Московский государственный университет имени М. В. Ломоносова.

Байесовские рассуждения.

Чванкина Дарья
Группа 417
Октябрь 2022

1 Введение

Заданы 2 вероятностные модели количества студентов, посещаемых курс. Предлагается сравнить две данные модели с помощью экспериментов и привести теоретические выкладки.

2 Аналитический вывод формул

2.1 Задание № 2

$$\mathbb{E}A = \sum_{i=1}^{\infty} x_i p_i = \frac{1}{a_{max} - a_{min} + 1}(75 + 76 + \cdots + 90) \tag{1}$$

$$\mathbb{E}B = \sum_{i=1}^{\infty} x_i p_i = \frac{1}{b_{max} - b_{min} + 1}(75 + 76 + \cdots + 90) \tag{2}$$

Модель 1

$$\begin{aligned} \mathbb{E}C = \mathbb{E}(\mathbb{E}(C|A,B)) = \mathbb{E}(\mathbb{E}(Bin(A,p_1) + Bin(B,p_2))) = \mathbb{E}(\mathbb{E}Bin(A,p_1) + \mathbb{E}Bin(B,p_2)) = \\ = \mathbb{E}(Ap_1 + Bp_2) = p_1\mathbb{E}A + p_2\mathbb{E}B \end{aligned} \tag{3}$$

$$\begin{aligned} \mathbb{D}C = \mathbb{D}(\mathbb{E}(C|A,B)) + \mathbb{E}(\mathbb{D}(C|A,B)) = \{Cov(Bin(A,p_1), Bin(B,p_2)) = 0, Cov(A,B) = 0\} = \\ = \mathbb{D}(Ap_1 + Bp_2) + \mathbb{E}(Ap_1(1-p_1) + Bp_2(1-p_2)) = \\ = p_1^2\mathbb{D}A + p_2^2\mathbb{D}B + p_1(1-p_1)\mathbb{E}A + p_2(1-p_2)\mathbb{E}B \end{aligned} \tag{4}$$

Модель 2

$$\mathbb{E}C = \mathbb{E}(\mathbb{E}(C|A,B)) = \mathbb{E}(\mathbb{E}(Poiss(Ap_1 + Bp_2)|A,B)) = \mathbb{E}(Ap_1 + Bp_2) = p_1\mathbb{E}A + p_2\mathbb{E}B \tag{5}$$

$$\begin{aligned} \mathbb{D}C = \mathbb{D}(\mathbb{E}(C|A,B)) + \mathbb{E}(\mathbb{D}(C|A,B) = \mathbb{D}(Ap_1 + Bp_2) + \mathbb{E}(Ap_1 + Bp_2) = \\ = p_1^2\mathbb{D}A + p_2^2\mathbb{D}B + p_1\mathbb{E}A + p_2\mathbb{E}B \end{aligned} \tag{6}$$

Верно для обоих моделей

$$\mathbb{E}D = \mathbb{E}(\mathbb{E}(D|C)) = \mathbb{E}(\mathbb{E}(C + Bin(C,p_3)|C)) = \mathbb{E}(\mathbb{E}C + Cp_3) = \mathbb{E}C + p_3\mathbb{E}C = (p_3+1)\mathbb{E}C \tag{7}$$

$$\begin{aligned} \mathbb{D}D = \mathbb{D}(\mathbb{E}(D|C)) + \mathbb{E}(\mathbb{D}(D|C)) = \{Cov(const,..) = 0\} = \mathbb{D}(\mathbb{E}C + p_3C) + \mathbb{E}(\mathbb{D}C + Cp_3(1-p_3)) = \\ = p_3^2\mathbb{D}C + \mathbb{D}C + p_3(1-p_3)\mathbb{E}C \end{aligned} \tag{8}$$

