

Тест отношения правдоподобия (Likelihood-ratio test)

Синонимы: Тест Уилкса, Wilks's test, LR-test

Раздел: Метрики

Тест отношения правдоподобия — это статистическая процедура для оценивания степени соответствия данным двух статистических моделей, одна из которых строится на множестве всех доступных переменных выборки, а другая только на некотором их подмножестве.

Иными словами, цель применения теста заключается в том, чтобы подтвердить или опровергнуть гипотезу, что модель с меньшим числом переменных (короткая модель) не будет значимо отличаться по качеству от модели с большим числом переменных (длинной модели). При этом модели полагаются **вложенными**, т.е. короткую модель можно получить из длинной путем наложения ограничений на ее параметры.

Цель применения теста в предсказательном моделировании — снижение размерности пространства входных признаков. Действительно, чем больше переменных мы используем для построения модели, тем больше информации привлекается, и тем лучше модель, как ожидается, будет соответствовать данным. Однако на практике не все входные переменные являются одинаково информативными и значимо влияют на выходную переменную. Включение таких переменных в модель только усложняет ее, не улучшая существенно качество.

Таким образом, если удастся обнаружить малозначимые переменные и доказать, что модель без них будет работать не хуже (или незначительно хуже), чем с ними, то эти переменные можно исключить без значимого ущерба для точности модели, заменив длинную модель короткой. Для этого и служит тест отношения правдоподобия.

Статистика теста отношения правдоподобия вычисляется по формуле:

$$\lambda_{LR} = -2 \ln \left(\frac{L(\Theta_0|x)}{L(\Theta|x)} \right),$$

где $L(\Theta|x)$ — это функции правдоподобия модели без ограничений на параметры (длинной модели), и $L(\Theta_0|x)$ — функция правдоподобия для модели с ограничениями на параметры (короткой модели).

Поскольку все вероятности положительны и ограниченный максимум не может превышать неограниченный максимум, отношение правдоподобия изменяется между нулем и единицей.

Если короткая и длинная модель работают одинаково хорошо, то значения их функции правдоподобия близки, и λ_{LR} стремится к 0. Если короткая модель работает хуже, чем длинная, т.е. ее функция правдоподобия уменьшается, то λ_{LR} становится больше 0. Статистическую значимость данного увеличения и должен подтвердить или опровергнуть тест.

Для параметров $\Theta_0 = \Theta \cap \Theta_0$ выдвигается простая гипотеза $H_0 : (\varphi(\Theta \setminus \Theta_0) = 0)$. Если H_0 верно, то статистика теста отношения правдоподобия имеет (возможно, асимптотическое) распределение по закону $\chi^2(q)$, по которому можно оценить статистическую значимость модели. $\Theta \setminus \Theta_0$ — это разность множеств, $\Theta \cap \Theta_0$ — пересечение множеств параметров Θ и Θ_0 , а q — число ограничений на параметры (т.е. количество параметров, которое нужно исключить из длинной модели, чтобы получить короткую). Оно же равно числу степеней свободы распределения.

На практике чаще всего используется ограничение $\Theta_0 = \{\theta_{i_1} = 0, \theta_{i_2} = 0, \dots, \theta_{i_{n-q}} = 0\}$ для параметров $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$.

Если значение статистики больше критического, которое определяется по таблицам критических значений распределения χ^2 для числа степеней свободы q при заданном уровне значимости, то ограничения отвергаются, и предпочтение отдается длинной модели, а в противном случае — короткой.