Московский Государственный Университет

Байесовские рассуждения

Выполнил: Курцев Д.В.

Группа: 417

Факультет Вычислительной математики и кибернетики

Кафедра Математических методов прогнозирования

Сентябрь 2022

1 p(a), p(b)

$a \sim \text{Unif}\ [a_{min}, a_{max}]$, $b \sim \text{Unif}\ [b_{min}, b_{max}]$

Найдём математическое ожидание и дисперсию для случайной величины a (для b всё аналогично):

$$\mathbb{P}(a = k) = \begin{cases} \frac{1}{n}, & n = a_{max} - a_{min} + 1, k \in [a_{min}, a_{max}] \\ 0, & \text{иначе} \end{cases}$$

Тогда:

$$\mathbb{E}\ a = \sum_{k=a_{min}}^{a_{max}} \frac{1}{n} k = \frac{1}{n} \sum_{k=a_{min}}^{a_{max}} k = \frac{1}{n} \frac{(a_{min} + a_{max})\, n}{2} = \frac{(a_{min} + a_{max})}{2}$$

Дисперсию вычислим, воспользовавшись свойством, по формуле:

$$\mathbb{D}a = \mathbb{E}a^2 - \mathbb{E}^2 a$$

$$\mathbb{E}a^2 = \frac{1}{n} \sum_{k=a_{min}}^{a_{max}} k^2 = \frac{1}{n} (\sum_{k=1}^{a_{max}} k^2 - \sum_{k=1}^{a_{min}-1} k^2) = \{\sum_{k=1}^{N} k^2 = \frac{N(N+1)(2N+1)}{6}\} =$$

$$= \frac{1}{6n}(a_{max}(a_{max}+1)(2a_{max}+1) - a_{min}(a_{min}-1)(2a_{min}-1)) = \frac{1}{6n}(3a_{max}^2 + 2a_{max}^3 + a_{max} -$$

$$-2a_{min}^3 + 3a_{min}^2 - a_{min}) \quad (=)$$

$$a_{max}^3 - a_{min}^3 = (a_{max} - a_{min} + 1)(a_{max}^2 + a_{max}a_{min} + a_{min}^2) - (a_{max}^2 + a_{max}a_{min} + a_{min}^2)$$

$$(=)\ \frac{1}{6n}(2n(a_{max}^2 + a_{max}a_{min} + a_{min}^2) - 2a_{max}^2 - 2a_{max}a_{min} - 2a_{min}^2 + 3a_{max}^2 + 3a_{min}^2 + a_{max} - a_{min}) =$$

$$= \frac{1}{6n}(2n(a_{max}^2 + a_{max}a_{min} + a_{min}^2) + a_{max}^2 + a_{min}^2 - 2a_{max}a_{min} + a_{max} - a_{min}) =$$

$$= \frac{1}{6n}(2n(a_{max}^2 + a_{max}a_{min} + a_{min}^2) + a_{max}(a_{max} - a_{min} + 1) - a_{min}(a_{max} - a_{min} + 1)) =$$

$$= \frac{1}{6}(2a_{max}^2 + 2a_{max}a_{min} + 2a_{min}^2 + a_{max} - a_{min})$$

Таким образом, получим, что дисперсия равна:

$$\mathbb{D}a = \mathbb{E}a^2 - \mathbb{E}^2 a = \frac{1}{6}(2a_{max}^2 + 2a_{max}a_{min} + 2a_{min}^2 + a_{max} - a_{min}) - \frac{1}{4}(a_{max}^2 - 2a_{max}a_{min} + a_{min}^2) =$$

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

Тогда получим, что

$$\mathbb{E}a = \frac{a_{min} + a_{max}}{2}, \quad \mathbb{E}b = \frac{b_{min} + b_{max}}{2}$$

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

2 p(c)

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

$$p(c) = \sum_a \sum_b p(a,b,c) = \sum_a \sum_b p(c|a,b)p(a)p(b)$$

То есть

$$\mathbb{P}(c = k) = \sum_a \sum_b \mathbb{P}(c = k|a,b)\mathbb{P}(a)\mathbb{P}(b)$$

Найдём мат ожидание $p(c)$:

$$\mathbb{E}\ c = \sum_k \mathbb{P}(c = k)\ k = \sum_k \sum_a \sum_b \mathbb{P}(c = k|a,b)\mathbb{P}(a)\mathbb{P}(b)k \quad =$$

$$= \quad \sum_a \mathbb{P}(a) \sum_b \mathbb{P}(b) \sum_k \mathbb{P}(c = k|a,b)k = \sum_a \sum_b \mathbb{P}(a)\mathbb{P}(b)\mathbb{E}(c|a,b) \quad (=)$$

$\mathbb{E}(c|a,b)$ для обеих моделей равно $ap_1 + bp_2$. Тогда получим:

$$(=) \sum_a \sum_b \mathbb{P}(a)\mathbb{P}(b)(ap_1 + bp_2) = \sum_b \mathbb{P}(b) \sum_a \mathbb{P}(a)ap_1 + \sum_a \mathbb{P}(a) \sum_b \mathbb{P}(b)bp_2 =$$

$$= p_1 \sum_b \mathbb{P}(b)\mathbb{E}a + p_2 \sum_a \mathbb{P}(a)\mathbb{E}b = p_1\mathbb{E}a + p_2\mathbb{E}b$$

Чтобы найти дисперсию, найдём мат ожидание квадрата c

$$\mathbb{E}c^2 = \sum_k \mathbb{P}(c = k)\ k^2 = \sum_k \sum_a \sum_b \mathbb{P}(c = k|a,b)\mathbb{P}(a)\mathbb{P}(b)k^2 = \sum_a \sum_b \mathbb{P}(a)\mathbb{P}(b)\mathbb{E}(c^2|a,b) =$$

$$= \sum_a \sum_b \mathbb{P}(a)\mathbb{P}(b)(\mathbb{D}(c|a,b) + \mathbb{E}^2(c|a,b))$$

Для 1 модели:

$$\sum_a \sum_b \mathbb{P}(a)\mathbb{P}(b)\mathbb{D}(c|a,b) = \sum_a \sum_b \mathbb{P}(a)\mathbb{P}(b)(ap_1(1-p_1) + bp_2(1-p_2)) =$$

$$p_1(1-p_1)\sum_b \mathbb{P}(b)\sum_a \mathbb{P}(a)a + p_2(1-p_2)\sum_a \mathbb{P}(a)\sum_b \mathbb{P}(b)b =$$

$$= p_1(1-p_1)\mathbb{E}a\sum_b \mathbb{P}(b) + p_2(1-p_2)\mathbb{E}b\sum_a \mathbb{P}(a) = p_1(1-p_1)\mathbb{E}a + p_2(1-p_2)\mathbb{E}b$$

