Тема # 4

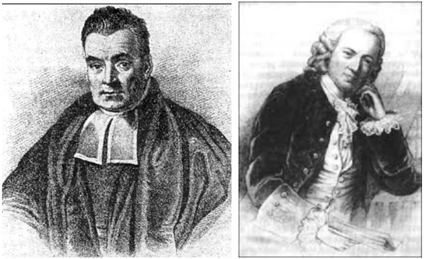
ФОРМУЛА ПОЛНОЙ ВЕРОЯТНОСТИ. ФОРМУЛА БАЙЕСА И СХЕМА ИСПЫТАНИЙ БЕРНУЛЛИ

Большая часть фото используемых в презентации взята из сцен к/ф "Достать ножи" , 2019 г.

Задания для самостоятельного анализа

ВНИМАНИЕ!!!

Сюжетная линия детективного характера, которая представлена в презентации, не имеет ничего общего с реальностью. Сюжет полностью выдуман, как и непосредственная связь между его персонажами. Несмотря на то что в материалах задействованы имена великих ученых (Байес и Бернулли), следует уточнить, что в действительности Якоб Бернулли (1655-1705 г.) и Томас Байес (1702-1761 г.) не были знакомы, а информация, представленная о них, крайне недостоверна. Исключение составляют лишь их заключения в контексте теории вероятностей.



: слева: Томас Байес, справа: Якоб Бернулли

Для того, чтобы не вводить читателя в заблуждение и оставить право выбора прочтения (только научный текст или же в контексте детективной истории) вымышленные сюжетные линии далее будут представлены синим шрифтом, а научный текст – черным.

Представим себе классический сценарий детективной истории!

Загородный особняк. На торжественный ужин собралась весьма обеспеченная семья Байеса, он же владелец особняка, который решил объявить о своем решении относительно наследства. Почти никто из гостей не скрывает, что его визит обусловлен интересом именно к этому объявлению. И каждый считает, что если кто и заслужил получить всё, то, конечно же, он и никто другой! Ужин проходил за весьма посредственными разговорами и притворным интересом гостей друг к другу. Разумеется, никто никого не слушал, а улыбки на лицах были настолько искусственными, что очевидным было то, что все стремятся не более чем хоть как-то заполнить время до долгожданной кульминации. Ну и, конечно же, произвести благочестивое впечатление на Байеса, покуда он, наконец, готов назвать имя приемника. Для каждого за столом время тянулось предательски медленно. Кто-то даже умудрился зевнуть. А кто-то, создавая видимость абсолютно сосредоточенного внимания к истории соседа справа, неожиданно для собеседника задремал и уронил со стола вилку! Звон упавшей вилки следом за собой породил абсолютную тишину ведь проявление такого неуважения явно не сулит ничем хорошим. Дабы в этом убедиться и насладиться осуждающим взором Байеса, он был в этом хорош, все повернулись к нему. Но ко всеобщему изумлению взгляд Байеса не ничуть не изменился, ибо к этому моменту он уже почил, не вставая со своего места.

Оживившись толи от смерти Байеса, толи от окончания нудного ужина, все бросились обвинять друг друга в убийстве. Каждый утверждал, что именно его имя указано в завещании, потому что он самый благородный и честный человек, который никогда бы не осмелился на столь страшное деяние, тем более в отношении любимого родственника!

Толи по счастливой случайности, толи по задумке Байеса, в доме присутствовал его адвокат, который был готов пролить свет истины на то, кто же унаследует всё! Вся семья затихла в ожидании. Однако все оказалось не так просто, как многие ожидали, ибо почивший завещал всё тому, кто первым верно установит имя преступника, которое должно совпасть с именем, что в конце вечера будет названо хорошим другом Байеса и по совместительству – бывшим одногруппником, который всю последнюю неделю гостит в его особняке!

Услышав это, Бернулли опешил! Ведь он не знал имени убийцы, но при этом должен был в конце вечера его назвать. Бернулли ничего не оставалось, как выйти к "опечаленной" смертью Байеса семье и отказаться от столько серьезных обязанностей, которые Байес завещал ему!

– Если я не ошибаюсь, то Вы также как и отец являетесь математиком! Ну так посчитайте чего там надо и назовите имя. Если имя будет "правильным" , то может и вам на старости лет больше ни о чем думать не придется! - произнесла дочь Байеса томно отводя взгляд при откровенной улыбке куда-то к потолку.

Бернулли оставил всех в холле разбираться между собой, в то время как сам решил удалиться в кабинет Байеса, чтобы изучить подсказки, которые Байес мог установить. Не даром же он был преисполнен уверенностью в том, что Бернулли разгадает имя убийцы.

