

Информационный критерий Акаике (Akaike's information criterion)

Синонимы: AIC

Разделы: [Метрики](#)

Критерий для выбора лучшей из нескольких статистических моделей, построенных на одном и том же наборе данных и использующих логарифмическую функцию правдоподобия. Предложен Хироцугу Акаикэ в 1974 году. Критерий является не статистическим, а информационным, поскольку основан на оценке потери информации при уменьшении числа параметров модели. Критерий позволяет найти компромисс между сложностью модели (числом параметров) и ее точностью. В общем случае AIC вычисляется по формуле:

$$AIC = 2k - 2\ln(L),$$

где k — число параметров модели, L — максимизированное значение функции правдоподобия модели. Лучшей признается та модель, для которой значение AIC минимально.

AIC тесно связан с байесовским информационным критерием BIC, но, в отличие от него, содержит функцию штрафа, линейно зависящую от числа параметров.

Если модель использует метод наименьших квадратов, то критерий может быть вычислен следующим образом:

$$AIC = \ln\left(\frac{RSS_p}{n}\right) + 2\frac{p}{n} + 1 + \ln 2\pi,$$

где RSS_p — сумма квадратов остатков модели, полученная при оценке коэффициентов модели методом наименьших квадратов, n — объем обучающей выборки.

Из выражения видно, что при фиксированном размере выборки рост критерия обусловлен в основном увеличением числа параметров модели, а не ее ошибкой. Т.е. за увеличение числа параметров модель «штрафуется» сильнее, чем за долю необъясненной дисперсии ошибки. Таким образом, задача заключается в том, чтобы выбрать модель с минимальным числом параметров, которые объясняют наибольшую долю дисперсии ошибки.

На практике это делается следующим образом. Берется «нулевая модель», которая содержит только свободный член, и для нее вычисляется значение критерия. Затем в нулевую модель поочередно добавляются параметры, и каждый раз AIC вычисляется вновь. Выбирается модель, для которой значение критерия окажется минимальным.

На малых выборках рекомендуется применять скорректированный критерий Акаике.

В Logitom существует специализированный визуализатор отчет по регрессии, который отображает статистические параметры и результаты статистических тестов. С его помощью производится оценка полученных метрик, в том числе и AIC — информационный критерий Акаике.