2.2 Для остальных заданий

$$\mathbb{P}(A = n) = \frac{1}{a_{max} - a_{min} + 1}$$
$$\mathbb{P}(B = n) = \frac{1}{b_{max} - b_{min} + 1} \tag{9}$$

$$\mathbb{P}(C = n) = \sum_{a' \in [a_{min}, a_{max}]} \sum_{b' \in [b_{min}, bmax]} \mathbb{P}(C = n | a = a', b = b')\mathbb{P}(a = a')\mathbb{P}(b = b')$$
для модели 1:
$$\mathbb{P}(C = n | a = a', b = b') = \sum_{k=1}^{n} \mathbb{P}(Bin(a', p_1) = k)\mathbb{P}(Bin(b', p_2) = n - k) \tag{10}$$
для модели 2:
$$\mathbb{P}(C = n | a = a', b = b') = \mathbb{P}(Poisson(a'p_1 + bp_2) = n)$$

$$\mathbb{P}(C = n | a = a') = \sum_{b' \in [b_{min}, bmax]} \mathbb{P}(C = n | a = a', b = b')\mathbb{P}(b = b') \tag{11}$$

$$\mathbb{P}(C = n | b = b') = \sum_{a' \in [a_{min}, a_{max}]} \mathbb{P}(C = n | a = a', b = b')\mathbb{P}(a = a') \tag{12}$$

$$\mathbb{P}(D = n) = \sum_{c' \in [c_{min}, c_{max}]} \mathbb{P}(D = n | c = c')\mathbb{P}(c = c')$$
$$\mathbb{P}(D = n | c = c') = \mathbb{P}(Bin(c', p_3) = n - c') \tag{13}$$

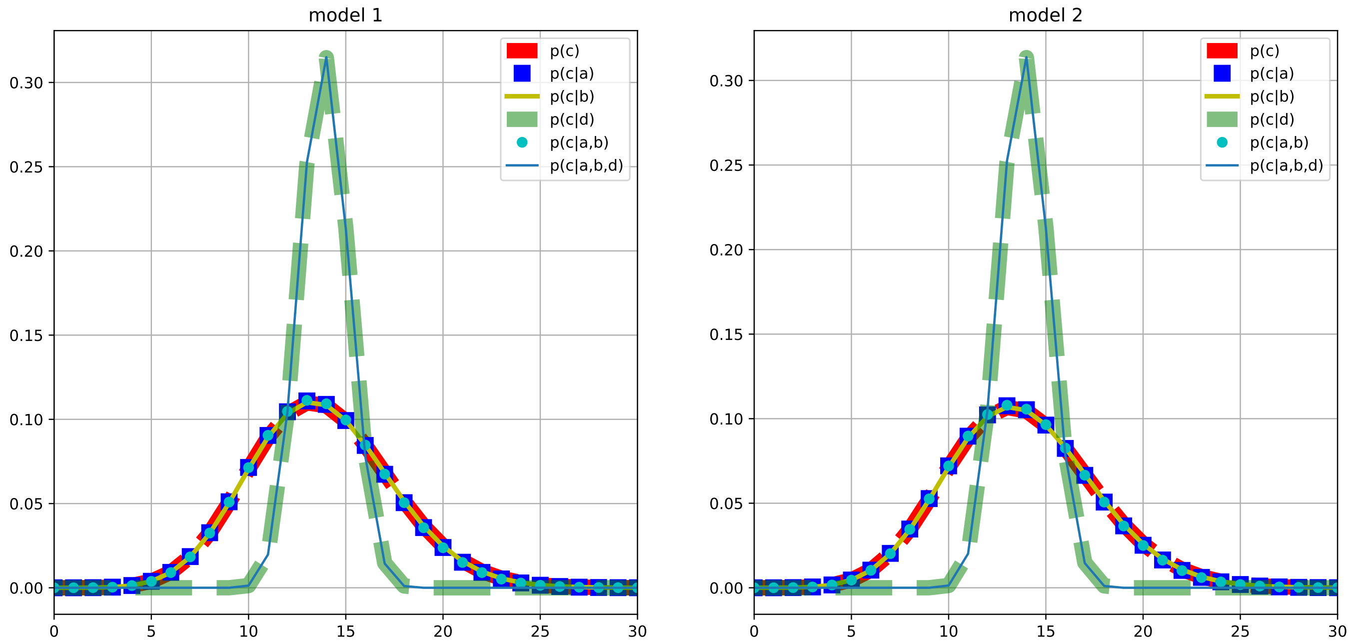
По Теореме Байеса:

$$p(c|d) = \frac{p(d|c)p(c)}{p(d)} \tag{14}$$

По формуле полной вероятности:

$$p(c|a, b, d) = \frac{p(a, b, c, d)}{p(a, b, d)}$$
$$p(a, b, c, d) = p(d|c)p(c|a, b)p(a)p(b) \tag{15}$$
$$p(a, b, d) = \sum_{c' \in [c_{min}, c_{max}]} p(a, b, c', d)$$

