



МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ М. В. ЛОМОНОСОВА
Факультет вычислительной математики и кибернетики

Отчет к первому практическому заданию по БММО: Байесовские рассуждения

Студент 4 курса ВМК (417 группа):
Оспанов А.М.

Москва, 2015

Содержание

1	Введение	2
2	Основная часть	3
2.1	Математические ожидания и дисперсии	3
2.2	Наблюдение $p(b d_1, \dots, d_N), p(d a, d_1, \dots, d_N)$	4
2.3	Замер времени работы	11
3	Заключение	12

1 Введение

Данный отчет написан к первому практическому заданию по БММО. Тема задания: Байесовские рассуждения. Отчет написан студентом 417 группы – Оспановым Аятом.

В данной работе были реализованы 3 и 4 вероятностные модели посещаемости курса. Были сделаны все исследования, требуемые в задании. Ipython Notebook-и поделены на 2, т.е. каждая модель в своем Notebook-е.

2 Основная часть

2.1 Математические ожидания и дисперсии

Математические ожидания и дисперсии для моделей считались аналитический. Применялись следующие свойства условного матожидания и условной дисперсии:

$$\mathbb{E}X = \mathbb{E}\mathbb{E}[X|Y]$$

$$\mathbb{D}X = \mathbb{E}\mathbb{D}[X|Y] + \mathbb{D}\mathbb{E}[X|Y]$$

В результате получились следующие формулы для вычисления матожиданий и дисперсий:

- Для модели 3:

$$\mathbb{E}a = \frac{a_{max} + a_{min}}{2}$$

$$\mathbb{E}b = \frac{b_{max} + b_{min}}{2}$$

$$\mathbb{E}c_n = p_1\mathbb{E}a + p_2\mathbb{E}b$$

$$\mathbb{E}d_n = (1 + p_3)\mathbb{E}c_n$$

$$\mathbb{D}a = \frac{(a_{max} - a_{min} + 1)^2 + 1}{12}$$

$$\mathbb{D}b = \frac{(b_{max} - b_{min} + 1)^2 + 1}{12}$$

$$\mathbb{D}c_n = p_1(1 - p_1)\mathbb{E}a + p_2(1 - p_2)\mathbb{E}b + p_1^2\mathbb{D}a + p_2^2\mathbb{D}b$$

$$\mathbb{D}d_n = p_3(1 - p_3)\mathbb{E}c_n + (1 + p_3)^2\mathbb{D}c_n$$

- Для модели 4 (различается только $\mathbb{D}c_n$):

$$\mathbb{D}c_n = p_1\mathbb{E}a + p_2\mathbb{E}b + p_1^2\mathbb{D}a + p_2^2\mathbb{D}b$$

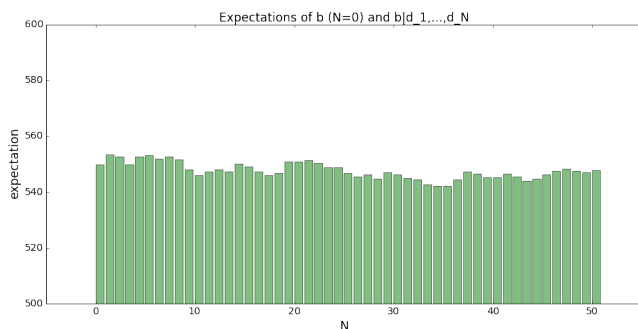
	Модель 3	Модель 4
$\mathbb{E}a$	82.5	82.5
$\mathbb{E}b$	550.0	550.0
$\mathbb{E}c_n$	13.75	13.75
$\mathbb{E}d_n$	17.875	17.875
$\mathbb{D}a$	21.25	21.25
$\mathbb{D}b$	850.0	850.0
$\mathbb{D}c_n$	13.1675	14.0475
$\mathbb{D}d_n$	25.1405	26.6277

