Теория вероятностей и математическая статистика

Введение. История

Салимов Рустем Фаридович

КФУ — ИБ

2023

Обо мне

Салимов Рустем Фаридович, кандидат физико-математических наук
доцент кафедры математической статистики
института математики и механики им. Лобачевского
ауд. 1205 (2 корпус)
tg: @NeRust1k
tg group: http://s.kpfu.ru/1tJ

План курса

Темы, которые планируется покрыть в рамках текущего курса:

1. Историческое введение и немного философии;
2. вероятностные пространства, общее определение;
3. условная вероятность, независимость; формулы полной вероятности и Байеса;
4. случайные величины и их распределения;
5. числовые характеристики случайных величин;
6. предельные теоремы;
7. элементы математической статистики;
8. введение в случайные процессы.

Как будут оцениваться ваши знания

- Практика — 50 баллов:
  - 1 к.р. (18 баллов) — классическая вероятность, комбинаторика, условная вероятность, формулы полной вероятности и Байеса, схема Бернулли;
  - 2 к.р. (18 баллов) — случайные величины, функции распределения и плотности, формула свёртки, числовые характеристики случайных величин;
  - математическая статистика (10 баллов) — вычисление числовых характеристик; проверка гипотез
  - на усмотрение практика (4 балла)
- Экзамен — 50 баллов:
  - вопросы будут делиться на три типа: на тройку, четверку и пятёрку (здесь будут более слож);
  - список вопросов будет выдан под конец семестра;
  - чтобы получить более высокую оценку необходимо вначале ответить на вопросы на более низкую оценку;
  - возможно будут автоматы от 36 баллов, полученных за практику, но можно будет получить максимум 3;

Что почитать?

- Миссаров М.Д. «Введение в теорию вероятностей»
- Володин И.Н. «Лекции по теории вероятностей и математической статистике»

Посложнее

- Ширяев А.Н. «Вероятность»
- Боровков А.А. «Теория вероятностей»
- Феллер В. «Введение в теорию вероятностей и ее приложения»
- Симушкин С.В. «Методы теории вероятностей. Учебное пособие»

Зачем?

Единственный предмет, так или иначе встречающийся на всех институтах университета.«Нет почти ни одной естественной науки, в которой так или иначе не применялись бы вероятностные методы».

В ИТ:

- Написание игр (случайность)
- Помехоустойчивые алгоритмы кодирования (теория информации и кодирования — ТВ, но с другой стороны)
- Data Science (статистика, байесовский подход)
- Machine learning (статистика, байесовский подход)
- Моделирование различных процессов (метод Монте-карло)

Не в ИТ:

- Описание различных процессов с помощью случайных процессов
- Проверка различных гипотез (математическая статистика)
- Азартные игры (покер, казино и т.п.)
- Букмекерские конторы
- Актуарные расчеты, ваша пенсия.

Историческая выжимка

- Долгое время теория вероятностей считалась чисто опытной наукой и «не совсем математикой», её строгое обоснование было разработано только в **1929** году
- Обычно выделяют несколько периодов:
    - Предыстория, до XVI века включительно. Какие-то задачи решались, но общего подхода не было
    - Начало формирования во второй половине XVII века основных понятий и методов теории вероятностей для случайных величин с конечным числом значений.
    - В XVIII веке появились монографии с систематическим изложением теории вероятностей. Первой из них стала книга Якоба Бернулли «Искусство предположений» (1713 год).
    - В начале XIX века усилиями Лапласа, Гаусса, Пуассона применение вероятностных методов в прикладной статистике значительно расширилось. Понятие вероятности стало определено и для непрерывных случайных величин.
    - В XX веке Карл Пирсон разработал алгоритмы МС, широко и повсеместно применяемые для анализа прикладных измерений, проверки гипотез и принятия решений. А.Н. Колмогоров дал классическую аксиоматику ТВ.

Случай (философия)

У древних греков совершенно разные взгляды:

- Левкипп (Древняя Греция, учитель Демокрита): «Nothing occurs at random, but everything for a reason and by necessity.»
- Аристотель («Метафизика»): возможно, что для какого-то события нет определенной причины, и это – случайность.
- Эпикур: некоторые вещи происходят по необходимости (некоторые, но не все), а другие, напротив, волей случая. Наконец, еще другие – по нашей собственной воле. Разумеется, если все происходит по необходимости, то и нет свободы воли.

