Отчет по Практическому заданию №1:

Байесовские рассуждения

Григорьев Илья Андреевич, 417 группа

Вариант №2

1 Теория

$a_{min} = 75, a_{max} = 90$

$b_{min} = 500, b_{max} = 600$

$p_1 = 0,1$; $p_2 = 0,01$; $p_3 = 0,3$

$c \in [0, a_{max} + b_{max}]$

$d \in [0, 2(a_{max} + b_{max})]$

№2.

Первая модель:

$p(a,b,c,d) = p(d|c)p(c|a,b)p(a)p(b)$

$d|c \sim c + Bin(c, p_3)$

$c|a,b \sim Bin(a, p_1) + Bin(b, p_2)$

$a \sim U[a_{min}, a_{max}]$

$b \sim U[b_{min}, b_{max}]$

$\mathbb{E}a = 82,5$, $\mathbb{D}a = \frac{16^2 - 1}{12} = 21,25$

$\mathbb{E}b = 550$, $\mathbb{D}b = \frac{101^2 - 1}{12} = 850$

$\mathbb{E}c = \mathbb{E}\, Bin(a, p_1) + \mathbb{E}\, Bin(b, p_2) = 0,1a + 0,01b$

$\mathbb{D}c = \mathbb{D}\, Bin(a, p_1) + \mathbb{D}\, Bin(b, p_2) = 0,09a + 0,0099b$

$$\mathbb{E}d = \mathbb{E}c + \mathbb{E}Bin(c, p_3) = 0{,}1a + 0{,}01b + 0{,}3c$$

$$\mathbb{D}d = \mathbb{D}c + \mathbb{D}Bin(c, p_3) + 2\,cov(c, Bin(c, p_3)) =$$

$$= 0{,}09a + 0{,}0099b + 0{,}21c + 2\,\mathbb{E}cBin(c, p_3) - 0{,}06ac - 0{,}006bc$$

Вторая модель:

$$p(a, b, c, d) = p(d|c)p(c|a, b)p(a)p(b)$$

$$d|c \sim c + Bin(c, p_3)$$

$$c|a, b \sim Poiss(ap_1 + bp_2)$$

$$a \sim \mathcal{U}[a_{min}, a_{max}]$$

$$b \sim \mathcal{U}[b_{min}, b_{max}]$$

$$\mathbb{E}a = 82{,}5 \quad ; \quad \mathbb{D}a = 21{,}25$$

$$\mathbb{E}b = 550 \quad ; \quad \mathbb{D}b = 850$$

$$\mathbb{E}c = 0{,}1a + 0{,}01b$$

$$\mathbb{D}c = 0{,}1a + 0{,}01b$$

$$\mathbb{E}d = 0{,}1a + 0{,}01b + 0{,}3c$$

$$\mathbb{D}d = \mathbb{D}c + \mathbb{D}Bin(c, p_3) + 2\,cov(c, Bin(c, p_3)) =$$

$$= 0{,}1a + 0{,}01b + 0{,}21c + 2\,\mathbb{E}cBin(c, p_3) - 0{,}06ac - 0{,}006bc$$

№3.

$$\mathbb{E}\, p(b) = 550, \quad \mathbb{D}\, p(b) = 850$$

$$p(b|a) = p(b), \quad \mathbb{E}\, p(b|a) = 550, \quad \mathbb{D}\, p(b|a) = 850$$

$$p(b|d) = \frac{p(d|b)\, p(b)}{\sum_{b=500}^{600} p(d|b)\, p(b)} = \frac{p(d|b)}{\sum_{b=500}^{600} p(d|b)}$$

$$p(b|d=\mathbb{E}d) = \frac{p(d=\mathbb{E}d\,|\,b)}{\sum_{b=500}^{600} p(d=\mathbb{E}d\,|\,b)} =$$

$$= \frac{p(d=\mathbb{E}d\,|\,c)\, p(c|a,b)\, p(a)}{\sum_{b=500}^{600} p(d=\mathbb{E}d\,|\,c)\, p(c|a,b)\, p(a)}$$

$$\mathbb{E}\, p(b|d=\mathbb{E}d) = \sum_{b=500}^{600} b \cdot p(b|d=\mathbb{E}d)$$

$$p(b|a,d) = \frac{p(a,d|b)\, p(b)}{\sum_{b=500}^{600} p(a,d|b)\, p(b)} = \frac{p(a,d|b)}{\sum_{b=500}^{600} p(a,d|b)} =$$