2.2 Наблюдение $p(b|d_1, \dots, d_N), p(d|a, d_1, \dots, d_N)$

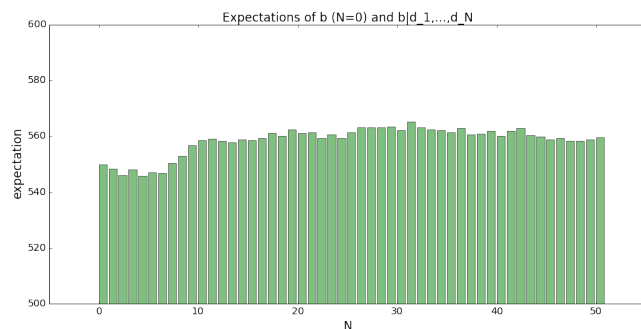
d_1, \dots, d_N сгенерирована из модели при параметрах $a = \mathbb{E}a, b = \mathbb{E}b$

Математические ожидания

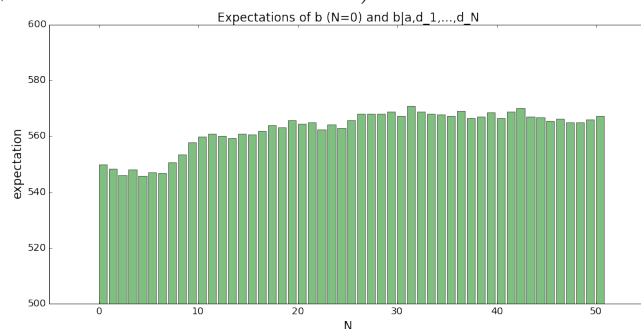
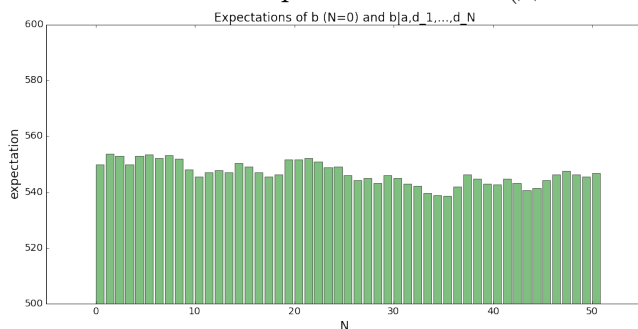
Модель 3



Модель 4

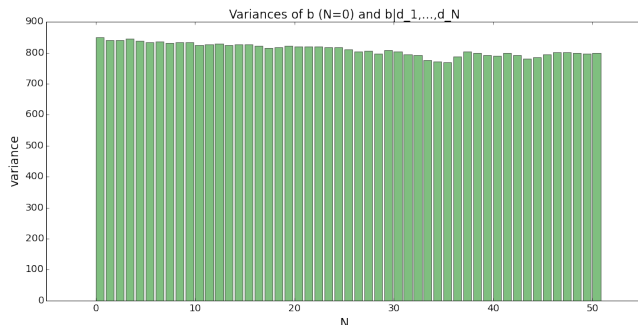


При известном a (для исследования считалось $a = \mathbb{E}a$)

