Задача. Построить модель для определения региона, где добыча принесёт наибольшую прибыль.

Решение.

Без бутстрепа из трех валидационных выборок по потенциальной прибыли побеждал регион 0. Однако, после процедуры изъятия из наших валидационных выборок тысячи подвыборок мы получили распределения потенциальной прибыли по регионам, из которых мы выбираем регион 1. Причины:

1. Средняя прибыль в регионе 1 выше, чем в остальных.
2. Риск убытков минимален (0.6%), тогда как в двух других регионах он выше заданной границы в 2.5%.
3. Если мы посмотрим на доверительные интервалы, то увидим, что в регионе 1 95%-й доверительный интервал полностью прибыльный, в отличие от двух других регионов, где мы видим отрицательные нижние границы интервалов. Также разброс вероятных значений прибыли в регионе 0 ниже, чем в остальных двух регионах.

Все эти плюсы обеспечены точностью модели. Фактически запасы нефти в регионе 1 меньше, чем у конкурентов.