

Дедлайн: 2022-11-06, 21:00.

1. Величина X принимает три возможных значения: 1, 2 и 3. Вероятность q приписывает им значения 0.2, 0.3 и 0.5 соответственно.

- а) Найдите вероятность p , которая максимизирует дивергенцию $KL(q||p)$.
- б) Найдите вероятность p , которая максимизирует дивергенцию $KL(p||q)$.

Комментарий: в условии чётко написано, что других значений у X вообще нет, но трактовка, что якобы они есть, а вероятность q приписывает им нулевые веса, также будет засчитана.

2. Илон Маск максимизирует свой долгосрочный доход. Каждый день он играет в приятную лотерею, где с вероятностью 0.3 ставка будет увеличена в три раза, с вероятностью 0.2 — в два раза, а с вероятностью 0.5 полностью сгорит.

- а) Какой процент своего благосостояния ему следует ставить каждый день?
- б) Чему равна оптимальная долгосрочная дневная доходность?

3. В банке 10 независимых клиентских «окошек». В момент открытия в банк вошло 10 человек. Каждый клиент встал к отдельному окошку. Других клиентов банке в этот день не было. Время обслуживания одного клиента распределено экспоненциально с параметром λ и независимо между окошками.

Оцените параметр λ методом максимального правдоподобия и оцените дисперсию полученной оценки параметра в каждой из ситуаций:

- а) Менеджер наблюдал за окошками в течение получаса и записывал время обслуживания клиента. Окошко-1 обслужило своего клиента за 10 минут, окошко-2 обслужило своего клиента за 20 минут, остальные окошки еще обслуживали своих первых клиентов в тот момент, когда менеджер удалился.
- б) Клиент из окошка-10 обратил внимание, что окошко-1 обслужило своего клиента за 10 минут, окошко-2 обслужило своего клиента за 20 минут, а остальные окошки еще обслуживали своих первых клиентов в тот момент, когда клиент из окошка-10 ушел по окончании своего обслуживания.

Подсказка: при невозможности явного решения можно максимизировать полученную функцию правдоподобия на компьютере численно.

4. Винни-Пух стоит на пороге сенсационного открытия. Вполне возможно, что вид дерева влияет на правильность пчёл. Он собрал наблюдения за 200 дней в одну табличку, в каждой клетке указано количество дней. События, происходящие в разные дни, независимы.

	Дуб	Берёза	Ёлка
Правильные	80	30	20
Неправильные	30	20	20

- Сформулируйте гипотезу о том, что вид дерева не связан с правильностью пчёл с помощью уравнений. Проверьте эту гипотезу с помощью LR, LM и теста Вальда на уровне значимости 5%.
 - Постройте 95% доверительный интервал для разницы доли правильных пчёл на берёзах и дубах.
5. Винни-Пух *точно знает*, что среди осмотренных им за неделю 7 деревьев с мёдом, *ровно* 5 деревьев были с правильными пчёлами, и *ровно* 2 дерева — с неправильными. Логарифм количества меда для хороших и плохих пчёл распределен нормально, $\mathcal{N}(\mu_g, 1)$ и $\mathcal{N}(\mu_b, 1)$, соответственно.
- Аккуратно опишите ЕМ-алгоритм для данной ситуации.
 - В чем отличие от ситуации с независимыми наблюдениями?
6. Риши Сунак по вступлении в должность сразу принялся одновременно тестировать четыре верных гипотезы H_0^i по независимым наборам данным. Он поставил в качестве цели $FWER \leq 0.05$ и использует для реализации две поправки: Бонферонни и Холма-Бонферонни. Каким будет фактический $FWER$ в каждом из этих случаев?
7. [бонус] Здесь можно написать всё, что угодно. Поделиться классной новой задачкой, книжкой или фильмом, пожаловаться на что-то, никаких ограничений!