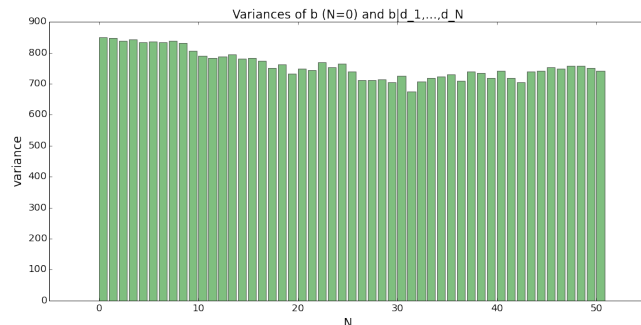


Дисперсии

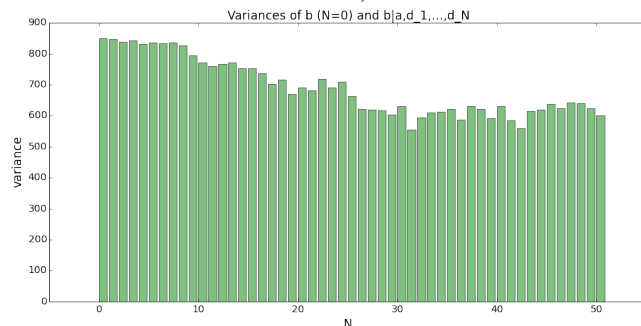
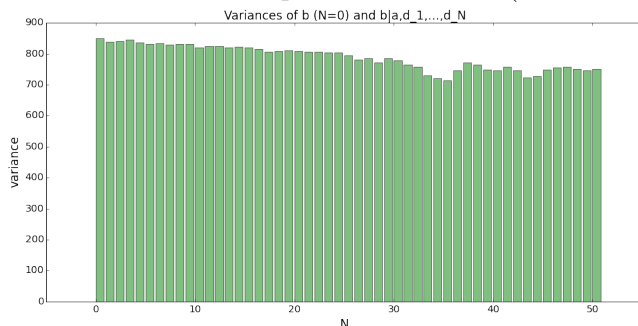
Модель 3



Модель 4



При известном a (для исследования считалось $a = \mathbb{E}a$)



Из математических ожиданий $b|d_1, \dots, d_N$ для модели 3 можно увидеть, что среднее количество студентов других факультетов уменьшается с добавлением d_i . Но при этом, в среднем, матожидание находится около 550, т.е. матожидания априорного распределения. А в модели 4 матожидание завышено (в этом мы убедимся и в дальнейших исследованиях). Если дополнительно известно значение a (в нашем случае берется матожидание априорного рас-

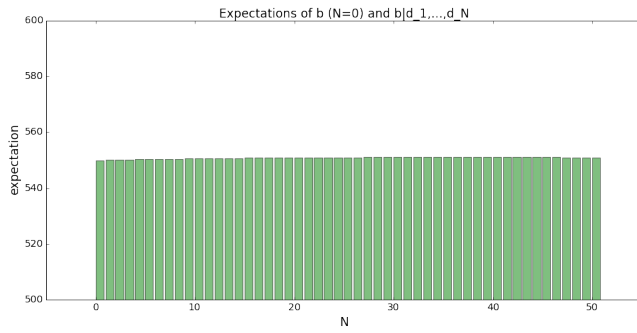
пределения a), то матожидание практически не меняется. Это можно объяснить тем, что a и b независимы.

Дисперсии же убывают с добавлением d_i , т.е. чем больше нам известно об отметившихся студентах, тем с большей точностью мы можем предсказать b .

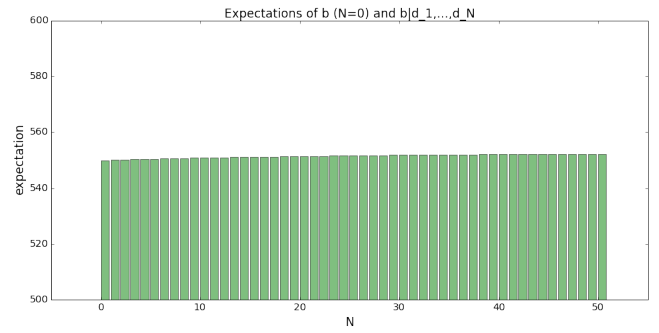
$$d_1 = \dots = d_N = \mathbb{E}d_N$$

Математические ожидания

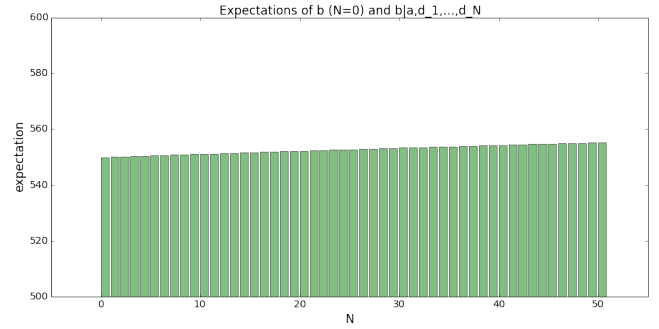
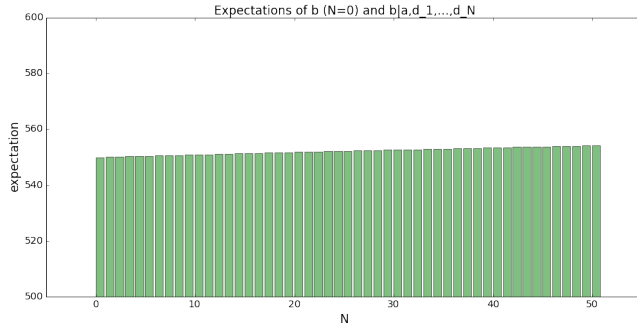
Модель 3



