

51 А была ли ошибка?

Задача. Боб написал программу для шифровальной машины **Amgine**. Алиса оценила объем программы и стиль программирования Боба. Алиса утверждает, что программа содержит ошибку с вероятностью α . В ответ Боб разработал систему из n тестов. Тест номер i обнаруживает ошибку с вероятностью p_i независимо от других тестов. Все тесты прошли успешно. Какова вероятность того, что в программе все-таки есть ошибка? Изменяются ли выводы, если учитывать результат очередного теста сразу после его выполнения?

Решение. Пусть событие A состоит в том, что программа содержит ошибку, событие B_i — ошибка будет обнаружена i -м тестом, $i = 1, 2, \dots, n$. Тогда $\mathbf{P}\{A\} = \alpha$ (априорная вероятность ошибки), $\mathbf{P}\{B_i | A\} = p_i$ и требуется оценить апостериорную (после тестирования) вероятность $\mathbf{P}\{A | \bar{B}_1, \bar{B}_2, \dots, \bar{B}_n\}$.

По формуле Байеса

$$\begin{aligned} \mathbf{P}\{A | \bar{B}_1, \bar{B}_2, \dots, \bar{B}_n\} &= \frac{\mathbf{P}\{\bar{B}_1, \bar{B}_2, \dots, \bar{B}_n | A\} \mathbf{P}\{A\}}{\mathbf{P}\{\bar{B}_1, \bar{B}_2, \dots, \bar{B}_n\}} = \\ &= \frac{\mathbf{P}\{\bar{B}_1, \bar{B}_2, \dots, \bar{B}_n | A\} \mathbf{P}\{A\}}{\mathbf{P}\{\bar{B}_1, \bar{B}_2, \dots, \bar{B}_n | A\} \mathbf{P}\{A\} + \mathbf{P}\{\bar{B}_1, \bar{B}_2, \dots, \bar{B}_n | \bar{A}\} \mathbf{P}\{\bar{A}\}} = \\ &= \frac{\alpha \prod_{i=1}^n (1 - p_i)}{\alpha \prod_{i=1}^n (1 - p_i) + 1 - \alpha}. \end{aligned}$$

Пусть результаты теста учитываются сразу после его выполнения, α_k — апостериорная вероятность ошибки после выполнения k тестов и одновременно априорная перед $(k+1)$ -м, $\alpha_0 = \alpha$. Снова по формуле Байеса:

$$\alpha_k = \frac{\alpha_{k-1}(1 - p_k)}{\alpha_{k-1}(1 - p_k) + (1 - \alpha_k)}$$

или

$$\frac{1}{\alpha_k} - 1 = \left(\frac{1}{\alpha_{k-1}} - 1 \right) \cdot \frac{1}{1 - p_k}$$

или

$$\frac{1}{\alpha_k} - 1 = \left(\frac{1}{\alpha_0} - 1 \right) \cdot \frac{1}{\prod_{i=1}^k (1 - p_i)}.$$

Отсюда $\alpha_n = \mathbf{P}\{A | \bar{B}_1, \bar{B}_2, \dots, \bar{B}_n\}$, т. е. адаптивный учет результатов тестов не меняет окончательную апостериорную вероятность. \square