Для 2 модели:

$$\sum_a\sum_b \mathbb{P}(a)\mathbb{P}(b)\mathbb{D}(c|a,b) = \sum_a\sum_b \mathbb{P}(a)\mathbb{P}(b)(ap_1 + bp_2) = p_1\mathbb{E}a + p_2\mathbb{E}b$$

$$\sum_a\sum_b \mathbb{P}(a)\mathbb{P}(b)\mathbb{E}^2(c|a,b) = \sum_a\sum_b \mathbb{P}(a)\mathbb{P}(b)(ap_1 + bp_2)^2 =$$

$$= \sum_a\sum_b \mathbb{P}(a)\mathbb{P}(b)(a^2p_1^2 + 2abp_1p_2 + b^2p_2^2) = p_1^2\sum_b \mathbb{P}(b)\sum_a \mathbb{P}(a)a^2 +$$

$$+p_2^2\sum_a \mathbb{P}(a)\sum_b \mathbb{P}(b)b^2 + 2p_1p_2\sum_a \mathbb{P}(a)a\sum_b \mathbb{P}(b)b = p_1^2\mathbb{E}a^2 + p_2^2\mathbb{E}b^2 + 2p_1p_2\mathbb{E}a\mathbb{E}b$$

$$\mathbb{E}^2c = (p_1\mathbb{E}a + p_2\mathbb{E}b)^2 = p_1^2\mathbb{E}^2a + p_2^2\mathbb{E}^2b + 2p_1p_2\mathbb{E}a\mathbb{E}b$$

Тогда в итоге получим:

$$\mathbb{D}_1c = p_1(1-p_1)\mathbb{E}a + p_2(1-p_2)\mathbb{E}b + p_1^2\mathbb{E}a^2 + p_2^2\mathbb{E}b^2 + 2p_1p_2\mathbb{E}a\mathbb{E}b - p_1^2\mathbb{E}^2a - p_2^2\mathbb{E}^2b - 2p_1p_2\mathbb{E}a\mathbb{E}b =$$

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

$$\mathbb{D}_2c = p_1\mathbb{E}a + p_2\mathbb{E}b + p_1^2\mathbb{D}a + p_2^2\mathbb{D}b$$

В итоге имеем:

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

$$\mathbb{D}_1\ c = p_1(1-p_1)\mathbb{E}a + p_2(1-p_2)\mathbb{E}b + p_1^2\mathbb{D}a + p_2^2\mathbb{D}b$$

$$\mathbb{D}_2\ c = p_1\mathbb{E}a + p_2\mathbb{E}b + p_1^2\mathbb{D}a + p_2^2\mathbb{D}b$$

3 p(c|a)

$$p(c|a) = \frac{p(a,c)}{p(a)} = \frac{\sum_b p(a,b,c)}{p(a)} = \frac{\sum_b p(c|a,b)p(a)p(b)}{p(a)} = \sum_b p(c|a,b)p(b)$$

$$\Rightarrow \quad \mathbb{P}(c=k|a) = \sum_b \mathbb{P}(c=k|a,b)\mathbb{P}(b)$$

Аналогично $p(c)$ найдём мат ожидание и дисперсию для $p(c|a)$:

$$\mathbb{E}\ (c|a) = \sum_k \sum_b \mathbb{P}(c=k|a,b)\mathbb{P}(b)\ k = \sum_b \mathbb{P}(b)\mathbb{E}(c|a,b) = \sum_b \mathbb{P}(b)(ap_1 + bp_2) =$$

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

$$\mathbb{E}(c^2|a) = \sum_k \sum_b \mathbb{P}(c=k|a,b)\mathbb{P}(b)k^2 = \sum_b \mathbb{P}(b)\mathbb{E}(c^2|a,b) = \sum_b \mathbb{P}(b)(\mathbb{D}(c^2|a,b) + \mathbb{E}^2(c|a,b))$$

Для 1 модели:

$$\sum_b \mathbb{P}(b)\mathbb{D}(c^2|a,b) = \sum_b \mathbb{P}(b)(p_1(1-p_1)a + p_2(1-p_2)b) = p_1(1-p_1)a + p_2(1-p_2)\mathbb{E}b$$

Для 2 модели:

$$\sum_b \mathbb{P}(b)\mathbb{D}(c^2|a,b) = \sum_b \mathbb{P}(b)(p_1 a + p_2 b) = p_1 a + p_2 \mathbb{E}b$$

$$\sum_b \mathbb{P}(b)\mathbb{E}^2(c|a,b) = \sum_b \mathbb{P}(b)(ap_1 + bp_2)^2 = p_1^2 a^2 + p_2^2 \mathbb{E}b^2 + 2p_1 p_2 a \mathbb{E}b$$

$$\Rightarrow D_1(c|a) = p_1(1-p_1)a + p_2(1-p_2)\mathbb{E}b + p_1^2 a^2 + p_2^2 \mathbb{E}b^2 + 2p_1p_2a\mathbb{E}b - p_1^2a^2 - p_2^2\mathbb{E}^2 b - 2p_1p_2a\mathbb{E}b =$$

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

$$D_2(c|a) = p_1 a + p_2\mathbb{E}b + p_2^2\mathbb{D}b$$

В итоге имеем:

$$\mathbb{E}\ (c|a) = p_1 a + p_2 \mathbb{E}b$$

$$\mathbb{D}_1\ (c|a) = p_1(1-p_1)a + p_2(1-p_2)\mathbb{E}b + p_2^2\mathbb{D}b$$

$$\mathbb{D}_2\ (c|a) = p_1 a + p_2\mathbb{E}b + p_2^2\mathbb{D}b$$

4 p(c|b)

Абсолютно аналогично $p(c|a)$ получим, что

$$p(c|b) = \sum_a p(c|a,b)p(a) \qquad \mathbb{P}(c=k|b) = \sum_a \mathbb{P}(c=k|a,b)\mathbb{P}(a)$$

$$\mathbb{E}\ (c|b) = p_1\mathbb{E}a + p_2 b$$

$$\mathbb{D}_1\ (c|b) = p_1(1-p_1)\mathbb{E}a + p_2(1-p_2)b + p_1^2\mathbb{D}a$$

$$\mathbb{D}_2\ (c|b) = p_1\mathbb{E}a + p_2 b + p_1^2\mathbb{D}a$$

5 p(d)

$$p(d,c) = p(d|c)p(c) \quad \Rightarrow \quad p(d) = \sum_c p(d,c) = \sum_c p(d|c)p(c) \quad \Rightarrow$$

$$\Rightarrow \quad \mathbb{P}(d=k) = \sum_c \mathbb{P}(d=k|c)\mathbb{P}(c)$$

Найдём мат ожидание $p(d)$:

$$\mathbb{E}\ d = \sum_k \mathbb{P}(d=k)\ k = \sum_k \sum_c \mathbb{P}(d=k|c)\mathbb{P}(c)k = \sum_c \mathbb{P}(c)\sum_k \mathbb{P}(d=k|c)k =$$

$$= \sum_c \mathbb{P}(c)\mathbb{E}(d|c) = \sum_c \mathbb{P}(c)(\mathbb{E}c + \mathbb{E}\ Bin(c,p_3)) = \sum_c \mathbb{P}(c)c + \sum_c \mathbb{P}(c)cp_3 =$$

$$\mathbb{E}c + p_3\sum_c \mathbb{P}(c)c = (1+p_3)\mathbb{E}c$$

Найдём мат ожидание d^2:

$$\mathbb{E}\ d^2 = \sum_k \mathbb{P}(d=k)\ k^2 = \sum_k \sum_c \mathbb{P}(d=k|c)\mathbb{P}(c)k^2 = \sum_c \mathbb{P}(c)\mathbb{E}(d^2|c) =$$

$$= \sum_c \mathbb{P}(c)(\mathbb{D}(d|c) + \mathbb{E}^2(d|c))$$

Так как в $\mathbb{D}(d|c)$ и $\mathbb{E}^2(d|c)$ c является константой, то:

$$\mathbb{E}^2(d|c) = (\mathbb{E}c + \mathbb{E}Bin(c,p_3))^2 = (c+cp_3)^2 = c^2(1+p_3)^2$$

$$\mathbb{D}(d|c) = \mathbb{D}(c + Bin(c,p_3)) = \mathbb{D}Bin(c,p_3) = cp_3(1-p_3)$$

$$\Rightarrow \quad \mathbb{E}\ d^2 = \sum_c \mathbb{P}(c)(c^2(1+p_3)^2 + cp_3(1-p_3)) = p_3(1-p_3)\mathbb{E}\ c + (1+p_3)^2\mathbb{E}\ c^2 =$$

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

Тогда получим, что:

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

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

В итоге имеем:

$$\mathbb{E}\ d = (1+p_3)\mathbb{E}c \quad \mathbb{D}\ d = p_3(1-p_3)\mathbb{E}\ c + (1+p_3)^2\mathbb{D}\ c$$

6 p(c|d)

Для вывода формулы распределения воспользуемся теоремой Байеса:

$$p(c|d) = \frac{p(d|c)p(c)}{\sum_c p(d|c)p(c)} = \frac{p(d|c)p(c)}{\sum_c p(c,d)} = \frac{p(d|c)p(c)}{p(d)}$$

$$\Rightarrow \quad \mathbb{P}(c=k|d) = \frac{\mathbb{P}(d|c=k)\mathbb{P}(c=k)}{\mathbb{P}(d)}$$

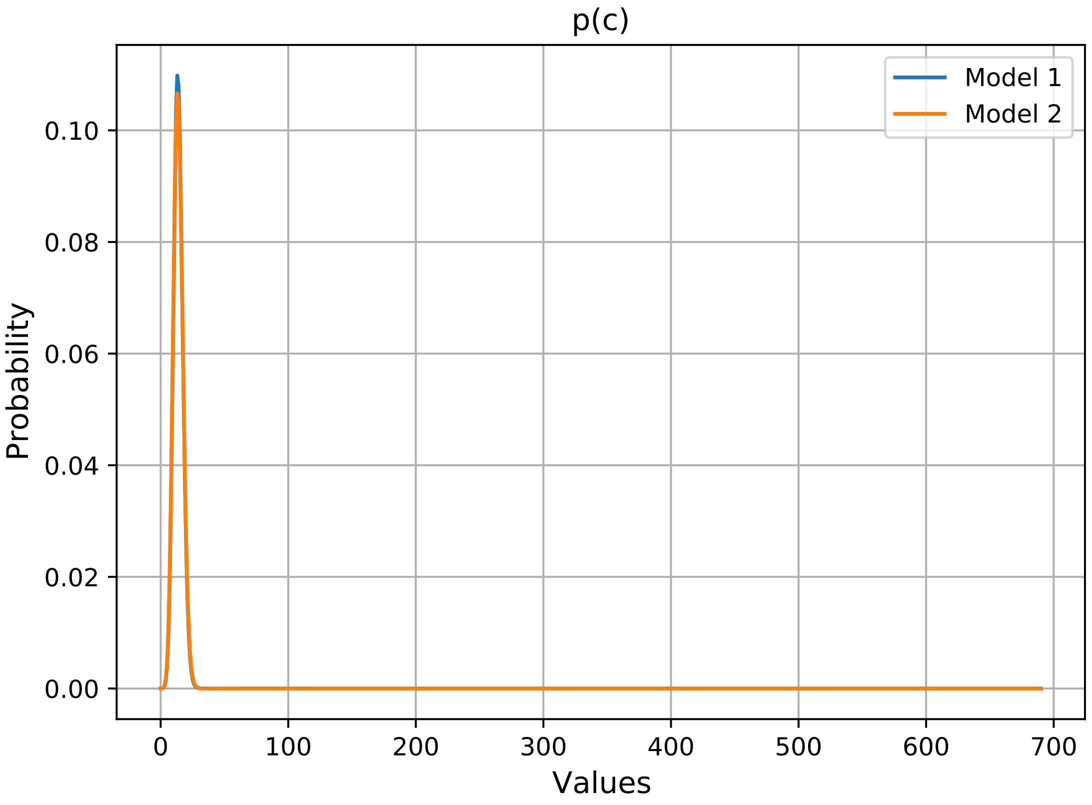
7 p(c|a, b, d)

