Задание 1

Задача 1 (обязательная для всех)

Пользуясь результатом, полученным на лекции для оценки максимального числа разбиений M точек в N мерном пространстве с помощью гиперплоскостей, проходящих через начало координат, получить оценку для разбиения таких точек с помощью

- произвольных гиперплоскостей
- при K ограничениях (K<N) в виде линейных неравенств на направляющий вектор гиперплоскости.
- сфер, имеющих общую точку в начале координат.

Задача 2

Пользуясь материалами курса доказать теорему Гливенко-Кантелли:

$$\sup_x |F_n(x) - F(x)| \underset{n\to\infty}{\to} 0 \text{ с вероятностью 1.}$$

Здесь $F_n(x)$ - эмпирическая функция распределения, построенная по независимой выборке n значений случайной величины, имеющей функцию распределения $F(x)$.

Задача 3.

Задачу приближения функции регрессии $M(y \mid x)$ в классе функций $F = \{f(x,a)\}$, где a - вектор параметров, можно понимать как задачу минимизации на F функционала среднего риска $J_m(a) = \int (y - f(x,a))^2 dP(x,y)$. Функционал эмпирического риска при этом имеет вид $J_e(a) = \frac{1}{l}\sum_{i=1}^{l}(y_i - f(x_i,a))^2$, где $(x_1, y_1, ..., x_l, y_l)$ - независимая выборка значений пары (x,y), имеющей распределение $P(x,y)$. При построении равномерной по классу F оценки величины уклонения эмпирического риска от среднего риска рассматривают множества вида $A_{a,c} = \{x, y : (y - f(x,a))^2 < c\}$.

Оценить функцию роста системы $A_{a,c}$ (параметры c и a пробегают все возможные значения) в задаче построения регрессии методом минимизации эмпирического риска в классе

- полиномов степени не выше n (x и y - скаляры)
- линейных функций вида $f(x,a) = \sum_{i=1}^{n} a_i z_i$, где z_i - i-ая координата вектора x.