$$= \frac{p(d|a,b)\, p(a)}{\sum_{b=500}^{600} p(d|a,b)\, p(a)} = \frac{p(d|a,b)}{\sum_{b=500}^{600} p(d|a,b)} = \frac{p(d|c)\, p(c|a,b)}{\sum_{b=500}^{600} p(d|c)\, p(c|a,b)}$$

$$p(b \mid a = \mathbb{E}a, d = \mathbb{E}d) = \frac{p(d = \mathbb{E}d \mid c)\, p(c \mid a = \mathbb{E}a, b)}{\sum_{b=500}^{600} p(d = \mathbb{E}d \mid c)\, p(c \mid a = \mathbb{E}a, b)}$$

$$\mathbb{E}\, p(b \mid a = \mathbb{E}a, d = \mathbb{E}d) = \sum_{b=500}^{600} b \cdot p(b \mid a = \mathbb{E}a, d = \mathbb{E}d)$$

№ 5.

$$p(c) = p(c \mid a, b)\, p(a)\, p(b)$$

$$p(c \mid a) = p(c \mid a, b)\, p(b)$$

$$p(c \mid b) = p(c \mid a, b)\, p(a)$$

$$p(b \mid a) = p(b)$$

$$p(b \mid d) = \frac{p(d \mid c)\, p(c \mid a, b)\, p(a)}{\sum_{b=500}^{600} p(d \mid c)\, p(c \mid a, b)\, p(a)}$$

$$p(b \mid a, d) = \frac{p(d \mid c)\, p(c \mid a, b)}{\sum_{b=500}^{600} p(d \mid c)\, p(c \mid a, b)}$$

$$p(d) = p(d \mid c)\, p(c \mid a, b)\, p(a)\, p(b)$$

2 Анализ

2.1 Задание 2

Первая модель:

$E[p(a)] = 82.5 \quad Var[p(a)] = 21.25$

$E[p(b)] = 550 \quad Var[p(b)] = 850$

$E[p(c)] = 13.75 \quad Var[p(c)] = 13.1675$

$E[p(d)] = 17.875 \quad Var[p(d)] = 25.14$

Вторая модель:

$E[p(a)] = 82.5 \quad Var[p(a)] = 21.25$

$E[p(b)] = 550 \quad Var[p(b)] = 850$

$E[p(c)] = 13.75 \quad Var[p(c)] = 14.0475$

$E[p(d)] = 17.875 \quad Var[p(d)] = 26.6278$

Вторая модель является упрощенным приближением первой модели. Все математические ожидания у них совпадают, что говорит о несмещенности данного приближения. Дисперсии случайных величин c и d у второй модели больше, чем у первой.

2.2 Задание 3

По мере прихода новой косвенной информации происходит уточнение прогноза для величины b в обеих моделях. Параметры a и d равны матожиданиям своих априорных распределений, округленных до ближайшего целого. На графиках видно, что с появлением новой информации распределение b меняется – значения близкие к матожиданию b становятся более вероятными, в то время как априорное распределение b равномерное. Информация про случайную величину a не меняет распределение b, так как эти случайные величины независимы. Информация про случайную величину d уточняет прогноз b, а информация про случайные величины a и d уточняет прогноз b еще сильнее. Это хорошо видно из матожиданий и дисперсий этих распределений. Чем меньше дисперсия, тем точнее прогноз.

Первая модель:

$E[p(b)] = 550 \quad Var[p(b)] = 850$

$E[p(b|a)] = 550 \quad Var[p(b|a)] = 850$

$E[p(b|d)] = 550.0727 \quad Var[p(b|d)] = 848.0379$

$E[p(b|a, d)] = 550.0363 \quad Var[p(b|a, d)] = 848.0311$

Вторая модель:

$E[p(b)] = 550 \quad Var[p(b)] = 850$

$E[p(b|a)] = 550 \quad Var[p(b|a)] = 850$

$E[p(b|d)] = 550.0978 \quad Var[p(b|d)] = 848.1280$

$E[p(b|a, d)] = 550.0635 \quad Var[p(b|a, d)] = 848.1231$

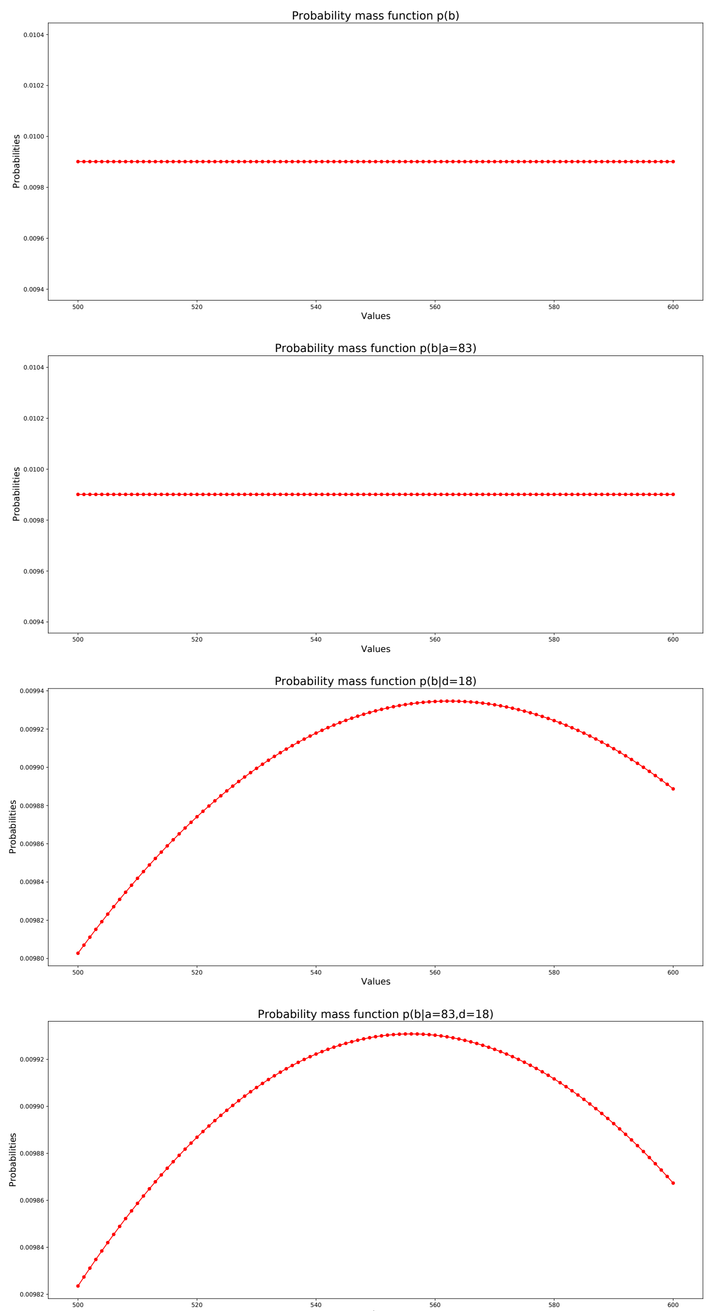


Рис. 1: Первая модель

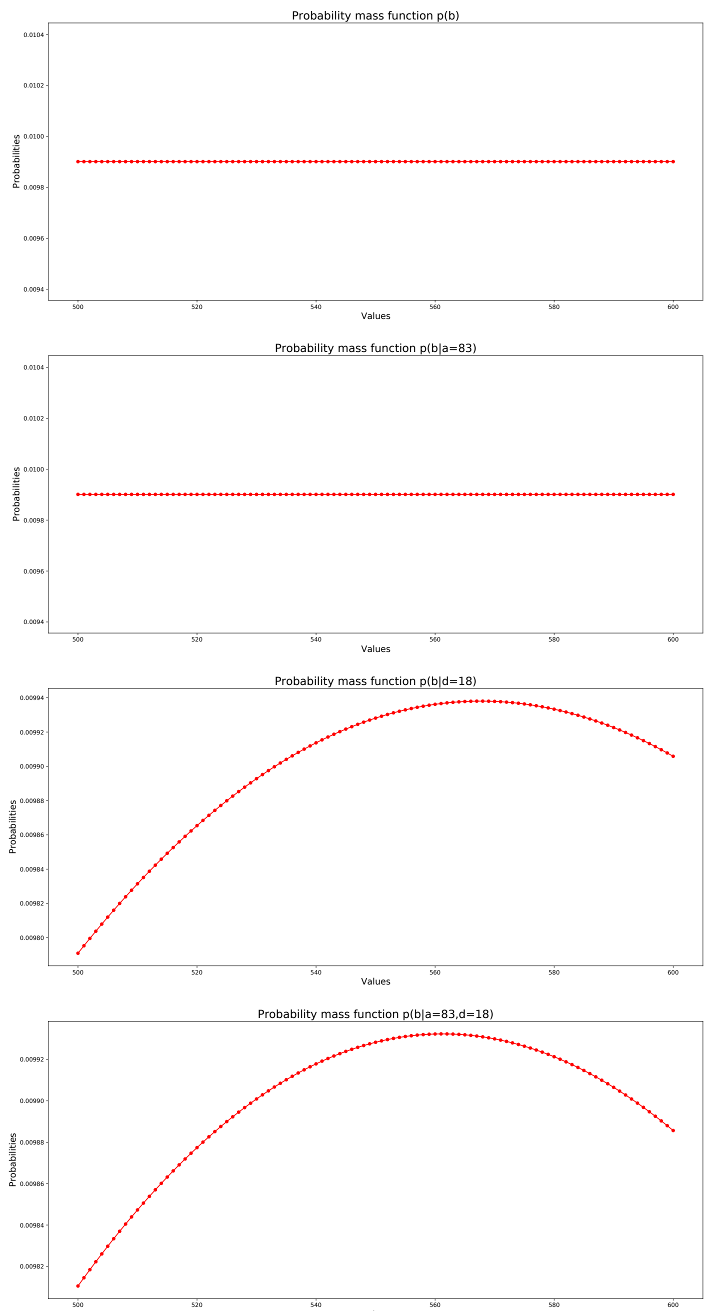
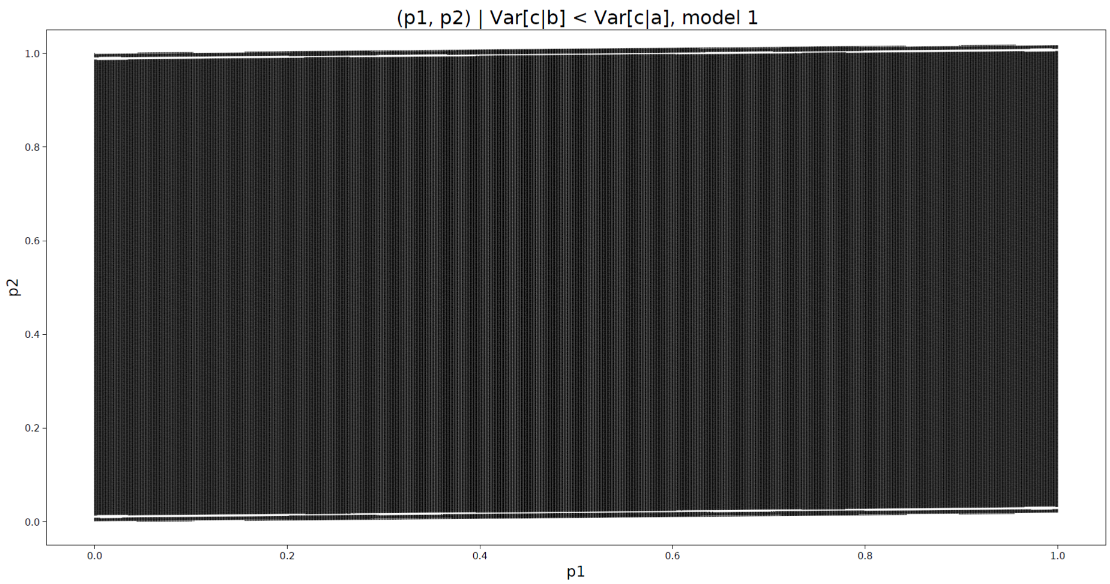