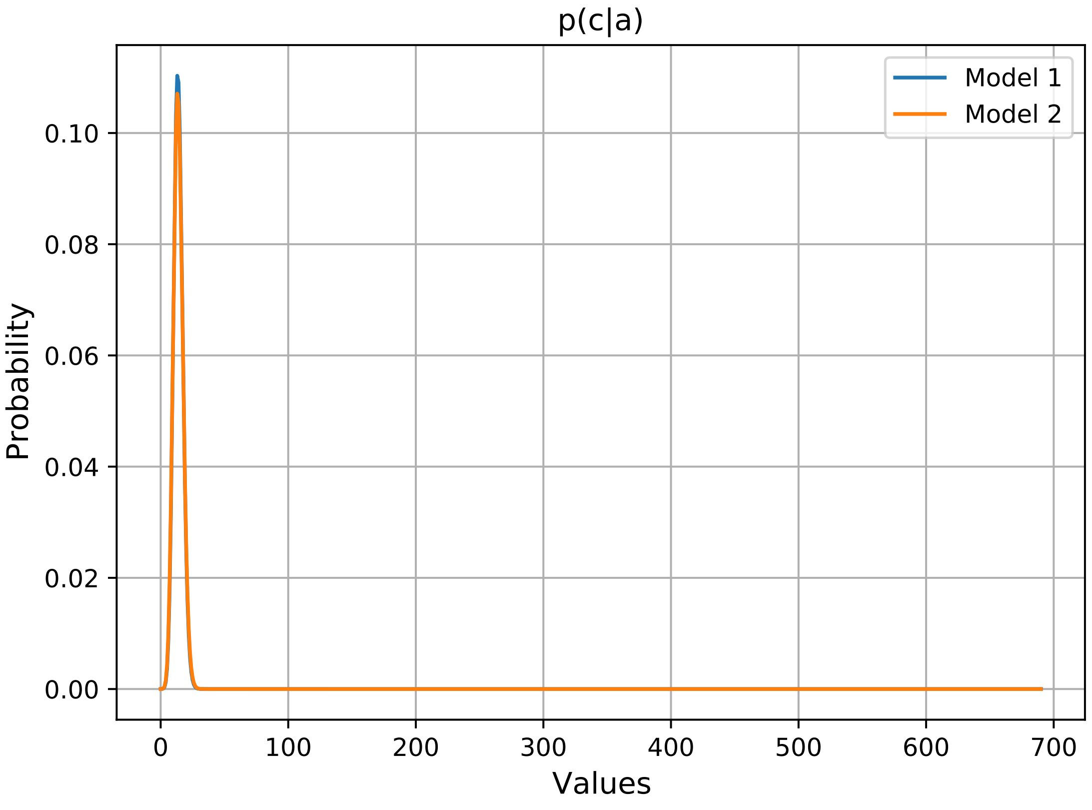
$$p(a,b,c,d) = p(c|a,b,d)p(a,b,d) \quad \Rightarrow \quad p(c|a,b,d) = \frac{p(a,b,c,d)}{p(a,b,d)} = \frac{p(a,b,c,d)}{\sum_c p(a,b,c,d)}$$

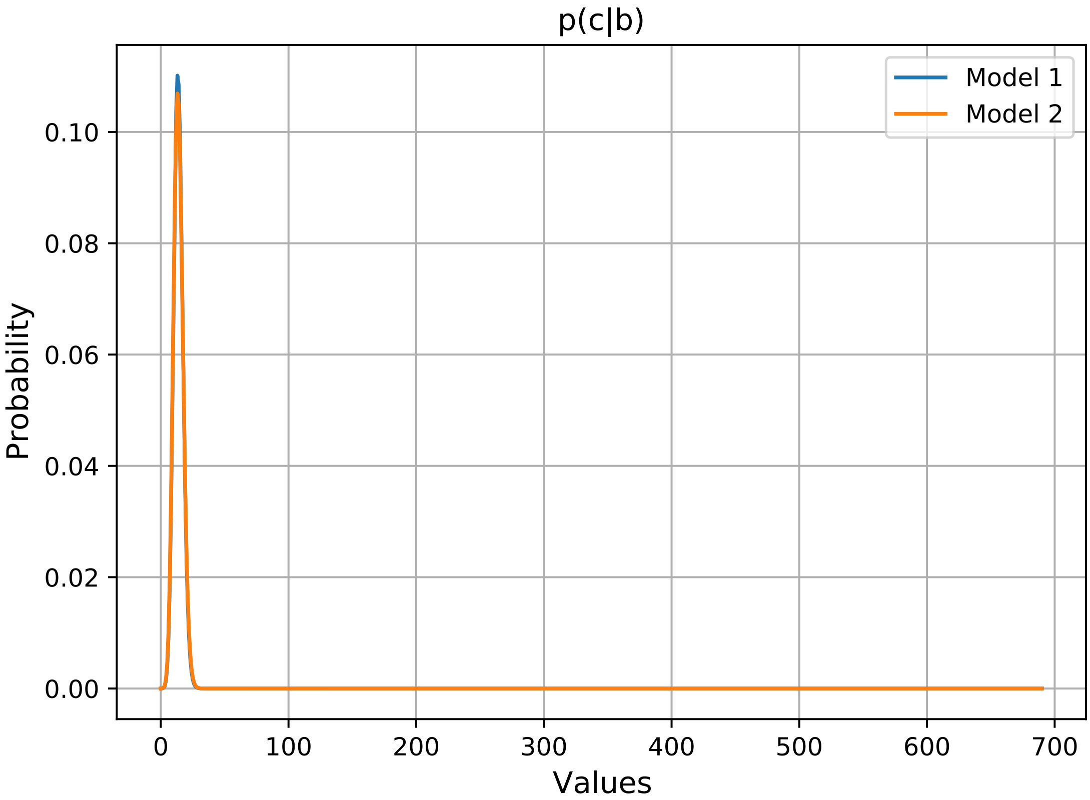
$$\Rightarrow \quad \mathbb{P}(c=k|a,b,d) = \frac{\mathbb{P}(a)\mathbb{P}(b)\mathbb{P}(c=k|a,b)\mathbb{P}(d|c=k)}{\sum_c \mathbb{P}(a)\mathbb{P}(b)\mathbb{P}(c=k|a,b)\mathbb{P}(d|c=k)}$$

