the estimator δ that for each x minimizes $\int L(\delta, h(\theta)) \pi(\theta|x) d\theta = \int \|h(\theta) - \delta(x)\|^2 \pi(\theta|x) d\theta = E[\|h(\theta) - \delta(x)\|^2 | x]$ and it is known (by a result that will be proven at the end), that the estimator minimizing that expectation is $\delta^{\pi}(x) = E^{\pi}[h(\theta)|x]$.

c) For $L(\delta, h(\theta)) = |h(\theta) - \delta|$, show that the Bayes estimator of $h(\theta)$ is the median of the posterior distribution.

Solution: According to part a) again, the Bayes estimator of $h(\theta)$ is the estimator δ that for each x minimizes $\int L(\delta, h(\theta)) \pi(\theta|x) d\theta = \int |h(\theta) - \delta(x)| \pi(\theta|x) d\theta = E[|h(\theta) - \delta(x)| | x]$. But from a result that will be proven below, this expectation is minimized by the median of the posterior distribution (of $h(\theta)$).

Results:

Note: for simplicity, we will provide the proof in the unidimensional case (for the multidimensional case, the result follows from the definition of the norm and the linearity of expectation).

1) Let Y be a random variable and M a constant. The value that minimizes $E[(Y-M)^2]$ is $M = E[Y]$.

$$\text{Proof: } E[(Y-M)^2] = E[(Y-E(Y)+E(Y)-M)^2] = E[(Y-E(Y))^2] + 2E[(Y-E(Y))(E(Y)-M)] + E[(E(Y)-M)^2].$$

The middle term is 0 since $E(Y)-M$ is a constant and $E[(Y-E(Y))]=0$.

The other two terms are positive and the expression is minimized when the last term is 0 i.e. for $M = E(Y)$.

2) Let Y be a real valued random variable with a median M and $E|Y| < \infty$.

Then $E|Y-M| = \inf_{a \in \mathbb{R}} E|Y-a|$.

$$a \in \mathbb{R}$$

Proof: Note that $E|Y-a| \leq E(|Y| + |a|) < \infty$ and thus it makes sense to talk about the right hand

side.

Suppose first that $a > M$.

$$|Y-a|-|Y-M| = \begin{cases} M-a & \text{if } M < a \leq Y \\ M+a-2Y & \text{if } M < Y < a \\ a-M & \text{if } Y \leq M < a. \end{cases}$$

$$\begin{aligned} \text{Thus, } E[|Y-a|-|Y-M|] &= (M-a)P(Y \geq a) + \int_M^a (M+a-2y) dF(y) + (a-M)P(Y \leq M) \geq (M-a)P(Y \geq a) \\ &+ (M-a) \int_M^a dF(y) + (a-M)P(Y \leq M) = (M-a)P(Y \geq M) + (a-M)P(Y \leq M) = (a-M)[2P(Y \leq M) - 1] \\ &\geq (a-M)(2 \cdot 1/2 - 1) = 0. \end{aligned}$$

This implies that $E|Y-a| \geq E|Y-M|$ for $a > M$.

The case when $a < M$ is obtained similarly and thus we got the desired result.