Модель 4

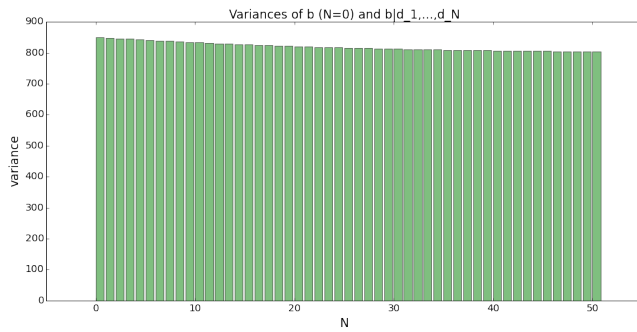


При известном a (для исследования считалось $a = \mathbb{E}a$)

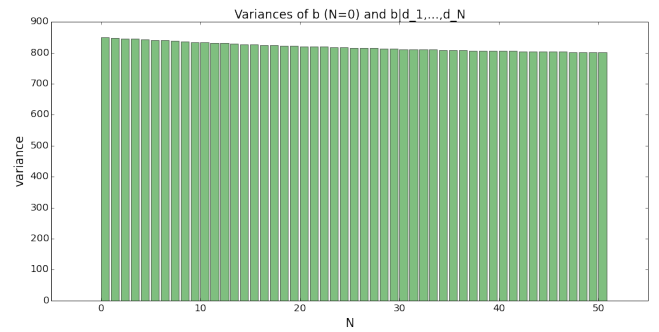


Дисперсии

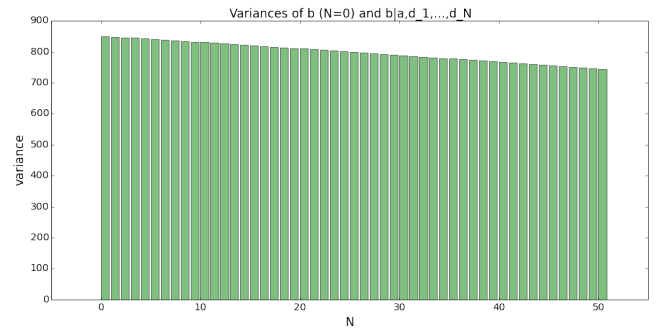
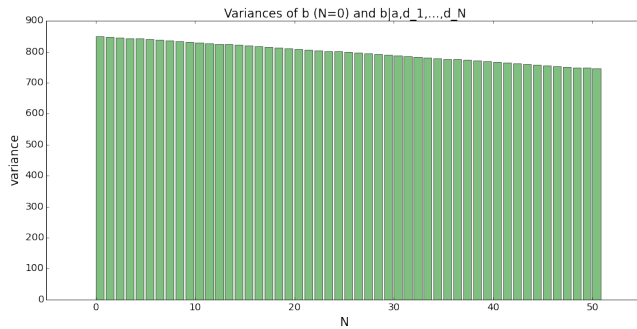
Модель 3



Модель 4



При известном a (для исследования считалось $a = \mathbb{E}a$)