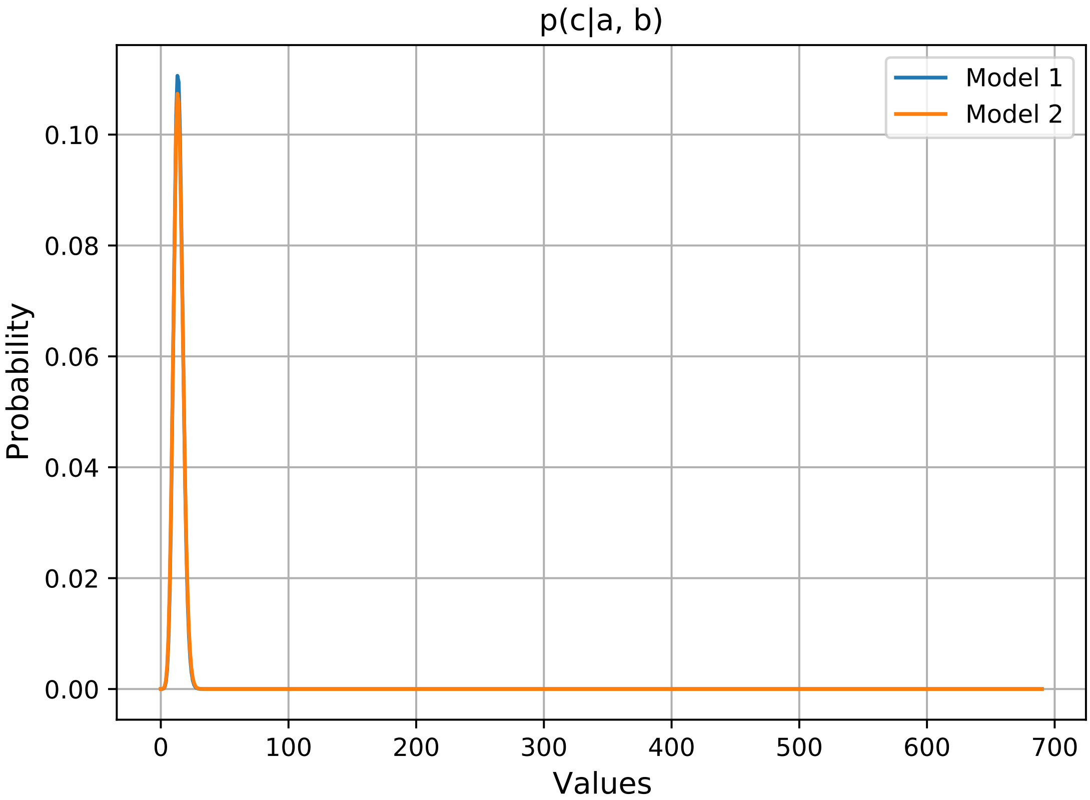
8 Графики распределений

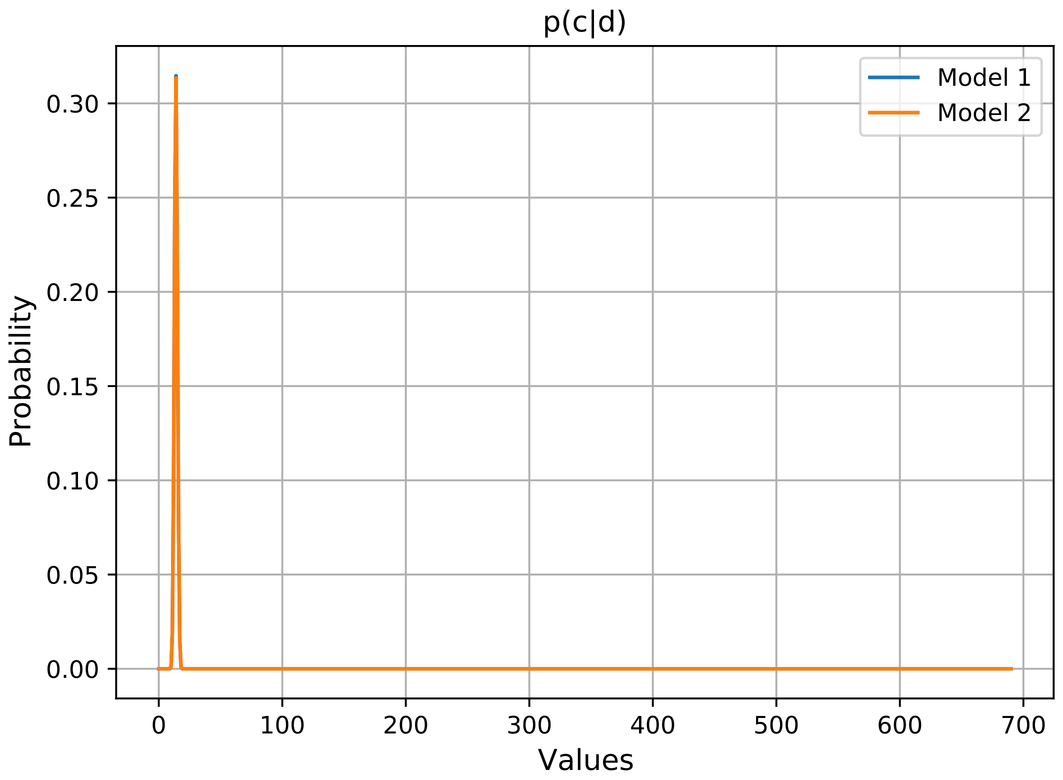
Мат ожидания для обеих моделей совпадают. Обозначим за $\mathbb{D}_1$ дисперсию для первой модели, а за $\mathbb{D}_2$ для второй. За t_i будем принимать время работы $i-$ой модели

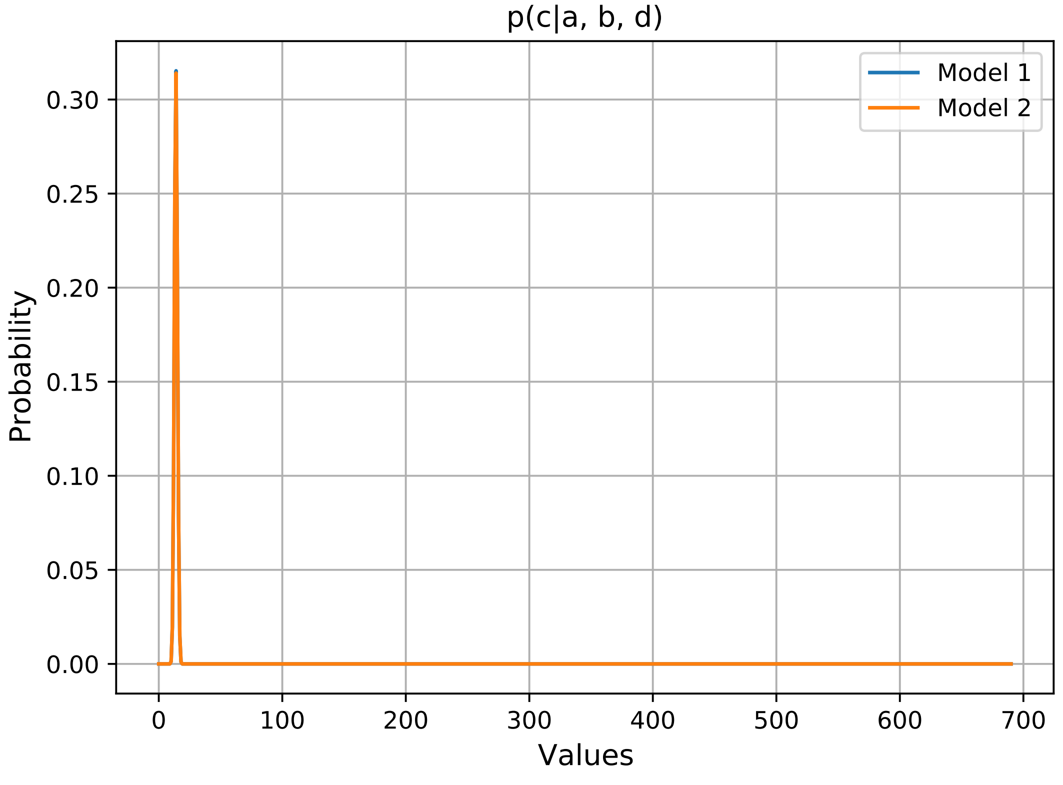

p(c)
Model 1
Model 2
Probability
0.10
0.08
0.06
0.04
0.02
0.00
0
100
200
300
400
500
600
700
Values


p(c|a)
Model 1
Model 2
Probability
0.10
0.08
0.06
0.04
0.02
0.00
0
100
200
300
400
500
600
700
Values


p(c|b)
Model 1
Model 2
Probability
0.10
0.08
0.06
0.04
0.02
0.00
0
100
200
300
400
500
600
700
Values


p(c|a, b)
Model 1
Model 2
Probability
0.10
0.08
0.06
0.04
0.02
0.00
0
100
200
300
400
500
600
700
Values





