ЛЕКЦИЯ 21

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ ВАПНИКА-ЧЕРВОНЕНКИСА

Эта лекция будет посвящена доказательству теоремы Вапника – Червоненкиса (теорема 66).

Напомним формулировку теоремы. Пусть:

$$VC(S, R) = d,$$

а $A \subset S$ — это любое конечное подмножество.

Пусть также:

$$\epsilon > 0.$$

Тогда существует ϵ-сеть размера:

$$m \leqslant \left\lceil \frac{8d}{\epsilon} \log_2 \frac{8d}{\epsilon} \right\rceil.$$

Док-во: Обозначим за n — мощность множества A:

$$n = |A|.$$

Будем строить ϵ-сеть по схеме случайного выбора с возвращением. Это значит, что друг за другом будем извлекать m случайных независимых возможно совпадающих элементов из множества A.

В итоге получается **мультимножество** N. Вероятность:

$$P(N) = \frac{1}{n^m}.$$

Нужно доказать, что с положительной вероятностью N является ϵ-сетью:

$$P\Big(N \text{ является } \epsilon\text{-сетью}\Big) > 0,$$

то есть можно доказать наоборот, рассмотрев вероятность:

$$P\Big(N \text{ не является } \epsilon\text{-сетью}\Big) = P\Big(\exists r \in R: \ \big|r \cap A\big| \geqslant \epsilon n, \ r \cap N = \emptyset\Big).$$

Обозначим вредоносное событие:

$$E_1 = \exists r \in R: \ \big|r \cap A\big| \geqslant \epsilon n, \ r \cap N = \emptyset.$$

Тогда цель — доказать, что:

$$P(E_1) < 1.$$

Непосредственно оценить такую вероятность довольно сложно, поэтому Вапник и Червоненкис решили усложнить событие E_1, тем самым упростив доказательство в целом.

Рассмотрим еще одно случайное мультимножество T, которое строится точно так же, как и N, но независимо от него. Теоретически эти два множества могут даже совпасть. В новом мультимножестве тоже содержится m элементов.

Запишем иначе в терминах нового мультимножества:

$$E_1 = \Big\{N: \ \exists r \in R: \ \big|r \cap A\big| \geqslant \epsilon n, \ r \cap N = \emptyset\Big\} = \\ = \Big\{(N, T): \ \exists r \in R: \ \big|r \cap A\big| \geqslant \epsilon n, \ r \cap N = \emptyset\Big\}.$$

Теперь запишем событие E_2:

$$E_2 = \left\{(N, T): \ \exists r \in R: \ \big|r \cap A\big| \geqslant \epsilon n, \ r \cap N = \emptyset, \ \left|r \cap T \geqslant \frac{\epsilon m}{2}\right|\right\}.$$

Заметим, что:

$$E_2 \Rightarrow E_1.$$

Тогда:

$$P(E_2) \leqslant P(E_1).$$

Для дальнейшего доказательства потребуется обосновать две леммы, которые будут рассмотрены далее.

Лемма 12

$$P(E_2) \geqslant \frac{5}{6} P(E_1).$$

Док-во: Распишем следующую вероятность:

$$P(E_2 | E_1) = \frac{P(E_2 \cap E_1)}{P(E_1)} = \frac{P(E_2)}{P(E_1)}.$$

Тогда достаточно доказать, что условная вероятность:

$$P(E_2 | E_1) \geqslant \frac{5}{6}.$$

На рисунке 21.1 показано событие E_2 при условии E_1. Фактически:

$$P(E_2|E_1) = P\left(\left|r \cap T\right| \geqslant \frac{\epsilon m}{2}\right),$$

то есть фиксируется какой-то r, что:

$$r \in R: \ \left|r \cap A\right| \geqslant \epsilon n.$$

Мультимножество T выбиралось так:

$$T = (x_1, \dots, x_m),$$

то есть по схеме испытаний Бернулли.