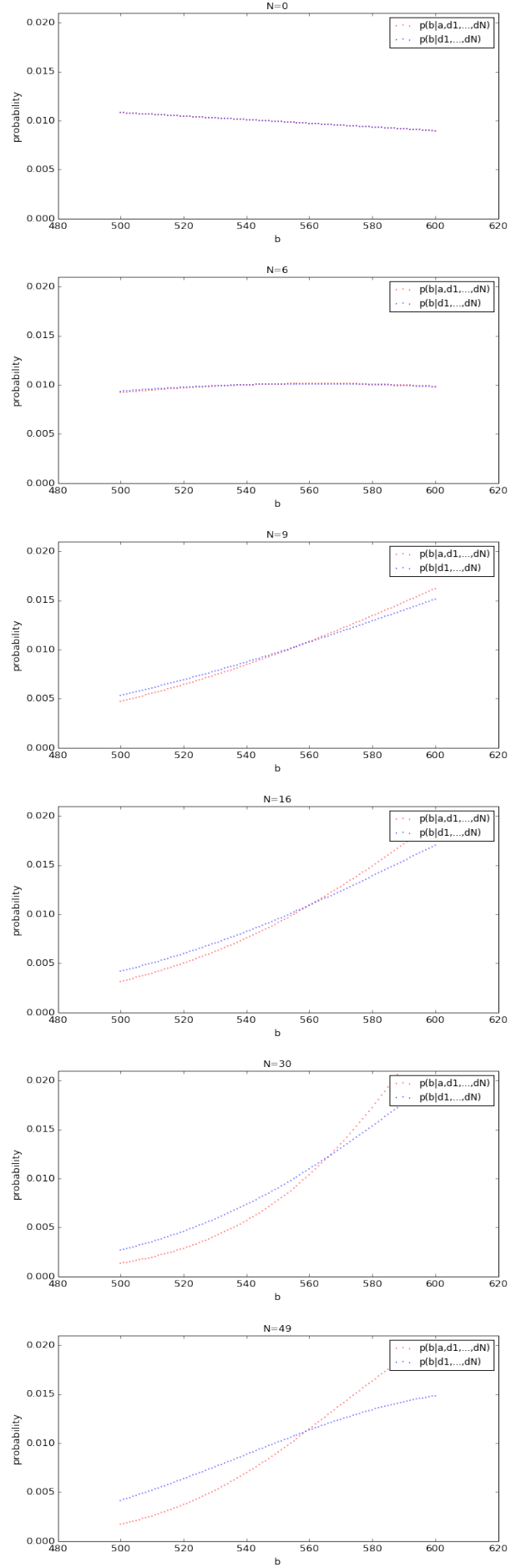
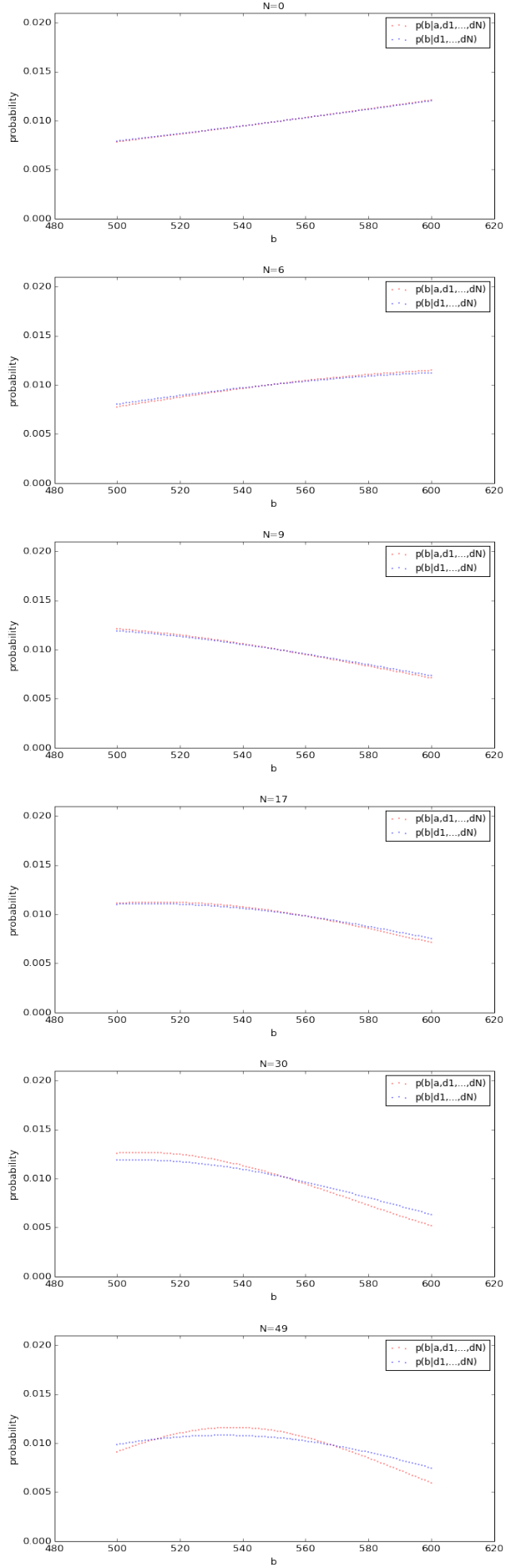