$\mathbb{E}\ c = 13.75 \quad \mathbb{D}_1\ c = 13.1675 \quad \mathbb{D}_2\ c = 14.0475 \quad t_1 = 208\ ms \quad t_2 = 8.82\ \mu s$

$\mathbb{E}\ (c|a = 83) = 13.8 \quad \mathbb{D}_1\ (c|a = 83) = 13.0 \quad \mathbb{D}_2\ (c|a = 83) = 13.885 \quad t_1 = 117\ ms \quad t_2 = 9.3\ \mu s$

$\mathbb{E}\ (c|b = 550) = 13.75 \quad \mathbb{D}_1\ (c|b = 550) = 13.0825 \quad \mathbb{D}_2\ (c|b = 550) = 13.9625 \ \ t_1 = 11.9\, ms \ \ t_2 = 4\, ms$

$\mathbb{E}\ (c|a = 83, b = 550) = 13.8 \quad \mathbb{D}_1\ (c|a = 83, b = 550) = 12.915 \quad \mathbb{D}_2\ (c|a = 83, b = 550) = 13.8 \ \ t_1 = 15.4\, ms \ \ t_2 = 2.03\, ms$

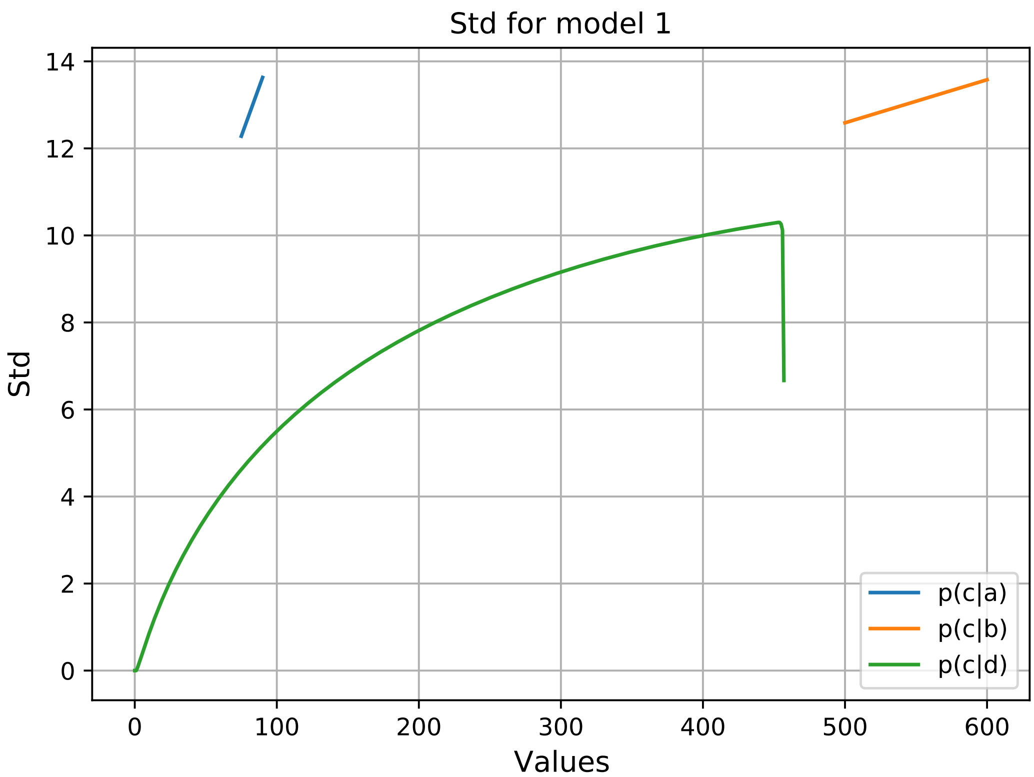
$\mathbb{E}\ (c|d = 18) = 13.9 \quad \mathbb{D}_1\ (c|d = 18) = 1.534 \quad \mathbb{D}_2\ (c|d = 18) = 1.544 \ \ t_1 = 1.23\, s \ \ t_2 = 685\, ms$

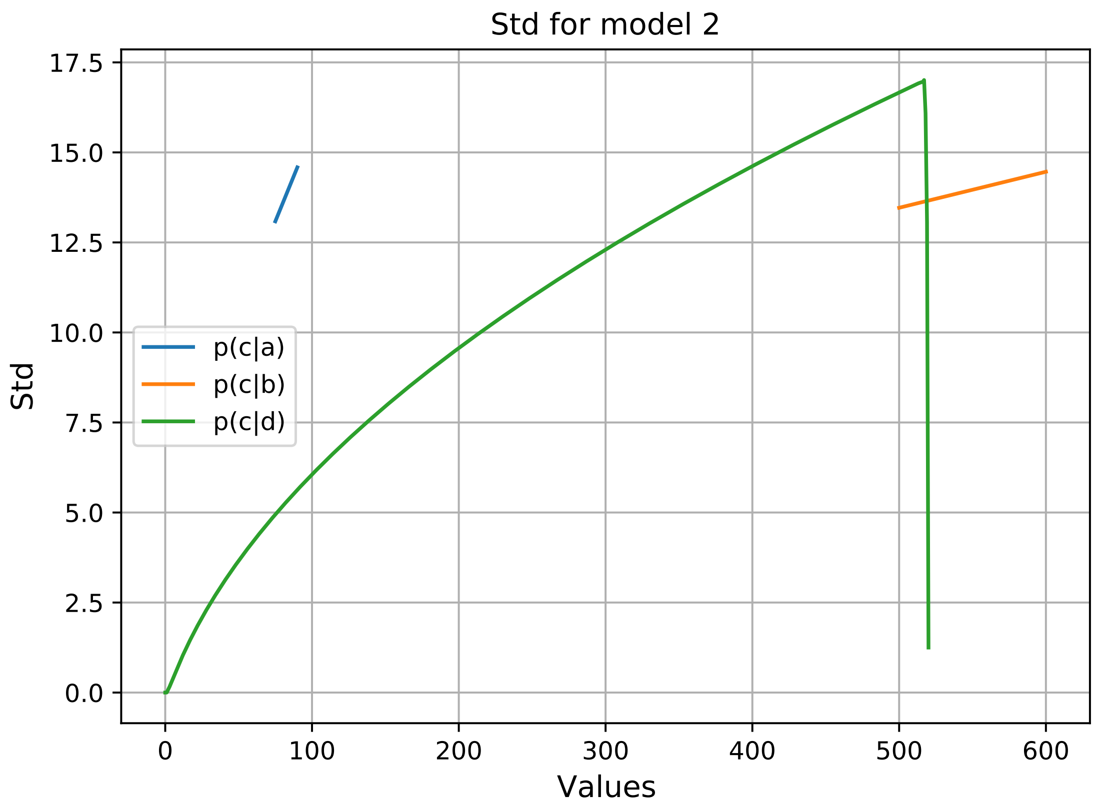
$\mathbb{E}\ (c|a = 83, b = 550, d = 18) = 13.9 \quad \mathbb{D}_1\ (c|a = 83, b = 550, d = 18) = 1.530 \ \ \mathbb{D}_2\ (c|a = 83, b = 550, d = 18) = 1.540 \ \ t_1 = 423\, ms \ \ t_2 = 253\, ms$

