2.1.4 Поправка Бонферрони

Осуществляя поиск незначимых категорий предиктора для объединения, CHAID выполняет большое количество статистических тестов для различных комбинаций категорий предиктора. Однако число таких комбинаций зависит от количества категорий, которое у каждой переменной разное. По одной переменной может оцениваться 2 варианта объединения, рассматриваться 2 таблицы сопряженности и выполняться 2 статистических теста, а по другой переменной – 10 вариантов объединения, 10 таблиц сопряженности и 10 статистических тестов. Вероятность того, что из 10 тестов хи-квадрат для второй переменной по крайней мере один из тестов дает ложное отклонение нулевой гипотезы составляет . Групповая вероятность ошибки намного больше индивидуальной вероятности ошибки *ai*. Например, если индивидуальная вероятность ошибки (*ai*) по каждому тесту = 0,05, то групповая вероятность ошибки 1–0,9510 = 0,401. Таким образом, при осуществлении множественных проверок гипотез при помощи хи-квадрат (одна проверка на каждое возможное объединение), *р*-значения недооценивают риск отклонения нулевой гипотезы, когда она верна. Например, вы можете сделать ошибочный вывод, что заемщики с разными профессиями отличаются по кредитоспособности, тогда как они на самом деле не отличаются.

Допустим, мы получили три p-значения: 0.01, 0.02 и 0.005. Если мы хотим, чтобы групповая вероятность ошибки при этом не превышала определенный уровень значимости α (например, 0.05), то, согласно методу Бонферрони, мы должны сравнить каждое из полученных *p*-значений не с α, а с α /*m*, где m – это количество возможных вариантов объединения *I* исходных категорий предиктора в *r* итоговых категорий. Деление исходного уровня значимости α на  m – это и есть *классическая* *поправка Бонферрони*. В рассматриваемом примере каждое из полученных *p*-значений необходимо было бы сравнить с 0.05/3 = 0.017. В результате мы получили, что *p*-значение для второй гипотезы (0.02) превышает 0.017 и, соответственно, у нас нет оснований отвергнуть эту гипотезу.

Вместо деления изначально принятого уровня значимости на число возможных вариантов объединения, мы могли бы умножить каждое исходное *p*-значение на это число и получить скорректированное *p*-значение. Именно этот вариант чаще всего используют в современных реализациях алгоритмов CHAID. Сравнив такие скорректированные *p*-значения с α, мы пришли бы к точно тем же выводам, что и при использовании классического варианта поправки Бонферрони:

0.01 \* 3 = 0.03 < 0.05: гипотеза отклоняется;

0.02 \* 3 = 0.06 > 0.05: нет оснований отклонить нулевую гипотезу;

0.005 \* 3 = 0.015 < 0.05: гипотеза отклоняется.

В ряде случаев при умножении исходных *p*-значений на уровень значимости результат может превысить 1. По определению, вероятность не может быть больше 1, и если это происходит, то получаемое значение просто приравнивают к 1.