В случае равных значений d_i модели практически одинаковые, а матожидания и дисперсии монотонны.

d_1, \dots, d_N сгенерирована из модели при параметрах $a = \mathbb{E}a, b = \mathbb{E}b$

Модель 3

Модель 4

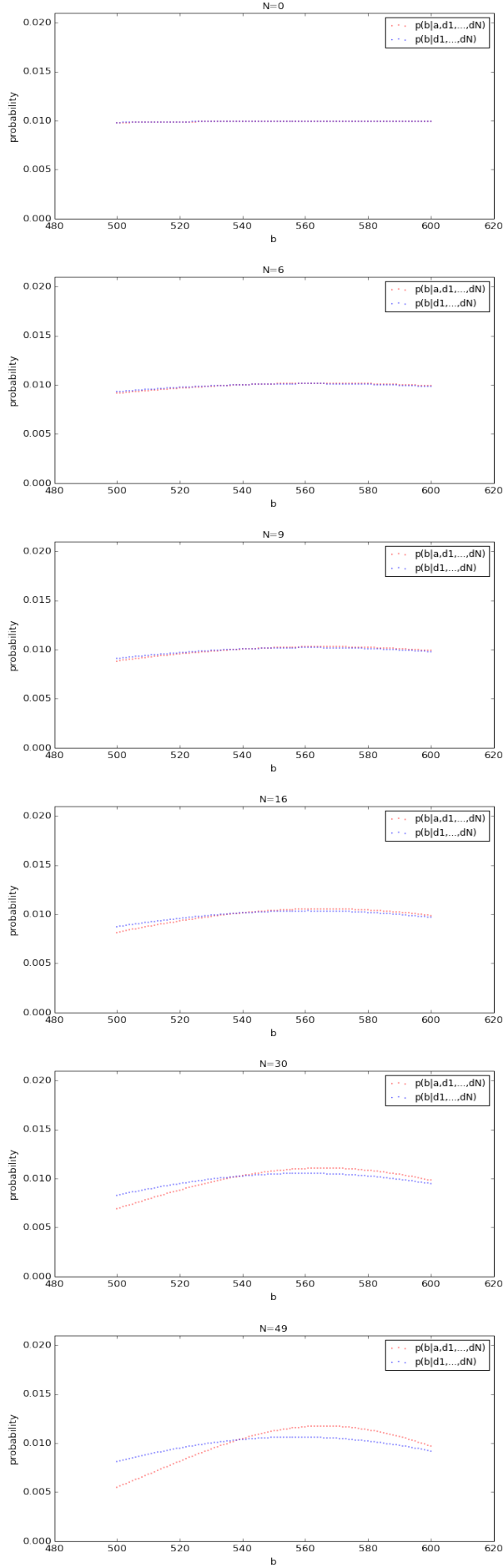


Теперь рассмотрим плотности распределений $b|a, d_1, \dots, d_N$ и $b|d_1, \dots, d_N$. Здесь мы можем видеть, что в модели 3, при добавлении d_i , вероятность того, что студентов других факультетов будет больше матожидания b уменьшается. Но в модели 4, эта вероятность растет. Т.е. погрешность прибиближения суммы биномиальных распределений Пуассоновским распределением - большая.