Рис. 1:



Модель 1		
Распределение	Мат. ожидание	Дисперсия
p(c)	13.75	13.17
p(c|a)	13.7	12.91
p(c|b)	13.75	13.08
p(c|d)	13.9	1.53
p(c|a,b)	13.7	12.83
p(c|a, b, d)	13.89	1.53

Модель 2		
Распределение	Мат. ожидание	Дисперсия
p(c)	13.75	14.05
p(c|a)	13.7	13.79
p(c|b)	13.75	13.96
p(c|d)	13.89	1.54
p(c|a,b)	13.7	13.7
p(c|a, b, d)	13.89	1.54

Таблица 1:

3 Эксперименты

Пронаблюдаем как происходит уточнение прогноза для величины c по мере прихода новой косвенной информации. Для этого построим графики для распределений p(c), p(c|a), p(c|b), p(c|d), p(c|a, b), p(c|a, b, d) при параметрах a, b, d, равных мат. ожиданиям своих априорных распределений, округленных до ближайшего целого (см. Рис. 1).

Мы видим, что параметры a, b не вносят никакую информацию о распределении. Совсем другой результат показывает параметр d. Условные распределения с ним более уверены в своем прогнозе, так как купол распределения стал более узким и более высоким.

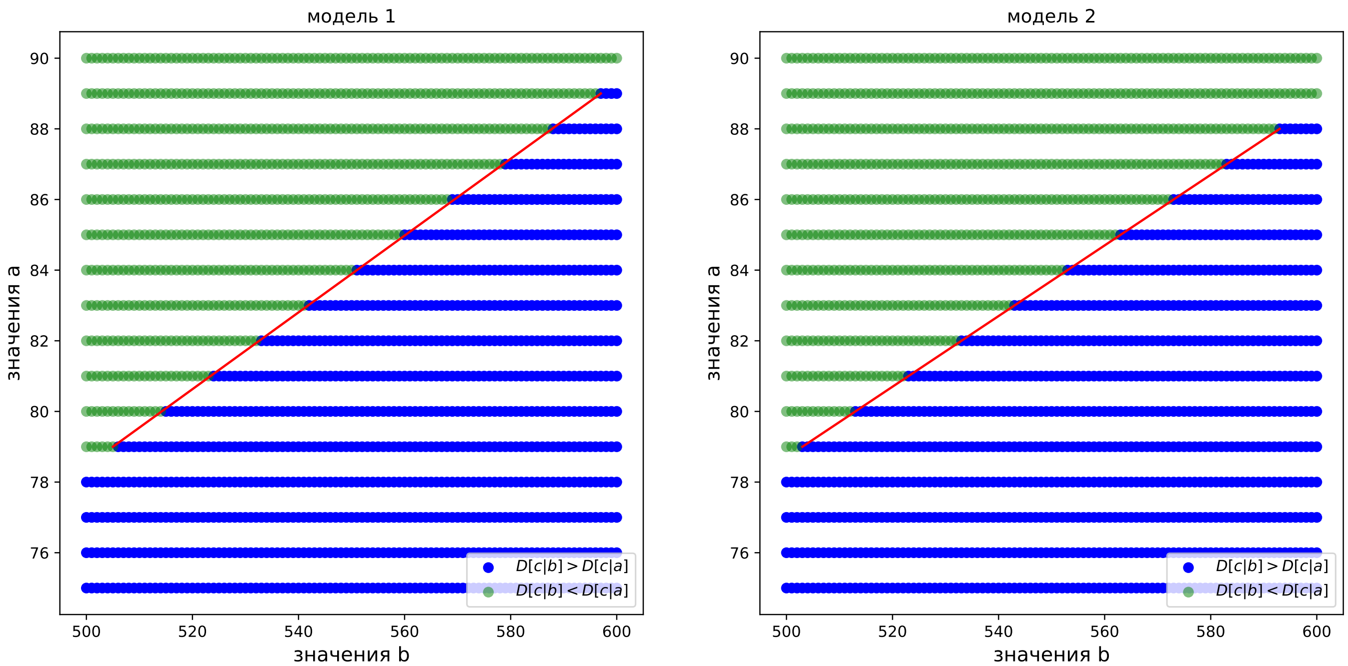
Посмотрим на то, как изменяются мат. ожидания и дисперсии, посчитанные численно.(См. Таб. 1) Мат. ожидания изменяются, но не намного. Зато заметно, что дисперсия при информации из параметра a меньше, чем при параметре b. Также заметно, что примерно в 8 раз уменьшается дисперсия, когда приходит информация из d.