– Я ухожу, на время! Но будьте готовы к тому, что я захочу с каждым из вас побеседовать лично...

В кабинете Байеса, Бернулли нашел дневник, в котором каждый член семьи имел свое обозначение: $H_1, H_2, \ldots, H_n$, а также было указано, что убийца один из них.

Бернулли стал рассматривать совокупность указанных множеств как события, которые можно было бы трактовать как: H_i – убийца это $i \in \{1, 2, \ldots, n\}$, причем

1. $\forall\, i, j \in \{1, 2, \ldots, n\} : j \neq i \to H_i \cdot H_j = \varnothing$, то есть события несовместны (убийца действовал один. Неудивительно, раз они все друг друга ненавидят! – подумал с ухмылкой Бернулли);
2. поскольку хотя бы один из подозреваемых сегодня состоялся как убийца, то $H_1 \cup H_2 \cup \ldots \cup H_n = \Omega$, то есть события (члены семьи) образуют полную группу событий.

Далее Бернулли вспомнил определение:

Определение 1

События $H_1, H_2, \ldots, H_n$ удовлетворяющие условиям 1-2 называются <u>гипотезами</u>

Поэтому всё, что было понятно с первой страницы дневника так это то, что Байес заранее рассматривал свое гипотетическое убийство каждым членом семьи.

Итого, что же нам известно? - подумал Бернулли почесывая затылок...:

Очевидно, что мотив был у каждого, тогда пусть $P(H_i) > 0,\ i = \overline{1, n}$.

Судя по посиневшим губам Байеса его вполне могли отравить фентанилом... ??! Перелистывая страницу дневника, Бернулли находит запись:

"A - передозировка фентанилом".

ДА КАКОГО ЛЕШЕГО?! - кричит Бернулли понимая, что Байес даже догадывался о том, чем именно его убьют. И на этот случай указал также и условные вроятности:

$$P(A|H_1), P(A|H_2), \ldots, P(A|H_n)$$

Но пока вопросов больше чем ответов и потому, было бы неплохо определиться с инструментом убийства – установить, действительно ли это фентанил?

Допустим, что это фентанил. Так как $A = A\Omega = A(H_1 + \ldots + H_n) = AH_1 + \ldots + AH_n$ тогда

$$P(A) = P\left(A\sum_{i=1}^{n} H_i\right) = P(AH_1) + \ldots + P(AH_n) = \sum_{i=1}^{n} P(AH_i),$$

но согласно формуле умножения имеем, что

$$\sum_{i=1}^{n} P(AH_i) = P(H_1)P(A|H_1) + \ldots + P(H_n)P(A|H_n) = \sum_{i=1}^{n} P(H_i)P(A|H_i).$$

Получается, что для некоторого события A и гипотез $\{H_i\}_{i=1}^{n}$ при известных $P(H_1) > 0, \ldots, P(H_n) > 0$ и $P(A|H_1), \ldots, P(A|H_n)$ можно заключить формулу полной вероятности события:

$$P(A) = \sum_{i=1}^{n} P(H_i)P(A|H_i).$$

Формула полной вероятности была получена задолго до них, так что Бернулли оказался далеко не первым кто до нее дошел...

"Если б я не пропускал пары в универе как Байес, то не пришлось бы формулу самому выводить... Зато, и с другой стороны, с синими губами на коврике сейчас лежу не Я!" – подискутировал сам с собой и успокоил немного себя Бернулли.

Переворачивая страницу дневника Бернулли находит рисунок Байеса. Однако рисунок по подсчету вероятности смерти от фентанила хоть и понятен, но лишен вычислений. Видимо Бернулли придется произвести их самостоятельно.

и здесь смотрим пример на доске!!!

Из примера видно, что Байес уже отмел варианты того, что некоторые родственники могли быть убийцами, так как они не имели доступа к фентанилу. Но списывать их со счетом все еще рано!

Итак, соберем все знания в кучу!

Пусть A может произойти с одним из событий $H_1, \ldots, H_n$, образующих полную группу попарно несовместных событий $\{H_i\}_{i=1}^{n}$, с известными вероятностями $P(H_i) > 0, i = \overline{1, n}$. Как и ранее нам известны $P(A|H_i)$ и то, что A произошло.

И что же дальше?! Какова условная вероятность самих гипотез?

Бернулли находит весьма содержательное заключение Байеса, которое тот приводит с полным доказательством, предоставляя тем самым любому желающему отменный инструмент для анализа случайных событий.

