Problem 2.

Proof. This is because each boolean function $f : \{0,1\}^n \to \{0,1\}$ can be written as

$$\bigvee_{a:f(a)=1} \left(\bigwedge_{i=1}^{n} (x_i = a_i) \right)$$

where $\{\alpha : f(\alpha) = 1\}$ is at most the size of $\{0,1\}^n$, i.e., 2^n. □

Problem 3.

Proof. Let R be the rectangle we want to learn and R_0 be the smallest axis-aligned rectangle containing all sample points, as described in the hint. Clearly $R_0 \subset R$ so the error simply corresponds to the region $R \backslash R_0$. For notational convenience we denote R as $[a,b] \times [c,d]$ and R_0 as $[a_0,b_0] \times [c_0,d_0]$. It follows that the four strips $[a,b] \times [d_0,d]$, $[b_0,b] \times [c,d]$, $[a,b] \times [c,c_0]$, and $[d,d_0] \times [c,d]$ cover $R \backslash R_0$, so it suffices to reduce the probability of each strip occurring to under $\epsilon/4$.

Let n be the sample size. The probability of all m samples not landing in the top strip is $(1-\epsilon/4)^m$. Therefore union bound implies the probability of all m samples lying in R_0 is bounded above by $4(1-\epsilon/4)^m$. Setting the upper bound of this to be δ, we see that the requirement on number of samples is $\log_{1-\epsilon/4}(\delta/4) \sim \log(\delta/4)$ which can be bounded by a polynomial for $\delta \leqslant 1$.

To generalize, in $\mathbb{R}^k$, we consider a similar approach, but we replace the four strips by 2^k k-dimensional boxes whose union cover the difference between the objective box and the "smallest" box. □

Problem 4.

Solution. (1) Consider $\mathcal{F}$, the set of functions from $\mathbb{R}$ to $\{0,1\}$ of form $1_{\{x>t\}}$. Any singleton is clearly shattered by $\mathcal{F}$, but conversely any set with $\geqslant 2$ (distinct) elements cannot be shattered by $\mathcal{F}$.

(2) For infinite VC-dimension consider $\mathcal{F}$, the collection of all functions from $\mathbb{R}$ to $\{0,1\}$.

(3) On one hand $\mathcal{F}$ contains 2^n functions so the VC dimension is at most n. On the pother hand, the vectors $e_1 := (1,0,0,\ldots)$, $e_2 := (0,1,0,\ldots)$, $e_3 := (0,0,1,\ldots)$ and so on form a set of n elements, which we call E, and E is shattered by $\mathcal{F}$. In particular, if $g : E \to \{-1,1\}$, and if we define I by

$$I := \{i \in \{1,\ldots,n\} : g(e_i) = 1\}$$

then h_I agrees with g on E. Therefore the VC dimension is precisely n.

Problem 6.

Proof. (1)

(2) This one is immediate, since the edge length is less than 1.

(3) First consider the balls $B_{1,n} := \{x \in \mathbb{R}^n : \|x\|_1 = \sum |x_i| \leqslant 1\}$. The ball should be divided into 2^n congruent parts based on the sign of each component. Looking at the principal 2^n-ant, the volumn of $\{x \in \mathbb{R}^n : \sum x_i \leqslant 1$ and $x_i \geqslant 0$ for all $i\}$ is $1/n!$, so the volumn of $B_{1,n}$ is $2^n/n!$.

Note that if $\sum |x_i|^2 \leqslant 1$ then by Cauchy-Schwarz $\sum |x_i| \leqslant \sqrt{n}$, so $B_n \subset \sqrt{n} B_{1,n}$. That is, the volumn of B_n should be bounded by $n^{n/2} 2^n / n!$ which converges to 0.

(4) Clearly D_n is a subset of the ball with center zero and radius $1/2$. From the previous part, even these larger balls have volumn converging to 0.

□

Problem 7.

Proof. □

$$\int_{-\infty}^{\infty} e^{-x^2/2}\,\mathrm{d}x = \sqrt{\pi}.$$

Theorem 0.0.1: Hoeffding Inequality

Let $X_1, X_2, \ldots$ be i.i.d. with $\mathbb{P}(X_1 = 1) = P(X_1 = -1) = 1/2$. Let $a_1, a_2, \ldots \in \mathbb{R}$. Then for $n \geqslant 1$ and $t \geqslant 0$,

$$\mathbb{P}\left(\sum_{i=1}^{n} a_i X_i \geqslant t\right) \leqslant \exp\left(-\frac{t^2}{2\sum_{i=1}^{n} a_i^2}\right) \qquad \text{and therefore} \qquad \mathbb{P}\left(\left|\sum_{i=1}^{n} a_i X_i\right| \geqslant t\right) \leqslant 2\exp\left(-\frac{t^2}{2\sum_{i=1}^{n} a_i^2}\right).$$

Proof. This proof is trivial. □