Заметим теперь, что:

$$\left|r \cap T\right| = Binom(m, \geqslant \epsilon).$$

Тогда:

$$P\left(\left|r \cap T\right| \geqslant \frac{\epsilon m}{2}\right) \geqslant P\left(Binom(m, \geqslant \epsilon) \geqslant \frac{\epsilon m}{2}\right).$$

Обозначим:

$$\eta = Binom(m, \geqslant \epsilon).$$

Перепишем в новых обозначениях:

$$P\left(Binom(m, \geqslant \epsilon) \geqslant \frac{\epsilon m}{2}\right) = P\left(-\eta \leqslant -\frac{\epsilon m}{2}\right) =$$
$$= P\left(M\eta - \eta \leqslant M\eta - \frac{\epsilon m}{2}\right) = P\left(M\eta - \eta \leqslant \frac{\epsilon m}{2}\right) \geqslant$$
$$\geqslant P\left(\left|M\eta - \eta\right| \leqslant \frac{\epsilon m}{2}\right) \geqslant 1 - \frac{D\eta}{\left(\frac{\epsilon m}{2}\right)^2}.$$

В данных рассуждениях последняя оценка получена по неравенству Чебышева, а D является дисперсией. Тогда запишем далее:

$$1 - \frac{D\eta}{\left(\frac{\epsilon m}{2}\right)^2} = 1 - \frac{m\epsilon(1-\epsilon)}{\left(\frac{\epsilon m}{2}\right)^2} \geqslant 1 - \frac{\epsilon m}{\left(\frac{\epsilon m}{2}\right)^2} = 1 - \frac{4}{\epsilon m}.$$

Так как:

$$\epsilon m \geqslant \epsilon \cdot \frac{8d}{\epsilon} \log_2 \frac{8d}{\epsilon} = 8d \log_2 \frac{8d}{\epsilon} > 8d \log_2 8d \geqslant 24,$$

потому что:

$$d \geqslant 1,$$

то получается в итоге:

$$P(E_2|E_1) \geqslant 1 - \frac{4}{\epsilon m} \geqslant \frac{5}{6}.$$

Лемма доказана. ■

Во второй лемме будет использована вероятностная схема, которую предлагается доказать в качестве упражнения:

Задача Следующая вероятностная схема эквивалентна схеме последовательного выбора мультимножеств (N, T) — сперва выбираем случайное мультимножество U:

$$U \subset A: \ |U| = 2m,$$

затем случайно разбиваем U на две части одинаковой мощности:

$$U = N \sqcup T.$$

При этом:

$$P(N) = \frac{1}{C_{2m}^m}, \quad U = (z_1, \dots, z_{2m}).$$

Лемма 13

$$P(E_2) \leqslant 2^{-\frac{\epsilon m}{2}} g(2m, d),$$

причем:

$$g(2m, d) = \sum_{i=0}^{d} C_{2m}^i.$$

Док-во: Заметим, что:

$$P(E_2) = \sum_U P(E_2|U) \cdot P(U),$$

что выполняется по формуле полной вероятности.

Следовательно, достаточно доказать, что:

$$P(E_2|U) \leqslant 2^{-\frac{\epsilon m}{2}} g(2m, d) \quad \forall U.$$

Запишем тогда:

$$E_2 = \bigcup_{r \in R: \ |r \cap A| \geqslant \epsilon n} E_{2,r},$$

где:

$$E_{2,r} = \left\{ (N, T): \ r \cap N = \emptyset, \ \big|r \cap T\big| \geqslant \frac{\epsilon m}{2} \right\}.$$

Пусть фиксируется какое-то U. Тогда некоторые из событий E_2 — тождественно совпадают (см. рис. 21.2).

Необходимо отметить, что на рис. 21.2 показано именно мультимножество U, а не подмножество. Пусть:

$$r_1 \cap U = r_2 \cap U$$

для некоторых:

$$r_1 \in R, \quad r_2 \in R.$$

Тогда:

$$E_{2,r_1} = E_{2,r_2}.$$

Вывод:

$$P(E_2|U) \leqslant \sum_{r \in P_{r_U}(R)} P(E_{2,r}|U).$$

Далее будет доказано, что:

$$P(E_2|U) \leqslant 2^{-\frac{\epsilon m}{2}}.$$

Тогда:

$$\sum_{r \in P_{r_U}(R)} P(E_{2,r}|U) \leqslant 2^{-\frac{\epsilon m}{2}} \cdot \sum_{r \in P_{r_U}(R)} 1.$$

