

Поиск отличия двух выборок с максимальным средним при проверке гипотезы однородности.

Научный руководитель – Шкляев Александр Викторович

Харченко М.А.¹, Зайцева А.В.²

1 - Московский государственный университет имени М.В.Ломоносова, Механико-математический факультет, Кафедра математической статистики и случайных процессов, Москва, Россия, *E-mail: maksim.kharchenko@math.msu.ru*; 2 - Московский государственный университет имени М.В.Ломоносова, Механико-математический факультет, Кафедра математической статистики и случайных процессов, Москва, Россия, *E-mail: anna.zaitceva@math.msu.ru*

<div>[12 pt, russian]article</div><div>[T1]fontenc</div><div>ulem</div><div>[utf8]luain

1. Тезисы к конференции Ломоносов

</div><div>Пусть есть набор независимых случайных величин $X_{i,j}$, $i \leq n_j$, $j \leq k$, где $X_{i,j} \sim F_j$. Имеется k выборок из различных распределений. Пусть μ_j – математическое ожидание распределения F_j , $\mu_{(1)} \leq \dots \leq \mu_{(k)}$ – вариационный ряд из них (то есть упорядоченные по возрастанию величины). Задача – проверить гипотезу $H_0 : \mu_{(k)} = \mu_{(k-1)}$ с альтернативой $H_1 : \mu_{(k)} > \mu_{(k-1)}$.

</div><div>Постановка задачи происходит из проблематики АБ тестирования – мы хотим понять является ли распределение с наибольшим средним наиболее "выгодным" или имеется целый блок из нескольких одинаково выгодных распределений. Например, если в реализации АБ-теста переход на дизайн с максимальным средним требует больших финансовых затрат, но не имеет значимых отличий от дизайна со вторым средним, то можно использовать и второй дизайн.

</div><div>При этом возможны несколько постановок: а) F_j – нормальные распределения с одинаковыми дисперсиями, но, возможно, разными средними, б) F_j – нормальные распределения с, вообще говоря, разными дисперсиями, в) F_j – произвольные, но обладают конечными 1-м и 2-м моментами, г) F_j – произвольные распределения.

</div><div>В докладе будет предложен критерий, построенный на основе критерия обобщенного отношения правдоподобий ([1]), для случая $k=2, 3$ в постановках а), б), в) при дополнительных ограничениях. Более конкретно, в случаях б) и в) предполагается, что у распределений с большими средними больше дисперсия, а в случае в) критерий асимптотический.</div><div>Моделирование работы критерия для различных k демонстрирует частоту ошибки 1-го рода на уровне 0.05.

</div><div>Для случая k нормальных выборок с равными дисперсиями были посчитаны размеры выборок на основании вероятности ошибки второго рода.

</div><div>В докладе будет описана теория построения такого критерия, описаны решенные и открытые задачи и приведены результаты численных экспериментов.

</div><div> </div><div>

Список литературы

</div><div>Кельберт М., Сухов Ю. Вероятность и статистика в примерах и задачах. Том 1. Основные понятия теории вероятностей и математической статистики. – Litres, 2022.</div><div>

</div><div></div><div> </div><div>