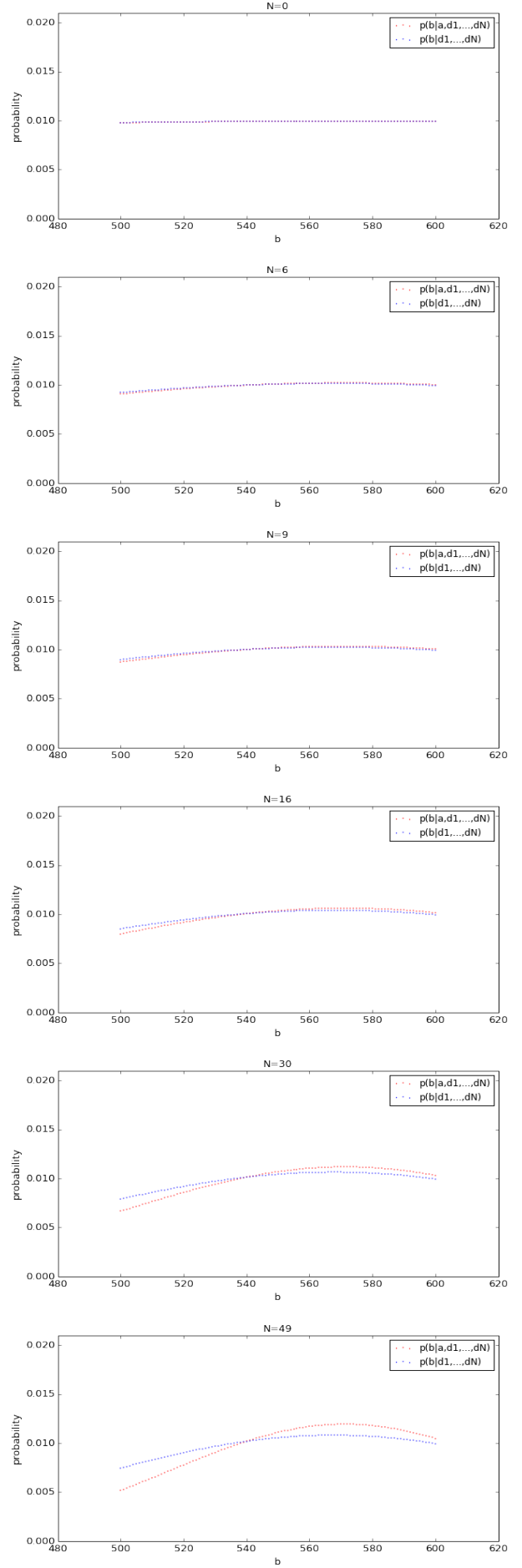
Также, если дополнительно известно значение a , то график практически не меняется. Это также объясняется независимостью a и b .

$$d_1 = \dots = d_N = \mathbb{E}d_N, a = \mathbb{E}a$$

Модель 3



Модель 4



Далее, если все d_i равны матожиданию своего априорного распределения, то погрешность между моделью 3 и моделью 4 пренебрежимо мала. Это можно объяснить тем, что берутся средние значения отметившихся студентов, тем самым оценка b выше около матожидания b . И так как сумма биномиальных распределений и распределение Пуассона имеют одинаковые матожидания, то модели являются практически одинаковыми. Также известное значение a практически не меняет распределение.

2.3 Замер времени работы

	Модель 3	Модель 4
$p(c_n)$	124 ms	35.4 ms
$p(d_n)$	140 ms	55.9 ms
$p(b d_1, \dots, d_N)$	35.1 s	36.1 s
$p(b a, d_1, \dots, d_N)$	2.17 s	2.25 s

Времена работы при нахождении $p(c_n)$ и $p(d_n)$ различаются в 2-3 раза, так как в модели 3 нужно делать свертку, что является трудоемкой задачей. А при подсчете $p(b|d_1, \dots, d_N)$ и $p(b|a, d_1, \dots, d_N)$ времена работы практически одинаковые, но большие. Поэтому разница времени для моделей из-за свертки не влияет на общее время.

3 Заключение

В итоге встает вопрос о выборе модели. Времена подсчета $p(b|d_1, \dots, d_N)$ и $p(b|a, d_1, \dots, d_N)$ для двух моделей одинаковые. Поэтому смотрим на погрешность приближения моделей. В случае, если $d_1 = \dots = d_N = \mathbb{E}d_N$, то особой разницы в моделях нет. А в общем случае когда выборка d_1, \dots, d_N генерится случайным образом, то модель 4 дает большие погрешности, что является причиной не использовать эту модель. В итоге, т.к. время работы моделей одинаковые, лучше использовать модель 3.

P.S. Если бы код писался более оптимально, то модель 4 была бы быстрее, т.к. не надо было бы считать свертку. Но при этом эта модель давала бы большие погрешности.