

Отчет о первой практической работе

Дарья Дятлова, вариант 2

Задание 1

Заданы 2 вероятностные модели:

$$\begin{aligned} p(a, b, c, d) &= p(d|c)p(c|a, b)p(a)p(b), \\ d|c &\sim c + \text{Bin}(c, p_3), \\ c|a, b &\sim \text{Bin}(a, p_1) + \text{Bin}(b, p_2), \\ a &\sim \text{Unif}[a_{\min}, a_{\max}], \\ b &\sim \text{Unif}[b_{\min}, b_{\max}]. \end{aligned} \tag{1}$$

$$\begin{aligned} p(a, b, c, d) &= p(d|c)p(c|a, b)p(a)p(b), \\ d|c &\sim c + \text{Bin}(c, p_3), \\ c|a, b &\sim \text{Pois}(ap_1 + bp_2), \\ a &\sim \text{Unif}[a_{\min}, a_{\max}], \\ b &\sim \text{Unif}[b_{\min}, b_{\max}]. \end{aligned} \tag{2}$$

Тогда априорные вероятности a, b : $p(v) = \frac{1}{v_{\max} - v_{\min} + 1}$

- $p(a) = \frac{1}{90-75+1} = \frac{1}{16}$;
- $p(b) = \frac{1}{600-500+1} = \frac{1}{101}$.

В первой вероятностной модели мы определили $p(c|a, b)$ как сумму биномиальных распределений на a и b , из чего следует, что:

- $p(c) = \sum_{a,b} p(c|a, b) \cdot p(a) \cdot p(b)$;
- $p(c) = \sum_a \cdot \sum_b \cdot \sum_{i:0:n} p(a) \cdot p(b) \cdot \binom{a}{i} \cdot p_1^i \cdot (1-p_1)^{a-i} \cdot \binom{b}{n-i} \cdot p_2^i \cdot (1-p_2)^{b-n+i}$.

Во второй модели $p(c|a, b)$ задано как пуассоновское распределение $(\frac{\lambda^k}{k!} e^{-\lambda})$, параметризованное в нашем случае $\lambda = p_1 \cdot a + p_2 \cdot b$, следовательно:

- $p(c) = \sum_a \sum_b p(a) \cdot p(b) \cdot \frac{p_1 \cdot a + p_2 \cdot b^k}{k!} e^{-(p_1 \cdot a + p_2 \cdot b)}$.

Вероятность $p(d) = \sum_c p(d|c) \cdot p(c)$:

- $p(d = k) = \sum_{i:0:a+b} p(i) \cdot p(\text{Bin}(i, p_3) = k - i)$.

Апостериорные вероятности:

- a и b независимы, значит $p(b|a) = p(b)$;
- $p(c|a) = p(c|a, b) \cdot p(b)$ (в зависимости от модели подставляем $p(c|a, b)$);
- $p(c|b) = p(c|a, b) \cdot p(a)$ (в зависимости от модели подставляем $p(c|a, b)$);
- $p(b|d) = \frac{p(b) \cdot p(d|b)}{p(d)} = \frac{\sum_c p(b) \cdot p(d|c) \cdot p(c|b)}{p(d)}$;
- $p(b|a, d) = \frac{\sum_c p(d|c) \cdot p(c|a, b) \cdot p(b)}{\sum_c p(d|c) \cdot p(c|a)}$.

Задание 2

Матожидания априорных распределений:

- $Ea = \frac{a_{max} + a_{min}}{2} = \frac{90 + 75}{2} = 82.5;$
- $Eb = \frac{b_{max} + b_{min}}{2} = \frac{500 + 600}{2} = 550;$
- $Ec = \sum_c c \cdot p(c) = 14;$
- $Ed = \sum_d d \cdot p(d) = 18.$

Дисперсии априорных распределений вычисляются по формуле $D(v) = Ev^2 - (Ev)^2$, код для вычисления значений приведен в основном файле; результаты вычислений:

- $D a = 21;$
- $D b = 850;$
- $D c = 13$ – для модели 1 и 14 для модели 2;
- $D d = 25$ – для модели 1 и 27 для модели 2.

graphicx

Задание 3

Графики уточнения прогноза для величины b – количество на курсе студентов с непрофильных факультетов, по мере поступления матожидания числа студентов поступивших на профильный факультет и/или матожидания общего числа студентов, записавшихся на данной лекции для моделей 1 и 2:

