Машинное обучение
Домашнее задание

13 марта 2018 г.

Непрерывная случайная величина y подчиняется показательному закону распределения вероятностей, если её плотность имеет следующий вид:

$$p(y|\theta) = \frac{1}{\theta} e^{-\dfrac{y}{\theta}} [y \geq 0], \quad a\theta > 0, \quad \mathbb{E}y = \theta$$

Параметр θ в статистическом смысле характеризует среднюю продолжительность времени до наступления нового события. Тогда параметр $\frac{1}{\theta}$ отвечает за среднее число событий в единицу времени. Примерами случайных величин, подчинённых показательному закону распределения, могут быть:

- Продолжительность телефонного разговора;
- Затраты времени на обслуживания покупателя;
- Период времени работы прибора между поломками;
- Промежутки времени между появлениями машин на автозаправочной станции.

Задача 0.1. *Стоит задача прогнозирования продолжительности безотказной работы некоторого устройства, среднее время работы до поломки которого равно:* $\mathbb{E}y = \theta$.

1. *Постройте линейный прогнозирующий алгоритм* $y^\star = a(\langle w, x \rangle)$, *позволяющий решить задачу.*
2. *Методом максимального правдоподобия определите функционал качества* $\mathcal{L}(y, w, x)$ *для оптимизации весов* w.
3. *И выпишите формулы обновления весов* w *с помощью метода полного градиентного спуска в матричной форме.*

Задача 0.2. *Пусть* $p(y|\eta) = \dfrac{h(y)}{g(\eta)} e^{\eta s(y)}$ *– экспоненциальная форма распределения. Докажите, что вторая производная логарифма нормировочной константы является дисперсией достаточной статистики:* $\dfrac{d^2}{d\eta^2} \ln g(\eta) = \mathbb{D}[s(y)|\eta] > 0$