Mathematical Statistics, Winter semester 2021/22
Solutions to Problem sheet 3

7) Show that the Hellinger affinity, and therefore the Hellinger distance as well, do not depend on the choice of a dominating σ-finite measure μ.

Hint: See the proof of Lemma 2.1. (Uniqueness of a maximum likelihood estimator)

Solution
Suppose that μ_1 and μ_2 are σ-finite measures such that $P_1, P_2 \ll \mu_1, \mu_2$. Hence, we have that

$$P_1, P_2 \ll \mu_1, \mu_2 \ll \mu_1 + \mu_2,$$

which means that P_1 and P_2 have densities w.r.t. μ_1, μ_2, and $\mu_1 + \mu_2$ such that

$$\frac{dP_i}{d(\mu_1+\mu_2)}(x) = \frac{dP_i}{d\mu_j}(x)\frac{d\mu_j}{d(\mu_1+\mu_2)}(x) \qquad (\mu_1+\mu_2)\text{a.e.}, \qquad \text{for } i,j = 1,2.$$

Now we obtain, for $i = 1, 2$, that

$$\begin{aligned}
\rho^{(i)}(P_1,P_2) &:= \int \sqrt{\frac{dP_1}{d\mu_i}(x)}\sqrt{\frac{dP_2}{d\mu_i}(x)}\, d\mu_i(x) \\
&= \int \sqrt{\frac{dP_1}{d\mu_i}(x)}\sqrt{\frac{dP_2}{d\mu_i}(x)}\, \frac{d\mu_i}{d(\mu_1+\mu_2)}(x)\, d(\mu_1+\mu_2)(x) \\
&= \int \sqrt{\frac{dP_1}{d\mu_i}(x)\frac{d\mu_i}{d(\mu_1+\mu_2)}(x)}\sqrt{\frac{dP_2}{d\mu_i}(x)\frac{d\mu_i}{d(\mu_1+\mu_2)}(x)}\, d(\mu_1+\mu_2)(x) \\
&= \int \sqrt{\frac{dP_1}{d(\mu_1+\mu_2)}(x)}\sqrt{\frac{dP_2}{d(\mu_1+\mu_2)}(x)}\, d(\mu_1+\mu_2)(x).
\end{aligned}$$

Therefore,

$$\rho^{(1)}(P_1,P_2) = \rho^{(2)}(P_1,P_2)$$

and

$$H^{(1)}(P_1,P_2) = \sqrt{1 - \rho^{(1)}(P_1,P_2)} = \sqrt{1 - \rho^{(2)}(P_1,P_2)} = H^{(2)}(P_1,P_2).$$

8) Let $X_1, \dots, X_n \sim \mathrm{Bin}(1, \theta)$ be independent random variables, $\theta \in \Theta := [0, 1]$.

Compute the maximum likelihood estimator of θ.

Solution

Let $X = (X_1, \dots, X_n)^T$ and let $x = (x_1, \dots, x_n)^T$ be a possible realization of X. Since X is a discrete random variable the likelihood function L is given by

$$L(\theta; x) \;=\; P_\theta(X = x) \;=\; \prod_{i=1}^{n} P_\theta(X_i = x_i) \;=\; \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}.$$

For all x, $L(\,\cdot\,; x)$ is a continuous function on the compact set $[0, 1]$. Therefore the maximum likelihood estimate exists and is obtained as the maximizer of $L(\,\cdot\,; x)$.

Let $k = \sum_{i=1}^n x_i$. We consider first the case of $1 < k < n$. To find the maximum point we compute the first derivative of $L(\,\cdot\,; x)$ on the interior $(0, 1)$ of the parameter space:

$$\frac{d}{d\theta} L(\theta; x) \;=\; \frac{d}{d\theta}\{\theta^k (1-\theta)^{n-k}\} \;=\; \underbrace{\theta^{k-1}(1-\theta)^{n-k-1}}_{>0} \Big\{k(1-\theta) \;-\; (n-k)\theta\Big\}.$$

We see that $\frac{d}{d\theta} L(\theta; x) = 0$ if and only if $\theta = k/n$. Since $L(\theta; x) = 0$ if $\theta \in \{0, 1\}$ we conclude that $\theta = k/n$ is the **global** maximizer of $L(\theta; x)$.

Now we consider the remaining cases. If $k = 0$, then

$$L(\theta; x) \;=\; \theta^k(1-\theta)^{n-k} \;=\; (1-\theta)^n$$

is (obviously) maximized by $\theta = 0$.
If $k = n$, then

$$L(\theta; x) \;=\; \theta^k(1-\theta)^{n-k} \;=\; \theta^n$$

is maximized by $\theta = 1$.

To summarize, if $X = x$, then the maximum likelihood **estimate** is

$$\widehat{\theta}_{ML}(x) \;=\; \frac{1}{n}\sum_{i=1}^{n} x_i,$$

and the corresponding **estimator**

$$\widehat{\theta}_{ML}(X) \;=\; \frac{1}{n}\sum_{i=1}^{n} X_i.$$

If we knew in advance that θ is strictly positive, then we could choose the parameter space by $\Theta = (0, 1]$. If $\theta < 1$, then it happens with positive probability that $\sum_{i=1}^n X_i = 0$. In this case, the maximum likelihood estimator is not defined.

9) Let $X_1, \ldots, X_n$ be i.i.d. with $X_i \sim \text{Uniform}([\theta_1, \theta_2])$, where $-\infty < \theta_1 < \theta_2 < \infty$.

(i) Compute the moment estimator of $\theta = (\theta_1, \theta_2)^T$.

(ii) Compute the maximum likelihood estimator of θ.

Solution

We have to estimate the two-dimensional parameter $\theta = \binom{\theta_1}{\theta_2}$ which means that we have to solve a system of two equations. The first two theoretical moments of X_1 are given by

$$
\begin{aligned}
E_\theta X_1 &= \frac{\theta_1 + \theta_2}{2}, \\
E_\theta[X_1^2] &= \text{var}_\theta(X_1) + (E_\theta X_1)^2 \\
&= \frac{(\theta_1 - \theta_2)^2}{12} + \frac{(\theta_1 + \theta_2)^2}{4}.
\end{aligned}
$$

Denote by $\widehat{\mu}_k = \frac{1}{n}\sum_{i=1}^n X_i^k$ $(k = 1, 2)$ the corresponding sample moments. Then the method of moments estimator $\widehat{\theta}_{MM} = (\widehat{\theta}_{1,MM}, \widehat{\theta}_{2,MM})^T$ is given by the solution to

$$
\begin{aligned}
\frac{\widehat{\theta}_{1,MM} + \widehat{\theta}_{2,MM}}{2} &= \frac{1}{n}\sum_{i=1}^n X_i, \\
\frac{(\widehat{\theta}_{1,MM} - \widehat{\theta}_{2,MM})^2}{12} &= \frac{1}{n}\sum_{i=1}^n X_i^2 - \Big(\frac{1}{n}\sum_{i=1}^n X_i\Big)^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2.
\end{aligned}
$$

(ii) Compute the maximum likelihood estimator of θ.

Solution

The likelihood function is given by

$$
L(\theta; X_1, \ldots, X_n) = \prod_{i=1}^n \frac{1}{\theta_2 - \theta_1} \mathbb{1}_{[\theta_1, \theta_2]}(X_i).
$$

We can easily find the maximum likelihood estimators by inspection:

$$
\begin{aligned}
\widehat{\theta}_{1,ML} &= \min\{X_1, \ldots, X_n\}, \\
\widehat{\theta}_{2,ML} &= \max\{X_1, \ldots, X_n\}.
\end{aligned}
$$