Теорема 1 (формула Байеса)

Пусть $P(A) > 0$ и для $\{H_i\}_{i=1}^n$ имеем $P(H_i) > 0$, $i = \overline{1, n}$, также известны $P(A|H_1), \ldots, P(A|H_n)$. Тогда условная вероятность гипотезы H_i определяется как

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{P(H_1)P(A|H_1) + \ldots + P(H_n)P(A|H_n)}.$$

Доказательство. ► Так как

$$P(A|H_i) = \frac{P(AH_i)}{P(H_i)} \Rightarrow P(AH_i) = P(H_i)P(A|H_i),$$

$$P(H_i|A) = \frac{P(H_iA)}{P(A)} \Rightarrow P(AH_i) = P(A)P(H_i|A),$$

тогда

$$P(H_i)P(A|H_i) = P(A)P(H_i|A) \Rightarrow P(H_i|A) = \frac{P(H_i)P(A|H_i)}{P(A)},$$

а далее по формуле полной вероятности. ◄ P.S.: я ж пары не пропускал! LOL

Простенько, незамысловато – тихо произнес Бернулли после прочтения доказательства и нахмурил брови. Его мимика выдавала не столь напряжение от насмешки мертвеца, сколько то, что на его памяти был случай, когда он бессознательно использовал доказанное заключение Байеса. Тот случай Бернулли и пытался вспомнить.

ДА-Да! Года три назад мне довелось принять участие в суде присяжных. Что-то там, эм, ..., речь шла о двух сыновьях, вроде, которые порешали своего отца зачем-то.... Как же их звали...? Эм, пусть 1 и 2 – цифры порой приятнее людей и никогда не врут! Улики в совокупности своей склоняли присяжных ко мнению о том, что виновен 1 на 40% и 60% - к тому, что виновен 2. Однако я отдал свой голос учитываю в отличии от остальных также агрессию и склонность каждого из них к насилию, которая была выражена у 1 на 90%, а у 2 - на 20%. С учетом этого мне было очевидно, что вероятность того, что виновен 1 вовсе не 0,4, а 0,75, в то время как у 2 - эт 0,25. С чем Байес б согласился!!!

ЗАДАЧА для разминки на применение формулы Байеса:

Врач после осмотра больного считает, что возможно одно из двух заболеваний, которые мы зашифруем номерами 1 и 2, причем степень своей уверенности в отношении правильности диагноза он оцениванет как 40% и 60% соответственно. Для уточнения диагноза больного направляют на анализ, исход которого дает положительную реакцию при заболевании 1 в 90% случаев и при заболевании 2 - в 20% случаев. Анализ дал положительную реакцию. Как изменится мнение врача после этого?

После припоминания того случая в суде, Бернулли решил собраться и еще раз посмотреть на все со стороны:

Вероятности $P(H_1), \ldots, P(H_n)$, которые оставил Байес были им получены ДО его смерти / опыта, то есть являются априорными.

В дневнике были указаны данные по тем членам семьи, которые имели доступ к фентанилу: пусть $P(H_1) = 0.25$, $P(H_2) = 0.4$, $P(H_3) = 0.15$, $P(H_4) = 0.2$; а также условные вероятности –
$P(A|H_1) = 0.5$, $P(A|H_2) = 0.3$, $P(A|H_3) = 0.6$, $P(A|H_4) = 0.4$.

Но после смерти / опыта Байеса можно рассмотреть апостериорные условные вероятности $P(H_1|A), \ldots, P(H_n|A)$ (т. е. полученные после опыта!), чтобы скорректировать подозрения Байеса с учетом новой информации.

Как изменятся подозрения Бернулли основанные ранее на подозрениях Байеса?

Проанализировав уровень недоверия к четырем персонам,
Бернулли решил провести серию допросов, которую назвал схемой Бернулли!

Повторные опыты — это последовательное проведение n раз одного и того же опыта или одновременное проведение n одинаковых опытов. Например, при контроле уровня надежности прибора могут либо проводить n испытаний с одним и тем же прибором, либо ставить на испытания n опытных идентичных образцов этого прибора.

Определение 2

Схемой Бернулли (или последовательностью независимых одинаковых испытаний, или биномиальной схемой испытаний) называют последовательность испытаний, удовлетворяющих следующим условиям:

1. при каждом испытании (допросе) различаются лишь два исхода – A и $\bar{A}$ (соврал или нет);
2. испытания независимые, то есть исход k-го испытания не зависит от результатов предшествующих;
3. вероятность успеха во всех испытаниях постоянна и равна p, то есть

$$P(A) = p, \qquad P(\bar{A}) = 1 - p = q.$$

Бернулли начал методично допрашивать одного подозреваемого за другим, а после, вновь по кругу приглашать всех на аудиенцию (всего 5 кругов). При этом, каждый раз он задавал серию одинаковых вопросов (сохраняя однообразие). Он стремился каждый раз получить однозначное заключение – говорит подозреваемый правду (проявляется постоянной вероятностью p) или лжет (проявляется постоянной вероятностью $1 - p = q$). При этом Байес указал в дневнике, что H_1 честен на 0,4, H_2 на 0,4, H_3 на 0,35, H_4 на 0,25.