Также было проведено сравнение дисперсий $\mathbb{D}[c|a]$ и $\mathbb{D}[c|d]$, и $\mathbb{D}[c|b]$ и $\mathbb{D}[c|d]$ при всех значениях параметров a, b, d. Эксперимент показал, что выполняются неравенства: $\mathbb{D}[c|a] <$

$\mathbb{D}[c|d]$ и $\mathbb{D}[c|b] < \mathbb{D}[c|d]$ для модели 1. В модели два этого не наблюдается.

Сравним как вносят свой вклад параметры a и b. Построим графики и посмотрим являются ли линейно разделимыми множества $\{(a,b)|\mathbb{D}[c|b] < \mathbb{D}[c|a]\}$ и $\{(a,b)|\mathbb{D}[c|b] > \mathbb{D}[c|a]\}$. Исходя из графиков на Рис.2 заметим, что множества линейно разделимые. Также можно заметить, что чем больше a, тем при большем количестве значений b условная дисперсия $\mathbb{D}(c|a)$ будет больше, чем $\mathbb{D}(c|b)$.

Рис. 2:



Теперь посмотрим на время выполнения оценки распределений.(См. Таб. 2) Дольше всех считаются оценки для распределения p(c|d). Время вычисления мат. ожидания и дисперсии существенно отличается от остальных распределений. Так как в реализации вычисления данного распределения используются вычисления других распределений, то это неудивительно.

Модель 1		
Время, сек	Мат. ожидание	Дисперсия
p(c)	0.14	0.28
p(c|a)	2.21	4.6
p(c|b)	14.28	34.17
p(c|d)	645.33	1248.61
p(c|a,b)	2.19	4.44
p(c|a, b, d)	7.57	14.48

Модель 2		
Время, сек	Мат. ожидание	Дисперсия
p(c)	0.33	0.33
p(c|a)	1.47	2.48
p(c|b)	6.77	13.47
p(c|d)	435.15	859.09
p(c|a,b)	1.1	2.18
p(c|a, b, d)	4.95	9.91

Таблица 2:

Сравним модели между собой на проведенных экспериментах. Визуально графики на Рис. 1 не отличаются друг от друга и модели ведут себя похоже. Если посмотреть на Таб. 1, то заметим, что мат. ожидания у распределений не отличаются, чего нельзя сказать о дисперсиях. Дисперсии у второй модели больше, чем у первой, следовательно, вторая модель менее уверена в прогнозе.

Посмотрим на время вычисления оценок распределений двух моделей.(См. Таб. 2) В данном случае представляет интерес только распределение p(c|d), так как оно считается дольше всех. Тут однозначно выигрывает модель номер два.