Так как U состоит из $2m$ элементов, то по следствию 7 из леммы 10:

$$2^{-\frac{\epsilon m}{2}} \cdot \sum_{r \in P_{r_U}(R)} 1 \leqslant 2^{-\frac{\epsilon m}{2}} g(2m, d).$$

Если:

$$\left|r \cap U\right| < \frac{\epsilon m}{2},$$

то:

$$P(E_{2,r}|U) = 0.$$

Пусть теперь:

$$\left|r \cap U\right| \geqslant \frac{\epsilon m}{2}.$$

Обозначим за x:

$$x = \left|r \cap U\right| \geqslant \frac{\epsilon m}{2}.$$

Тогда (см. рис. 21.3):

$$P(E_{2,r}|U) \leqslant P(N \cap X = \emptyset) = \frac{C^m_{2m-x}}{C^m_{2m}} = \frac{(2m-x)!\, m!\, m!}{m!\,(m-x)!\,(2m!\,)} =$$
$$= \frac{m(m-1) \cdot \ldots \cdot (m-x+1)}{2m(2m-1 \cdot \ldots \cdot (2m-x+1))} < \frac{1}{2},$$

так как:

$$\frac{m}{2m} < \frac{1}{2}, \quad \frac{m-1}{2m-1} < \frac{1}{2}, \quad \frac{m-x+1}{2m-x+1} < \frac{1}{2}.$$

Следовательно, лемма доказана. ■

Для завершения доказательства теоремы 66 остается вспомнить, что:

$$g(2m, d) \leqslant (2m)^d,$$

откуда следует, что:

$$P(E_2) \leqslant 2^{-\frac{\epsilon m}{2}} (2m)^d.$$

Заметим, что функция $2^{-\frac{\epsilon m}{2}}$ убывает экспоненциально по m, а функция $(2m)^d$ возрастает полиномиально по m, то есть с некоторого момента точно выполняется доказываемое утверждение.

Однако доказательство того, что именно следующий момент:

$$m = \left\lceil \frac{8d}{\epsilon} \log_2 \frac{8d}{\epsilon} \right\rceil$$

является тем самым моментом, приводиться не будет, считая, что эта точная выкладка — не самая важная в данном доказательстве (хотя это вычисление и несложное).

Тогда теорему можно считать доказанной. ■

Завершение доказательства можно найти в книге лектора — *Райгородский Андрей Михайлович*, «*Системы общих представителей в комбинаторике и их приложения в геометрии*».

На следующей лекции будет рассмотрено приложение данной темы к статистике.

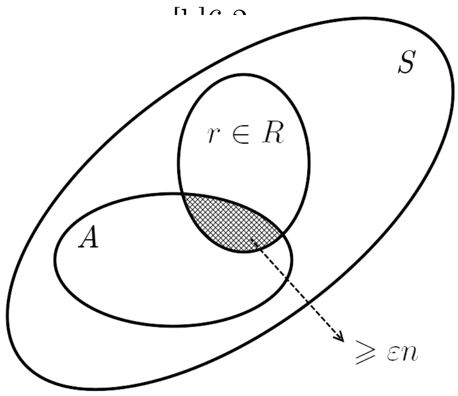

Рис. 21.1

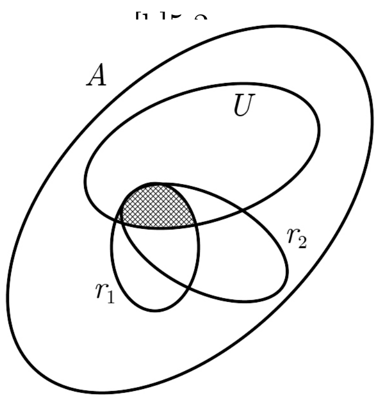

Рис. 21.2

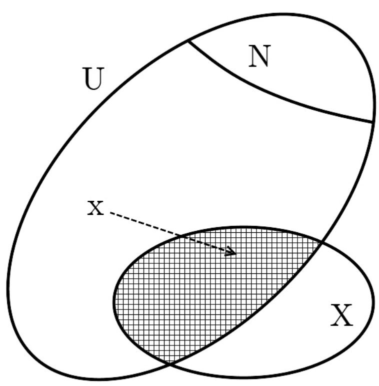

Рис. 21.3