$\mathbb{E}\ d = 17.875 \quad \mathbb{D}_1\ d = 25.14 \quad \mathbb{D}_2\ d = 26.63 \ \ t_1 = 622\, ms \ \ t_2 = 333\, ms$

Как видно из графиков и из значений дисперсий лучше всего на прогноз c влияет d. Что неудивительно, так как количество присутствующих студентов намного сильнее коррелирует с отметившимися, чем с общим количеством записанных на курс. Заметим так же, что a вносит чуть больше ясности, чем b.

9 Уточнение прогноза





Заметим, что во второй модели при достаточно больших d $\mathbb{D}(c|d) > \mathbb{D}(c|b)$. Следовательно нельзя сказать, что d вносит наибольший вклад в уточнение для прогноза c. Напротив, в первой модели $\mathbb{D}(c|d)$ растёт с ростом d, однако потом начинает уменьшаться. Что вполне предсказуемо, так как становится меньше возможных вариантов для c. Поэтому в первой модели можно утверждать, что d играет большую роль в оценке c. (Для больших значений d не получается вычислить дисперсию из-за точности вычислений на компьютере)

Найдём множества $A = \{(a, b) \mid \mathbb{D}(c|b) < \mathbb{D}(c|a)\}$ и $B = \{(a, b) \mid \mathbb{D}(c|b) \geq \mathbb{D}(c|a)\}$

Обозначим за $p'_i = p_i$ или $p'_i = p_i(1 - p_i)$, $i \in \{1, 2\}$ в зависимости от модели.

Тогда: $\mathbb{D}\ (c|a) = p'_1 a + p'_2 \mathbb{E}b + p_2^2 \mathbb{D}b$, $\mathbb{D}\ (c|b) = p'_1 \mathbb{E}a + p'_2 b + p_1^2 \mathbb{D}a$

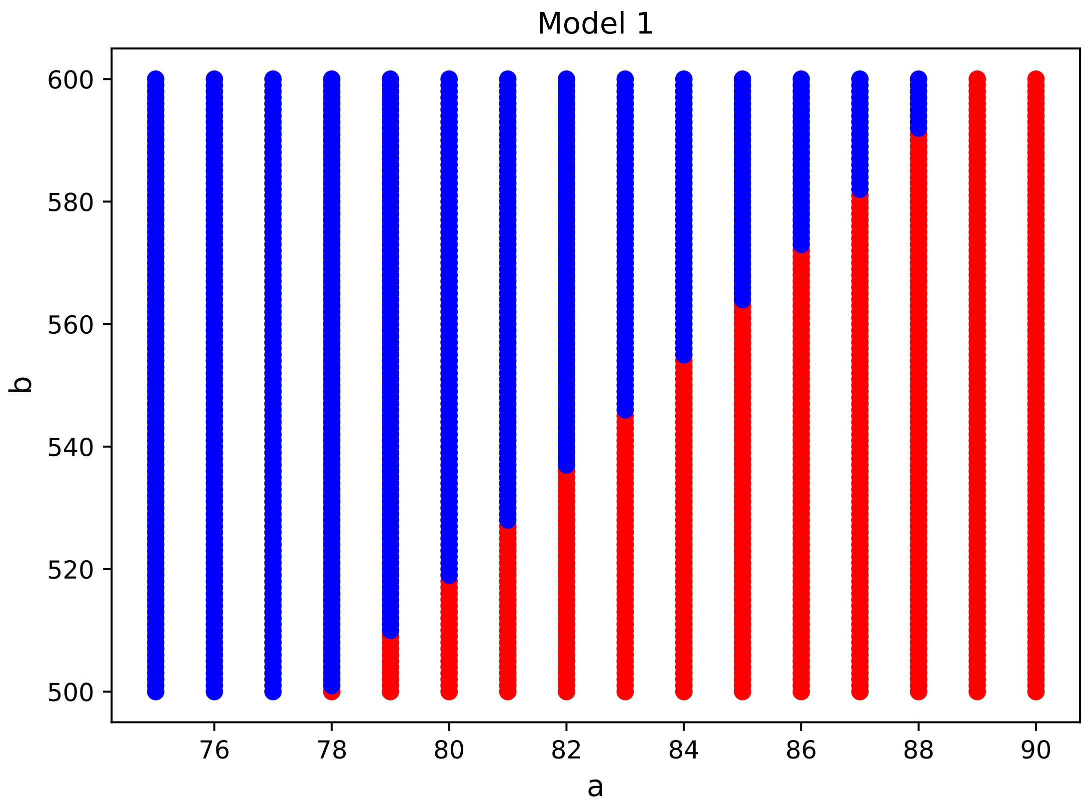
$\mathbb{D}(c|a) - \mathbb{D}(c|b) = p'_1(a - \mathbb{E}a) + p'_2(\mathbb{E}b - b) + p_2^2 \mathbb{D}b - p_1^2 \mathbb{D}a > 0$