Начали задумываться над решений простых теоретико-вероятностных задач:

- Лука Пачоли. Компания играет в мяч до шестидесяти очков и делает ставку в двадцать два дуката. По некоторым обстоятельствам игра не может быть закончена. Причём одна сторона имеет пятьдесят мячей, а другая тридцать. В какой пропорции следует справедливо разделить ставку?
- Другая постановка (переписка Паскаля и Ферма). Два игрока поставили на кон по десять золотых дукатов. Игра в орлянку идет до трех побед. Игра идет до трех побед, победитель забирает все. Игра по какой-то причине остановилась, если у одного игрока две победы, а у другого одна.

Подбрасывание костей

Одной из причин развития теория вероятностей стали азартные игры, в которые начали играть ещё в древнем мире.

- Бенедетто д'Имола считал, что за одну комбинацию, скажем, может на одной кости выпасть единичка, на другой шестёрка, а может на другой кости выпасть шестёрка, а может единичка. Эту комбинацию он считал за одну.
- Первый математик, который переводит эту дискуссию на новый уровень, это Джероламо Кардано, который пишет замечательный труд «Об игре случая». Кардано владеет исчислением комбинаций, у него есть идея вероятностного пространства — всевозможных исходов равновероятных.
- Задача Шевалье де Мере: бросается игральная кость 4 раза и оцениваются шансы, что хотя бы раз выпадет 6. Попробуйте посчитать, на что нужно поставить, чтобы выиграть.

Опять про «случай»

- «Искусство предположений» Якоба Бернулли содержит в себе 4 части: комментарии к книге Гюйгенса с дополнениями, элементы комбинаторики, двадцать четыре задачи с решениями и применения теории вероятностей к экономике и др. областям. Здесь приводится ЗБЧ Бернулли с доказательством.
- Бернулли относится к самому понятию случая схоже с Левкиппом и против Аристотеля. Он приводит пример с затмениями и погодой.
- Случай – это только мера нашего незнания.
- Пуанкаре приводит прекрасный пример на этот счет: если вы придете в агентство по страхованию жизни, вам зададут вопросы и назначат плату. При этом, разумеется, агентство не знает деталей вашей медицинской истории, и все равно существует и извлекает прибыль. Теперь представьте, что болтливый врач пришел к страховщику с медицинскими картами своих клиентов. Теперь страховщик точно знает, у кого печень, у кого сердце и т.п. Его знания стали точнее, но прибыль он извлекать не перестаёт.

«Вероятность»

- Я. Бернулли кроме того ввёл определение вероятности, как отношения числа исходов, связанных с этим событием, к общему числу исходов.
- Определение вероятности «по Бернулли» сразу стало общепринятым, его воспроизводили Абрахам де Муавр в книге «Учение о случаях» и все последующие математики. Единственное важное уточнение — о том, что все «элементарные исходы» обязаны быть равновероятны, — сделал Пьер-Симон Лаплас в 1812 году.
- Если для события невозможно подсчитать классическую вероятность (например, из-за отсутствия возможности выделить равновероятные исходы), то Бернулли предложил использовать статистический подход, то есть оценить вероятность по результатам наблюдений этого события или связанных с ним.

Геометрическая вероятность

- Томас Симпсон в книге «Природа и законы случая» (1740) фактически использовал третье (наряду с классическим и статистическим) определение вероятности — геометрическое, пригодное для исследования непрерывных случайных величин с бесконечным числом значений. Симпсон нашёл вероятность того, что наудачу брошенный на плоскость параллелепипед остановится на заданной своей грани.
- Подход Симпсона развил Жорж-Луи де Бюффон, который в 1777 году привёл классический пример задачи на геометрическую вероятность: плоскость разграфлена «в линейку», на неё наудачу бросается игла, требуется найти вероятность того, что игла пересечёт линию. Если длина иглы a меньше, чем расстояние между линиями l, то искомая вероятность равна $\frac{2a}{\pi l}$. В 1901 году итальянский математик Марио Лаццарини использовал её для опытного определения числа π (Метод Монте-Карло).

