

Ошибки I и II рода (Type I errors, type II errors)

Раздел: [Метрики](#)

Ошибки I и II рода — это ключевые понятия, используемые в [анализе данных](#) и [математической статистике](#).

- В анализе данных ошибки I и II рода используются для оценки точности моделей [бинарной классификации](#).
- В математической статистике ошибки I и II рода применяются при проверке статистических [гипотез](#).

Рассмотрим пример. Пусть банк использует систему классификации заемщиков на кредитоспособных и некредитоспособных. При этом первым кредит выдается, а вторые получают отказ. Таким образом, обнаружение некредитоспособного заемщика можно рассматривать как «сигнал тревоги», сообщающий о возможных [рисках](#). Подробнее об оценке вероятности дефолта потенциального заемщика с помощью Loginom в статье [«Machine learning в Loginom на примере задачи с Kaggle»](#).

Любой реальный классификатор совершает ошибки. В нашем случае таких ошибок может быть две:

- Кредитоспособный заемщик распознается моделью как некредитоспособный и ему отказывается в кредите. Данный случай можно трактовать как «ложную тревогу».
- Некредитоспособный заемщик распознается как кредитоспособный и ему ошибочно выдается кредит. Данный случай можно рассматривать как «пропуск цели».

Несложно увидеть, что эти ошибки неравноценны по связанным с ними проблемам. В случае «ложной тревоги» потери банка составят только проценты по невыданному кредиту. В случае «пропуска цели» можно потерять всю сумму выданного кредита. Поэтому системе важнее не допустить «пропуск цели», чем «ложную тревогу».

Поскольку с точки зрения логики задачи нам важнее правильно распознать некредитоспособного заемщика (чем ошибиться в распознавании кредитоспособного) будем называть соответствующий исход классификации **положительным** (заемщик некредитоспособен), а противоположный — **отрицательным** (заемщик кредитоспособен). Тогда возможны следующие исходы классификации:

- Некредитоспособный заемщик классифицирован как некредитоспособный, т.е. положительный класс распознан как положительный. Наблюдения, для которых это

имеет место называются **истинно-положительными** (true positive — TP).

- Кредитоспособный заемщик классифицирован как кредитоспособный, т.е. отрицательный класс распознан как отрицательный. Наблюдения, которых это имеет место, называются **истинно отрицательными** (true negative — TN).
- Кредитоспособный заемщик классифицирован как некредитоспособный, т.е. имела место ошибка, в результате которой отрицательный класс был распознан как положительный. Наблюдения, для которых был получен такой исход классификации, называются **ложно-положительными** (false positive — FP), а ошибка классификации называется **ошибкой I рода**.
- Некредитоспособный заемщик распознан как кредитоспособный, т.е. имела место ошибка, в результате которой положительный класс был распознан как отрицательный. Наблюдения, для которых был получен такой исход классификации, называются **ложно-отрицательными** (false negative — FN), а ошибка классификации называется **ошибкой II рода**.

Таким образом, ошибка I рода, или ложно-положительный исход классификации, имеет место, когда отрицательное наблюдение распознано моделью как положительное. Ошибкой II рода, или ложно-отрицательным исходом классификации, называют случай, когда положительное наблюдение распознано как отрицательное. Поясним это с помощью таблицы.

Классифицировано\Факт	Некредитоспособный (положительный)	Кредитоспособный (отрицательный)
Некредитоспособный (положительный)	Ошибки нет (истинно-положительный, True Positive)	Ошибка I рода (ложно-положительный, False Positive), ложная тревога
Кредитоспособный (отрицательный)	Ошибка II рода (ложно-отрицательный, False Negative), пропуск цели	Ошибки нет (истинно-отрицательный, True Negative)

После классификации моделью всех обучающих примеров определяется число правильных классификаций и число ошибок I и II рода. Затем вычисляются чувствительность и специфичность классификатора и на их основе оценивается точность бинарной модели с помощью ROC-кривой.

При проверке статистических гипотез ошибка I рода, называемая также «ложно-положительный вывод», заключается в отклонении нулевой гипотезы в случае, когда она является истинной. Ошибка II рода, называемая «ложно-отрицательный вывод» состоит в

принятии нулевой гипотезы, когда она является ложной. Поясним сказанное с помощью таблицы:

Нулевая гипотеза	Ложная	Истинная
Отклоняется	Ошибки нет	Ошибка I рода (ложно-положительный вывод)
Принимается	Ошибка II рода (ложно-отрицательный вывод)	Ошибки нет

Вероятность ошибки I рода в статистике называют уровнем значимости.

В Logitom существует визуализатор качество бинарной классификации на основе логистической регрессии, в котором рассчитывается и отображается множество метрик, в том числе и ошибки I и II рода.