

Семинар 1. Байесовские рассуждения

Курс: Байесовские методы в машинном обучении, 2016

1. В результате медицинского обследования один из тестов выявил у человека серьезное заболевание. Данный тест имеет высокую точность 99% (вероятность позитивного ответа при наличии заболевания 99%, вероятность отрицательного ответа при отсутствии заболевания также 99%). Однако, выявленное заболевание является достаточно редким и встречается только у одного человека на 10000. Вычислить вероятность того, что у обследуемого человека действительно есть выявленное заболевание.
2. Рассмотрим следующую вероятностную модель. Студент на семинаре отвечает плохо ($c = 1$) или хорошо ($c = 0$) в зависимости от наличия/отсутствия депрессии ($d = 1$ или $d = 0$) и от участия в вечеринке накануне ($v = 1$ или $v = 0$). Участие в вечеринке также может приводить к тому, что у студента болит голова ($\gamma = 1$), а плохой ответ на семинаре влечёт недовольство преподавателя ($\pi = 1$). Причинно-следственные связи в модели показаны на рис. 1, а вероятности задаются как:

| | | | | | | | |
|-----------------|---|---|-------------------|---|----------------|---|------------------|
| $p(c = 1 v, d)$ | v | d | $p(\gamma = 1 v)$ | v | $p(\pi = 1 c)$ | c | |
| 0.999 | 1 | 1 | 0.9 | 1 | 0.95 | 1 | $p(v = 1) = 0.2$ |
| 0.9 | 1 | 0 | 0.2 | 0 | 0.5 | 0 | $p(d = 1) = 0.4$ |
| 0.9 | 0 | 1 | | | | | |
| 0.01 | 0 | 0 | | | | | |

Требуется определить $p(v = 1|\gamma = 1)$, $p(v = 1|\pi = 1)$ и $p(v = 1|\pi = 1, \gamma = 1)$.

3. Рассмотрим две независимые случайные величины, распределённые по закону Пуассона: $x_1 \sim \text{Poiss}(\lambda_1)$, $x_2 \sim \text{Poiss}(\lambda_2)$, т.е. $p(x_i = k) = \exp(-\lambda_i) \frac{\lambda_i^k}{k!}$, $k = 0, 1, 2, \dots$. Доказать, что $x_1 + x_2 \sim \text{Poiss}(\lambda_1 + \lambda_2)$.
4. Пусть x_1, x_2, \dots, x_N – независимая выборка из распределения $\text{Poiss}(\lambda)$. Требуется оценить λ с помощью метода максимального правдоподобия.
5. Пусть x_1, x_2, \dots, x_N – независимая выборка из распределения $\text{Poiss}(\lambda)$. Требуется найти байесовскую оценку λ как мат.ожидание апостериорного распределения $p(\lambda|x_1, \dots, x_N)$, если в качестве априорного распределения выступает гамма-распределение $p(\lambda) = \mathcal{G}(\lambda|a, b) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} \exp(-b\lambda)$, $a, b > 0$.

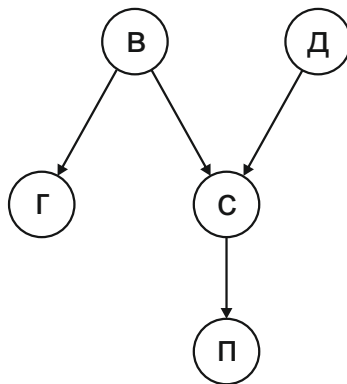


Рис. 1: Причинно-следственные связи для задачи 2

Замечание про сравнение оценок из задач 4 и 5:

1. про неинформативные априорные распределения можно подробнее прочесть в книге Kevin Patrick Murphy. Machine Learning: a Probabilistic Perspective. в главе 5.4. Priors.
2. то, что априорное распределение $p(\lambda) \sim \frac{1}{\lambda}$ является равномерным праером на $\log \lambda$, можно легко понять через законы распределения функций от случайной переменной, которые понятно описаны [здесь](#).