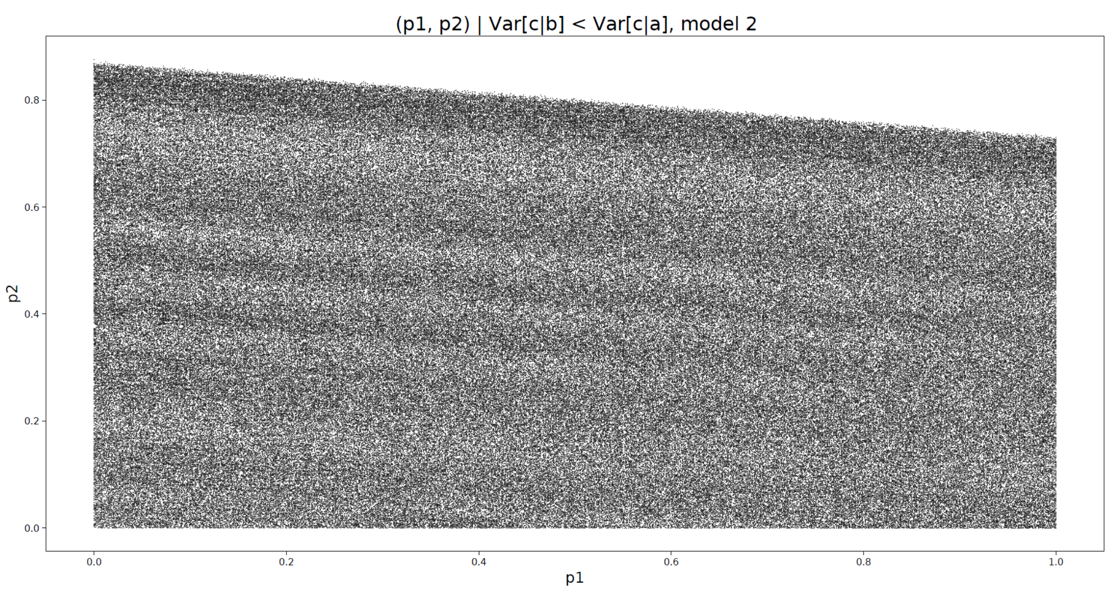
Рис. 2: Вторая модель

2.3 Задание 4

Было проведено исследование, при каких соотношениях параметров $p1$ и $p2$ изменяется относительная важность параметров a и b для оценки величины c.

Для обеих моделей были найдены множества точек $(p1, p2)$ при которых дисперсия распределения $p(c|b)$ меньше дисперсии распределения $p(c|a)$ при a, b равных матожиданиям своих априорных распределений, округленных до ближайшего целого.





Для первой модели в 99% точках $(p1, p2)$ важность параметра b выше важности параметра a для оценки величины c. В первой модели множества $(p1, p2)|Var[c|b] < Var[c|a]$ и $(p1, p2)|Var[c|b] >= Var[c|a]$ являются линейно разделимыми, это видно из графика (две белые горизонтальные полосы).

Для второй модели в 40% точках $(p1, p2)$ важность параметра b выше важности параметра a для оценки величины c. Во второй модели множества $(p1, p2)|Var[c|b] < Var[c|a]$ и $(p1, p2)|Var[c|b] >= Var[c|a]$ не являются линейно разделимыми, это так же видно из графика (точки перемешаны).

2.4 Задание 5

Были проведены временные замеры по оценке распределений p(c), p(c|a), p(c|b), p(b|a), p(b|d), p(b|a,d), p(d) на скалярных входах для обеих моделей.

Первая модель:

p(c) – 30 ms, p(c|a) – 7 ms, p(c|b) – 1 ms, p(b|a) – 0 ns, p(b|d) – 270 ms, p(b|a,d) – 276 ms, p(d) – 56 ms

Вторая модель:

p(c) – 21 ms, p(c|a) – 5 ms, p(c|b) – 1 ms, p(b|a) – 0 ns, p(b|d) – 724 ms, p(b|a,d) – 635 ms, p(d) – 46 ms

Все распределения на скалярных входах считаются быстрее секунды. Дольше всех считаются условные распределения p(b|d) и p(b|a,d), особенно для второй модели.

2.5 Задание 6

На основе предыдущих пунктов был проведен сравнительный анализ двух моделей. Дисперсии априорных распределений $p(c)$ и $p(d)$ у второй модели больше, что свидетельствует о ее приближенности. Математические ожидания всех априорных распределений у обеих моделей совпадают, что говорит о несмещенности приближения.

Максимальная разница между моделями проявляется в четвертом задании. Это можно объяснить тем, что параметры a и b в первой модели входят каждый в свое распределение, при этом абсолютные значения величины b больше. А во второй модели параметры a и b входят в линейную комбинацию параметра одного Пуассоновского распределения.