Solution to exercise 3

We want to compute $D(A_i)$ and $D^*(A_i)$, $i = 1, 2$, for the sets

$$A_1 = \left\{ \frac{1}{8}, \frac{1}{4}, \frac{1}{2}, \frac{3}{4} \right\}$$

and

$$A_2 = \left\{ \frac{1}{4}, \frac{1}{2}, \frac{5}{8}, \frac{3}{4} \right\}.$$

Looking at the definition of discrepancy and star discrepancy, for a one-dimensional set $\{x_1, \dots, x_n\}$ we have

$$D(x_1, \dots, x_n) = \sup_{a,b \in [0,1]} \left| \frac{|\{x_i \in [a,b]\}|}{n} - (b-a) \right| \tag{1}$$

and

$$D^*(x_1, \dots, x_n) = \sup_{b \in [0,1]} \left| \frac{|\{x_i \in [0,b]\}|}{n} - b \right|, \tag{2}$$

where $|\{x_i \in [a,b]\}|$ is the number of points in the interval $[a,b]$.

Let us start by computing the discrepancy $D\left(\left\{\frac{1}{8}, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}\right\}\right)$.

We want to consider sets with maximum number of points and minimum length (i.e., *clusters*) so that the value inside the absolute value in (1) is positive, as well as sets with minimal number of points and maximal length, so that the value inside the absolute value in (1) is negative. It's clear that possible good choices for this latter kind of sets are:

- $\left(\frac{1}{4}, \frac{1}{2}\right)$, that gives $\left|0 - \left(\frac{1}{4} - \epsilon\right)\right| = \frac{1}{4} - \epsilon$;
- $\left(\frac{1}{2}, \frac{3}{4}\right)$, that gives $\left|0 - \left(\frac{1}{4} - \epsilon\right)\right| = \frac{1}{4} - \epsilon$;
- $\left(\frac{1}{4}, \frac{3}{4}\right)$, that gives $\left|\frac{1}{4} - \left(\frac{1}{2} - \epsilon\right)\right| = \frac{1}{4} - \epsilon$,

for small $\epsilon > 0$. Here we introduce $\epsilon > 0$ because in the definition of the discrepancy the intervals are close. However, since ϵ can be chosen arbitrarily small, we know that $D(A_1) \geq \frac{1}{4}$.

Regarding possible choices of clusters, it's clear that they have to be sets including the point $\frac{1}{8}$. In particular, possible choices are:

- $\left[\frac{1}{8}, \frac{1}{4}\right]$, that gives $\left|\frac{2}{4} - \left(\frac{1}{4} - \frac{1}{8}\right)\right| = \frac{3}{8}$;
- $\left[\frac{1}{8}, \frac{1}{2}\right]$, that gives $\left|\frac{3}{4} - \left(\frac{1}{2} - \frac{1}{8}\right)\right| = \frac{3}{8}$;
- $\left[\frac{1}{8}, \frac{3}{4}\right]$, that gives $\left|\frac{4}{4} - \left(\frac{3}{4} - \frac{1}{8}\right)\right| = \frac{3}{8}$.

So we have $D(A_1) = \frac{3}{8}$.

Now we want to compute $D\left(\left\{\frac{1}{4}, \frac{1}{2}, \frac{5}{8}, \frac{3}{4}\right\}\right)$, proceeding as before. The only nice choices for sets with few points and large length are:

- $\left(\frac{1}{4}, \frac{1}{2}\right)$, that gives $\left|0 - \left(\frac{1}{4} - \epsilon\right)\right| = \frac{1}{4} - \epsilon$;
- $\left(\frac{1}{4}, \frac{5}{8}\right)$, that gives $\left|\frac{1}{4} - \left(\frac{1}{2} - \epsilon\right)\right| = \frac{1}{4} - \epsilon$;

So we know that $D(A_2) \geq \frac{1}{4}$. Regarding possible choices of clusters, it's clear that they to be sets including the point $\frac{5}{8}$. In particular, possible choices of this kind are:

- $\left[\frac{1}{2}, \frac{5}{8}\right]$, that gives $\left|\frac{2}{4} - \left(\frac{5}{8} - \frac{1}{2}\right)\right| = \frac{3}{8}$;
- $\left[\frac{1}{2}, \frac{3}{4}\right]$, that gives $\left|\frac{3}{4} - \left(\frac{3}{4} - \frac{1}{2}\right)\right| = \frac{1}{2}$.

