

Task 4.3. Бурнышев Павел

Правительство города М. испытывает систему обнаружения нежелаемых лиц по камерам в метро. В качестве демонстрации работоспособности системы была поставлена цель: найти и задержать опасного рецидивиста по имени Николай Вальный, а также его соучастников. Была собрана выборка из 5000 снимков лица, для которых была проверена гипотеза о несовпадении этого снимка с лицами участников команды Н. Вального. Для 100 фотографий нулевая гипотеза была отвергнута на уровне значимости $\alpha = 0.05$.

****Требуется:**** 1. Определить, в чем недостаток данного подхода и как можно его улучшить. 2. Предложить наилучший, на ваш взгляд, способ для повышения качества данного решения. 3. Какую меру качества контролирует данный способ? Какие гарантии он предоставляет? 4. В чём недостатки данного способа? 5. Как изменилась мощность при использовании предложенного вами способа относительно изначальной процедуры проверки? 6. Известно, что все 5000 фотографий были сделаны для разных людей, и правительство хочет, чтобы система ни в коем случае не упустила преступников. Ответьте на те же вопросы из пунктов 2-4. 7. Как изменилась мощность при использовании предложенного вами способа относительно предыдущего способа?

Решение:

1. Недостаток: отсутствие поправки на множественную проверку.
2. Надо ввести поправку наиболее мощную поправку FWER - Холма.
3. Метод Холма контролирует FWER - групповую ошибку первого рода. Вероятность ошибки первого рода не выше α .
4. Надо учесть характер зависимости статистик
5. Мощность стала ниже, так как уменьшилась вероятность отклонить нулевую гипотезу.
6. Правительство хочет снизить ошибку вероятности первого рода, используем FDR вместо FWER и метод Бенджамини-Хохберга, который обеспечивает контроль над FDR на уровне α .
7. Мощность увеличилась (теперь следим за долей ложно отвергнутых гипотез).