Ещё 18 век

- Томас Байес:
  - теорема сложения для несовместных событий
  - формулы Байеса
  - условная вероятность
- Леонард Эйлер. Применение ТВ к:
  - различным лотереям
  - страхованию
  - оценки ошибок измерения, округления и т.п.

19 век

В XIX веке число работ по теории вероятностей продолжало расти, были даже компрометирующие науку попытки распространить её методы далеко за разумные пределы — например, на область морали, психологии, правоприменения и даже богословия

- Валлийский философ Ричард Прайс, а следом Лаплас, считали возможным рассчитать по формулам Байеса вероятность предстоящего восхода Солнца.
- Пуассон пытался провести вероятностный анализ справедливости судебных приговоров и достоверности показаний свидетелей

Однако при этом ТВ получила значительное развитие:

- Карл Фридрих Гаусс изучал нормальное распределение и показал, что оно во многих практических ситуациях является предельным для случайных значений.
- Лаплас исследовал как дискретные, так и непрерывные случайные величины (ещё не вводя термина «случайная величина»), причём для непрерывных дал ключевое понятие плотности распределения вероятности, ранее неявно и ограниченно использованное Даниилом Бернулли.
- Симеон Дени Пуассон обобщил закон больших чисел Бернулли, сняв условие о том, что вероятность события в каждой игре одна и та же; при этих новых условиях статистическая частота будет сходиться к среднему арифметическому для вероятностей отдельных игр. Он же опубликовал формулу Пуассона, удобную для описания схемы Бернулли в том случае, когда вероятность события близка к нулю или к единице.

Российская школа

Первыми русскими математиками мирового уровня в теории вероятностей стали П. Л. Чебышёв и его ученики А. А. Марков и А. М. Ляпунов.

- Чебышёв с самого начала своей научной карьеры уделял наибольшее внимание теории вероятностей (наряду с теорией чисел), а с 1860 года сменил Буняковского на кафедре теории вероятностей и начал свой цикл лекций. Опубликовал всего 4 работы, но они были фундаментального характера. В частности доказал неравенство Чебышёва, усиленное позже Марковым.
- А.А. Марков расширил поле исследований, рассматривая и случай, когда новое случайное значение зависит от старого. Марков доказал вариант закона больших чисел для некоторых распространённых типов зависимых величин, введя в терминологию мировой науки «цепи Маркова».
- А.М. Ляпунову принадлежит введение метода характеристических функций в учение о предельных теоремах теории вероятностей.

- Линдеберг и Колмогоров находят условия, необходимые и достаточные для выполнения закона больших чисел.
- Вероятность стали рассматривать как меру (в рамках теории меры). Такой подход позволяет описывать и исследовать свойства вероятности на хорошо разработанном языке теории множеств
- На рубеже 19-20 вв значительное продвижение в математической статистике была выполнено Карлом Пирсоном. Виднейшим продолжателем работ Пирсона по прикладной математической статистике в первой половине XX века стал Рональд Эйлмер Фишер.
- Понятие случайного (или стохастического) процесса, возникшее в начале XX века, стало одним из центральных, быстро развивающихся и наиболее полезных применений теории вероятностей.

Обоснование и аксиоматизация

По мере развития теории вероятностей не прекращались споры о том, можно ли считать идеализированное событие математическим понятием (и тогда теория вероятностей есть часть математики) или же это факт, наблюдаемый в опыте (и тогда теорию вероятностей следует отнести к естественным наукам).

- П. Л. Чебышёв уверенно считал теорию вероятностей математической дисциплиной, задача которой — по известным вероятностям некоторых событий определить неизвестную вероятность исследуемого события.
- По мнению Давида Гильберта, теория вероятностей родственна механике, то есть представляет собой математизированную «физическую дисциплину»
- Август де Морган и его последователь У. С. Джевонс считали базовым понятием «субъективную вероятность» — количественную меру понимания предмета исследования, и связывали теорию вероятностей с логикой.

Назрела необходимость разработать аксиоматику для теории вероятностей, поскольку старое, полуинтуитивное и неформальное обоснование Бернулли и Лапласа давно устарело Общепризнанным в науке стал вариант А. Н. Колмогорова, опубликованный в 1929—1933 годах и основанный на идеях теории меры