В контексте детективной истории: Бернулли и Байес всегда сходились во мнение, что виновному приходится врать больше всех и потому, Бернулли решил в конце вечера назвать имя того, произведение лжи которого (врал 4 или 5 раз) на соответствующий уровень подозрения $P(H_i|A)$ окажется наибольшим.

В действительности: при рассмотрении схемы испытаний Бернулли основной задачей является нахождение вероятности события A_k, состоящего в том, что в n испытаниях успех наступит ровно k раз, $k = \overline{1,n}$. Для решения этой задачи используют следующую теорему, обозначая вероятность $P(A_k)$ через $P_n(k)$.

Теорема 2

Вероятность $P_n(k)$ того, что в n испытаниях по схеме Бернулли произойдет ровно k успехов, определяется формулой Бернулли

$$P_n(k) = C_n^k \, p^k \, q^{n-k} \tag{1}$$

Доказательство. ▶ Результат каждого опыта можно записать в виде последовательности $TFF\ldots T$, состоящей из и букв "T" и "F" , причем буква "T" на i-м месте означает, что в 1-м испытании произошел успех (правда), а "F" — неудача (неправда). Пространство элементарных исходов Ω состоит из 2^n исходов, каждый из которых отождествляется с определенной последовательностью УНН...У. Каждому элементарному исходу $\omega = TFF\ldots T$ можно поставить в соответствие вероятность

$$P(\omega) = P(TFF\ldots T).$$

В силу независимости испытаний события $TFF\ldots T$ являются независимыми в совокупности, и потому по теореме умножения вероятностей имеем

$$P(\omega) = p^i q^{n-i}, \qquad i = \overline{1, n}$$

если в n испытаниях успех "T" имел место i раз, а неуспех "F" , следовательно, $n - i$ раз. Событие A_k происходит всякий раз, когда реализуется элементарный исход ω, в котором $i = k$. Вероятность любого такого элементарного исхода равна $p^k q^{n-k}$. Число таких исходов совпадает с числом способов, которыми можно расставить k букв "T" на n местах, не учитывая порядок, в котором их расставляют. Число таких способов равно C_n^k. Так как A_k есть объединение (сумма) всех указанных элементарных исходов, то окончательно получаем для вероятности $P(A_k) = P_n(k)$. ◀

Формулу (1) называют также биномиальной, т.к. ее правая часть представляет собой $(k+1)$-й член формулы бинома Ньютона

$$(p+q)^n = C_n^0 p^0 qn + C_n^1 p^1 q^{n-1} + \ldots + C_n^n p^n q^0.$$

Набор вероятностей $P_n(k), k = \overline{1, n}$,
называют <u>биномиальным распределением вероятностей</u>.

Из формулы Бернулли вытекают два следствия:

1. Вероятность появления успеха (события А) в n испытаниях не более k_1 раз и не менее k_2 раз равна:
$$P\{k_1 \leqslant k \leqslant k_2\} = \sum_{k=k_1}^{k_2} C_n^k\, p^k\, q^{n-k}.$$
Это следует из того, что события A_k при разных k являются несовместными.
2. В частном случае при $k_1 = 1$ и $k_2 = n$ из следствия 1 получаем формулу для вычисления вероятности хотя бы одного успеха в n испытаниях:
$$P\{k \geqslant 1\} = 1 - q^n.$$

1. Монету (симметричную) подбрасывают 10 раз. Определим вероятность выпадения „герба“ :
   a) ровно пять раз; b) не более пяти раз; c) хотя бы один раз..

2. Вероятность выигрыша на один лотерейный билет равна 0,01. Определим, сколько билетов нужно купить, чтобы вероятность хотя бы одного выигрыша в лотерее была не менее заданного значения 0,9?

3. Студент Иванов выучил все $N = 30$ экзаменационных билетов, но из них на "5" – лишь $N_1 = 6$. Определите, зависит или нет вероятность извлечения "счастливого" билета (событие A) от того, первым или вторым выбирает Иванов свой билет.

4. Имя "какого события" H_i, $i = \overline{1,4}$, Бернулли назвал убийцей Байеса, придя к этому заключению до того, как долистал дневник до последней страницы. К слову, последняя страница дневника Байеса рассмешила Бернулли тем, что Байес написал имя, которое Бернулли в итоге выбрал и при этом подчеркнул следующие слова: а этого, в любом случае ничего не досталось бы!