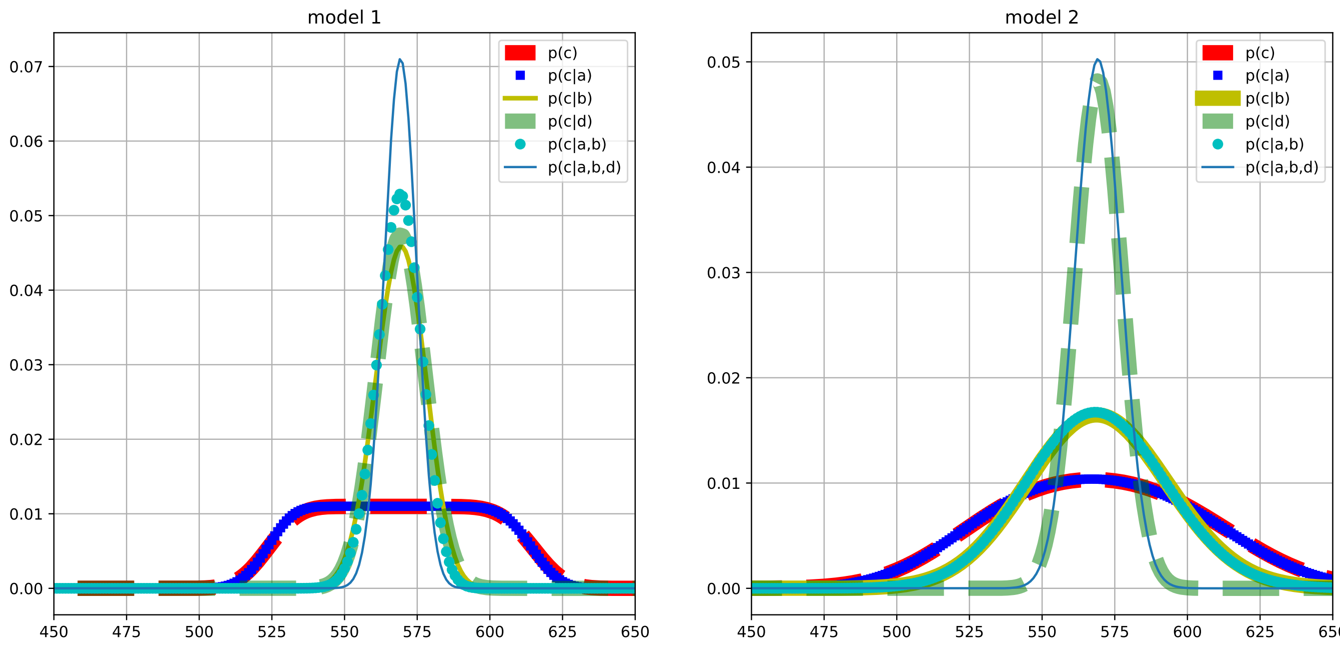
Поменяем немного параметры на $a_{min} = 75, a_{max} = 90, b_{min} = 500, b_{max} = 600, p_1 = 0.9, p_2 = 0.9, p_3 = 0.3$. Построим графики распределений p(c), p(c|a), p(c|b), p(c|d), p(c|a, b), p(c|a, b, d).(См. Рис. 3) Заметим, что модели ведут себя по-разному. Распределение p(c) у второй модели выглядит как колокол, а у первой модели есть отрезок, на котором p(c) ведет себя как равномерное распределение.

У двух моделей по-прежнему параметр a не дает информации. При это в первой модели параметр b дал много информации - график стал более узким и высоким, а в модели два мы видим не такой сильный эффект от параметра b.

Заметим, что в первой модели параметр d дал столько же информации, как и параметр b, а вот во второй модели наибольшую информацию дает только параметр d. В первой модели параметры a, b и d в совокупности дают информацию о распределении p(c) сопоставимую с информацией о c при параметре d в модели 2.

Различия в моделях объясняется в различии формул дисперсий. Если мат. ожидание у распределения p(c) одинаковое, то в формуле дисперсий модели разные. Они отличаются в коэффициентах при $\mathbb{E}A$ и $\mathbb{E}B$.

Рис. 3:



4 Вывод

Модель 1 и 2 при параметрах данных в задании различаются не сильно. Выгоднее использовать модель два, так как оценка распределения p(c|d) работает быстрее, чем в первой модели, а изменения в мат. ожиданиях и дисперсиях не существенны. Также вторая модель алгоритмически проще. Но при других параметрах данные модели могут существенно различаться, что продемонстрировано на Рис. 3.