So we have $D(A_2) = \frac{1}{2}$.
In order to compute the star discrepancy, we can proceed as in the lecture, moving from the left to the right. Consider first $A_1 = \left\{\frac{1}{8}, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}\right\}$. We have that $\frac{|\{x_i \in [0,b]\}|}{n} - b$ is:

- equal to $0 - \left(\frac{1}{8} - \epsilon\right) = -\frac{1}{8} + \epsilon$ for $b = \frac{1}{8} - \epsilon$, for small $\epsilon > 0$;
- equal to $\frac{1}{4} - \frac{1}{8} = \frac{1}{8}$ for $b = \frac{1}{8}$;
- equal to $\frac{1}{4} - \left(\frac{1}{4} - \epsilon\right) = \epsilon$ for $b = \frac{1}{4} - \epsilon$, for small $\epsilon > 0$;
- equal to $\frac{2}{4} - \frac{1}{4} = \frac{1}{4}$ for $b = \frac{1}{4}$;
- equal to $\frac{2}{4} - \left(\frac{1}{2} - \epsilon\right) = \epsilon$ for $b = \frac{1}{2} - \epsilon$, for small $\epsilon > 0$;
- equal to $\frac{3}{4} - \frac{1}{2} = \frac{1}{4}$ for $b = \frac{1}{2}$;
- equal to $\frac{3}{4} - \left(\frac{3}{4} - \epsilon\right) = \epsilon$ for $b = \frac{3}{4} - \epsilon$, for small $\epsilon > 0$;
- equal to $\frac{4}{4} - \frac{3}{4} = \frac{1}{4}$ for $b = \frac{3}{4}$.

The maximum absolute value of the values listed above is $D^*(A_1) = \frac{1}{4}$.
Considering now $A_2 = \left\{\frac{1}{4}, \frac{1}{2}, \frac{5}{8}, \frac{3}{4}\right\}$, we have that $\frac{|\{x_i \in [0,b]\}|}{n} - b$ is:

- equal to $0 - \left(\frac{1}{8} - \epsilon\right) = -\frac{1}{8} + \epsilon$ for $b = \frac{1}{8} - \epsilon$, for small $\epsilon > 0$;
- equal to $\frac{1}{4} - \frac{1}{4} = 0$ for $b = \frac{1}{8}$;
- equal to $\frac{1}{4} - \left(\frac{1}{2} - \epsilon\right) = -\frac{1}{4} + \epsilon$ for $b = \frac{1}{2} - \epsilon$, for small $\epsilon > 0$;
- equal to $\frac{2}{4} - \frac{1}{2} = 0$ for $b = \frac{1}{4}$;
- equal to $\frac{2}{4} - \left(\frac{5}{8} - \epsilon\right) = -\frac{1}{8} + \epsilon$ for $b = \frac{1}{2} - \epsilon$, for small $\epsilon > 0$;
- equal to $\frac{3}{4} - \frac{5}{8} = \frac{1}{8}$ for $b = \frac{5}{8}$;
- equal to $\frac{3}{4} - \left(\frac{3}{4} - \epsilon\right) = \epsilon$ for $b = \frac{3}{4} - \epsilon$, for small $\epsilon > 0$;
- equal to $\frac{4}{4} - \frac{3}{4} = \frac{1}{4}$ for $b = \frac{3}{4}$.

So we find $D^*(A_2) = \frac{1}{4}$.
Summing up, we have found $D(A_1) = \frac{3}{8}$, $D^*(A_1) = \frac{1}{4}$, $D(A_2) = \frac{1}{2}$, $D^*(A_2) = \frac{1}{4}$. Note that the relations

$$D^*(A_i) \leq D(A_i) \leq 2 \cdot D^*(A_i), \quad i = 1, 2,$$

hold as expected.