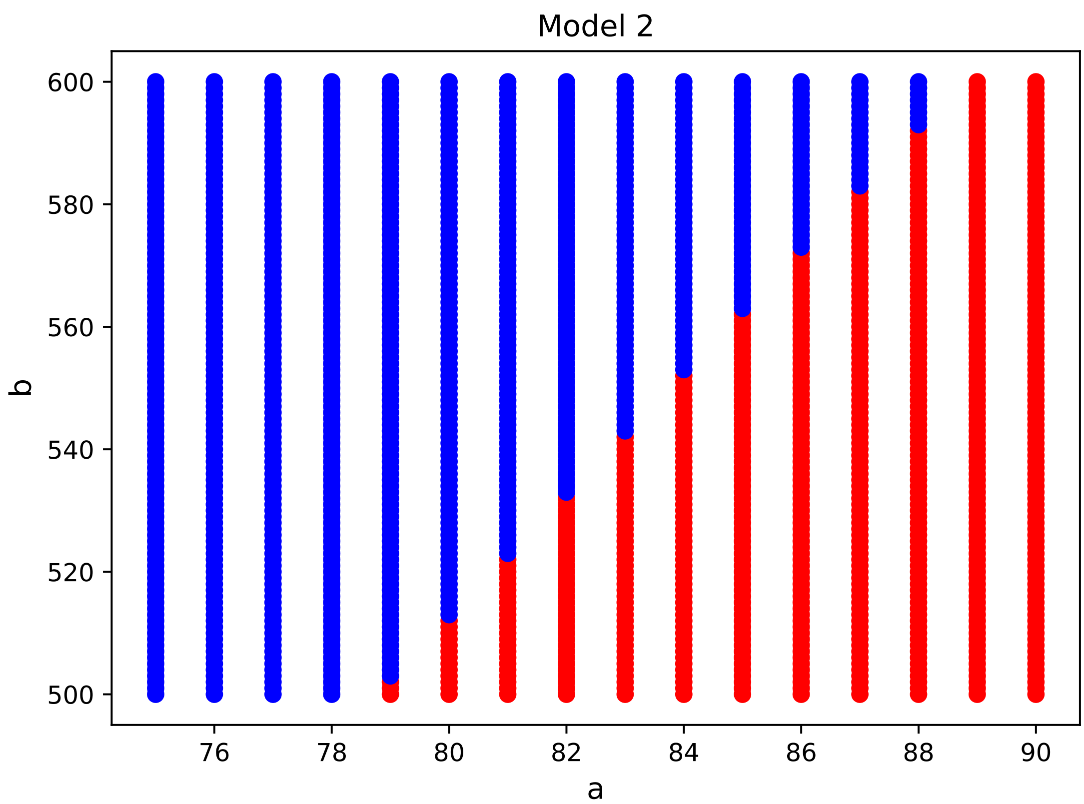
$\Rightarrow \quad p'_1 a - p'_2 b > p'_1 \mathbb{E}a - p'_2 \mathbb{E}b - p_2^2 \mathbb{D}b + p_1^2 \mathbb{D}a = const$

$\Rightarrow \quad p'_1 a - p'_2 b > const$

Получили линейную зависимость от a и b. Следовательно множества A и B линейно разделимы. И множество A состоит из пар удовлетворяющих уравнению $p'_1 a - p'_2 b > const$.

Изобразим на графике эти множества. Синими точками обозначены элементы B, а красными A. Как видно из рисунка данные множества линейно разделимы.


Model 1
600
580
560
540
520
500
b
76
78
80
82
84
86
88
90
a

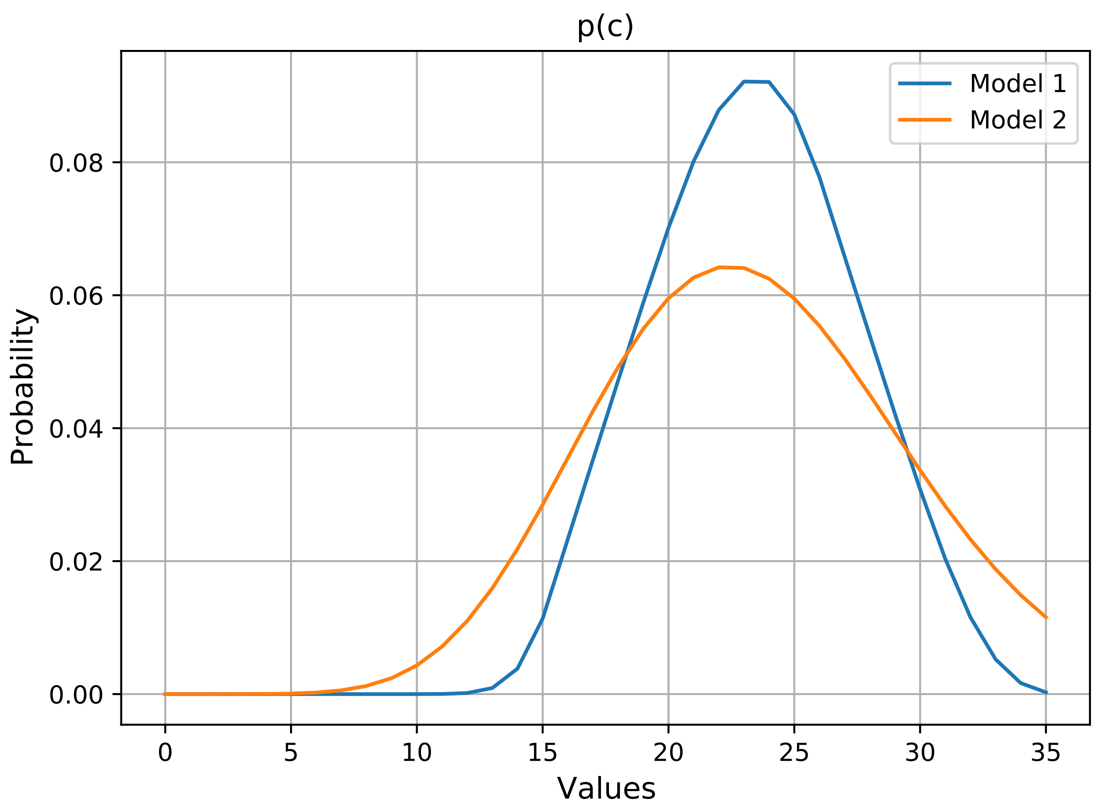

Model 2
600
580
560
540
520
500
b
76
78
80
82
84
86
88
90
a

10 Сравнение моделей

Как уже было получено в предыдущем разделе модели отличаются в уточнении прогноза c. Так же заметим, что вторая модель является лишь приближением первой. Известно, что биномиальное распределение $Bin(n, p)$ при большом количестве испытаний и маленькой вероятности успеха может быть с высокой точностью приближено пуассоновским распределением $Poiss(\lambda)$ с $\lambda = np$. Но если у нас мало испытаний и большая вероятность успеха, то тогда данные модели будут показывать совершенно разный результат.

Пусть на излюбленный курс по байесовским методам ходит только кафедра ММП, на которой из года в год 15-25 человек. А так как это не совсем простые студенты, то они посещают абсолютно все пары. Тогда вероятность их присутствия 0.95. Студенты других кафедр менее ответственные, поэтому их ходит до 10 человек с вероятностью 0.8 (среди них тоже есть матёрые).

Построим графики вероятностей распределения $p(c)$.



Как видно из рисунка, вероятности довольно сильно отличаются. Вторая модель менее точная, поэтому имеет более большую дисперсию.

P.s. автор не хотел никого обидеть, просто слегка утрировал реальность.