Задание:

Рассмотрите вариант алгоритма большинства для случая, когда существует эксперт, про которого известно, что он делает не более k ошибок. Получите оценку числа ошибок такого алгоритма большинства.

Решение:

Допустим, имеется N экспертов, и известно, что i-й эксперт делает не более k_i ошибок. Алгоритм большинства в этом случае будет выглядеть следующим образом:

$$a = \begin{cases} 1, \Sigma_{i=1}^{n} x_i > \frac{n}{2} \\ 0, \Sigma_{i=1}^{n} x_i < \frac{n}{2} \end{cases}$$

Алгоритм большинства для случая, когда существует эксперт, про которого известно, что он не делает более k ошибок, можно описать следующим образом:

- Подсчитываем голоса каждого эксперта за каждый возможный вариант решения.
- Если голоса большинства экспертов совпадают, выбираем их решение.
- Если голоса не совпадают, выбираем решение эксперта, который не делает более k ошибок.

Данный алгоритм позволяет учитывать экспертное мнение и при этом уменьшать вероятность ошибки, так как мы выбираем решение эксперта, который совершает меньше ошибок.

Предположим, что имеется n экспертов, из которых только один эксперт, обозначим его как E, делает не более k ошибок. Пусть вероятность того, что E допустит ошибку равна p, а вероятность правильного ответа равна (1-p).

Тогда для каждого эксперта, отличного от E, вероятность того, что он допустит не более k ошибок, может быть вычислена с помощью биномиального распределения:

$$P_i = \sum_{j=0}^{k} C_{n-1}^{j} p^j (1-p)^{n-1-j}$$

где P_i - вероятность того, что i-й эксперт допустит не более k ошибок, C_{n-1}^{j} - число сочетаний из n-1 по j (так как E не учитывается в подсчете), p - вероятность ошибки эксперта, (1-p) - вероятность правильного ответа эксперта.

Таким образом, вероятность P того, что E допустит не более k ошибок, равна 1 минус вероятность того, что хотя бы один из оставшихся n-1 экспертов допустит более k ошибок.

$$P_i = 1 - \sum_{i=0}^{n-1} \sum_{j=k+1}^{n} C_{n-2}^{i} C_{n-j}^{1} p^j p^{i+1} (1-p)^{n-1-j}$$

где первая сумма вычисляет вероятность того, что i экспертов из n-1 допустят не более k ошибок (включая E), а вторая сумма вычисляет вероятность того, что оставшийся эксперт, отличный от E и j-й по счету, допустит более k ошибок.

Теперь мы можем принимать решения, выбирая ответ, который дал эксперт E, если он допустил не более k ошибок, и ответ, который поддерживается большинством оставшихся экспертов, если E допустил более k ошибок.

Для оценки числа ошибок алгоритма большинства с учетом знания о максимальном числе ошибок k, которое может сделать эксперт, можно использовать неравенство Чернова.

Вероятность того, что количество ошибок, допущенных алгоритмом большинства, превысит определенный порог x, можно оценить следующим образом:

$$P(X > x) \leq \exp\left(-2n\epsilon^2\right)$$

где $\epsilon = \frac{(x-np)}{(np(1-p))^{1/2}}$ - параметр, который учитывает отклонение среднего значения от идеального значения np, а также вероятность ошибки p.

Таким образом, для алгоритма большинства с учетом максимального числа ошибок k можно использовать данную формулу, заменив значение p на вероятность того, что эксперт допустит не более k ошибок.

Для конкретного эксперта i вероятность того, что он допустит не более k ошибок, может быть вычислена с помощью биномиального распределения:

$$P_i = \sum_{j=0}^{k} C_{n-1}^{j} p^j (1-p)^{n-1-j}$$

где P_i - вероятность того, что i-й эксперт допустит не более k ошибок, C_{n-1}^{j} - число сочетаний из n-1 по j (так как i-й эксперт не учитывается в подсчете), p - вероятность ошибки эксперта, (1-p) - вероятность правильного ответа эксперта

Далее, значение ϵ_i для конкретного эксперта можно вычислить по формуле:

$$\epsilon_i = \frac{x - np_i}{(np_i(1-p_i))^{1/2}}$$

где x - пороговое значение числа ошибок, n - общее число экспертов, p_i - вероятность того, что i-й эксперт допустит не более k ошибок.

Таким образом, для оценки числа ошибок алгоритма большинства с учетом максимального числа ошибок k необходимо вычислить значение ϵ_i для каждого эксперта и затем подставить полученные значения в формулу для неравенства Чернова.