

За 10 лет преподавания теории вероятностей у меня накопилось несколько идей. Ими и хочу поделиться.

1. Не бывает событий без случайных величин!

Обычная программа по теории вероятностей начинается примерно так:

- (a) События
- (b) Условная вероятность, независимость событий
- (c) Случайные величины

Традиция вводить случайные величины после событий ошибочна! Во-первых, преподавателю, следующему этой традиции на первых этапах приходится сознательно избегать понятия «случайной величины» и говорить

Найдите \mathbb{P} (количество орлов при 10 подбрасываниях монетки равно 6)

Вместо Величина X — количество орлов при 10 подбрасываниях монетки. Найдите $\mathbb{P}(X = 6)$, $\mathbb{P}(X = 7)$.

Во-вторых, именно сравнивая понятия «случайной величины» и «события» проще объяснить, чем они отличаются. В-третьих, большинство вероятностей, которые реально приходится считать, формулируются именно с помощью случайных величин. В-четвертых, введение случайных величин на первом занятии позволяет не терять время и раньше приступить к другим сюжетам!

2. «Дерево» событий заменяет формулу Байеса

Общепринято давать студентам три формулы, связанные с условной вероятности: саму формулу условной вероятности, формулу полной вероятности и формулу Байеса. На мой взгляд, «дерево» событий и формула условной вероятности прекрасно заменяют собой все три формулы.

3. Английский алфавит привычнее греческого

Есть разные традиции обозначений в теории вероятностей: российская — строчные греческие буквы для случайных величин и M и D для мат. ожидания и дисперсии и европейская (американская?) — заглавные английские буквы для случайных величин и E и Var для мат. ожидания и дисперсии. Следует предпочесть английский вариант. Во-первых, для большинства студентов английские буквы привычнее греческих. Больше половины студентов второго курса не смогут с первой попытки правильно назвать буквы ϕ , ζ , ψ , и η ! Во-вторых, можно открыть arXiv и посмотреть свежие статьи по теории вероятностей. В моей случайной выборке из 25 статей 24 статьи оказались в традиции E , Var .

4. Обозначения должны быть строгими, но без занудства

Граница между строгостью и занудством проходит по моему так:

Вместо строгого обозначения с двойными скобками $P(\{X = 1\})$ вполне можно использовать обозначение $P(X = 1)$. Для строго изложения в учебнике можно один раз оговорить это в где-нибудь в начале. Вместе с тем недопустимы $P(A + B + C)$ вместо $P(A \cup B \cup C)$

5. Очень важны примеры с количествами.

С количествами студенту-нематематику работать гораздо проще, чем с дробями. Если задача на условную вероятность с началом «Вероятность выиграть лотерею равна 0.001...» вызывает трудности, то она станет гораздо понятнее при простой переделке «Из 10 тысяч лотерейных билетов 10 выигрышных...». Есть статистические исследования, которые говорят, что этот простой метод существенно лучше всех остальных подходов к объяснению условной вероятности, например, [2].

6. Индикатор события и разложение случайной величины в сумму.

Самая простая случайная величина, индикатор события A , могла бы появляться уже на первой лекции. A вместе с тем во многих курсах ее вообще нет...

Например, индикатор крайне полезен для разложения случайной величины в сумму более простых случайных величин. Практика показывает, что этот прием можно использовать на втором занятии. Используя его студенты легко решают задачи вида: «В 20 уток одновременно выстрелили 30 охотников. Каждый выбирал цель наугад и каждый стреляет без промаха. Сколько в среднем уток выжило?»

7. Сюжеты, которые можно рассказать...

Если есть время расширить курс, то имеет смысл рассказывать про Цепи Маркова, броуновское движение, подробнее про метод максимального правдоподобия. Традиционно включают производящие функции, которые студентам-нематематикам практически полностью не нужны.

8. Проблема нестрогости курса теории вероятностей без теории меры

В курсе теории вероятностей без теории меры можно говорить «любая функция» и «любое числовое подмножество» вместо «функция, измеримая по Борелю», «борелевское подмножество». В учебнике в определении можно дать сноску.

9. Независимость и некоррелированность

Большинство учебников так или иначе говорит: из независимости следует некоррелированность. Но мало кто указывает на два простых факта. Независимость величины X и Y — это то же самое, что некоррелированность $f(X)$ и $g(Y)$ для произвольных функций f и g . Равенство $E(Y|X) = E(Y)$ — это то же самое, что некоррелированность $f(X)$ и Y для произвольной функции f .

10. Разбор популярных интуитивных ошибок.

Есть вопросы, на которые человек не знающий теории вероятностей, скорее всего интуитивно даст неверный ответ. Их обязательно нужно разбирать. Для самого слабого студента это может стать единственной пользой от курса. Самая распространенная бытовая ошибка — это перепутывание условных вероятностей $\mathbb{P}(A|B)$ и $\mathbb{P}(B|A)$. Список других распространенных ошибок можно найти в [1].

11. Упрощенные таблицы для нормального распределения.

Традиционные таблицы для нормального распределения могут быть сложным барьером для слабого студента. В ней отсутствуют отрицательные числа, а десятые и сотые доли разнесены по строкам и столбцам. Чтобы посчитать вероятность того, что стандартная нормальная величина лежит от -1.7 до 0.5 нужно выполнить несколько преобразований. При текущей доступности компьютеров можно за 10 минут сделать таблицу нормального распределения, где число целиком, с десятками и сотыми долями, указано в одном столбике, а значение функции распределения — в соседнем. Причем диапазон аргументов идет от -3 до 3 . Может быть имеет смысл попытаться возродить номограммы?

12. Программа-минимум

Бывает, что на курс теории вероятностей выделено ужасающе малое количество времени. Например, для пяти занятий я бы попробовал такую программу-минимум:

- (a) События и случайные величины.
- (b) Вероятность и среднее.

- (с) Дисперсия и корреляция.
- (d) Условная вероятность и условное среднее.
- (е) Нормальное распределение и центральная предельная теорема.

13. Публикуйте варианты контрольных!

Во-первых, студентам будет на что ориентироваться. Во-вторых, исчезает неравенство: Вася достал прошлогодний вариант, а Петя — нет. У Васи — выше балл. Такого не будет! В-третьих, если прошлогодний вариант опубликован поневоле придется сочинить новый и от этого повысится качество преподавания!

14. Не стоит организовывать свой форум в интернете

Лучше сказать студентам форумы, где всегда много народа math.stackexchange.com, stats.stackexchange.com или на русском... Во-первых, в известных форумах участвует больше людей и студенты быстрее получают ответ на свой вопрос. Во-вторых, студенты увидят, что спрашивают другие. В-третьих, можно сэкономить свое время.

15. Полезные компьютерные программы и технологии.

В моей работе я часто использую:

- (a) L^AT_EX — язык программирования для создания документов.
- (b) mercurial, tortoisehg, bitbucket — система управления версиями: дает удобный доступ ко всем версиям текстового документа.
- (с) docs.google.com — дает возможность одновременного редактирования документов, полезна если несколько семинаристов проверяют контрольные работы
- (d) www.dropbox.com — позволяет создать на своем компьютере папку, которая автоматически синхронизируется с хранилищем в Интернете, удобно для передачи студентам свежей версии лекций или домашних работ

16. Софт связанный с теорией вероятностей.

Современная статистика немыслима без использования компьютеров. При этом для научной работы желательно использовать открытые бесплатные и кроссплатформенные программы. Их использование позволяет сделать любой эксперимент или исследование легко повторимыми. Вот мой выбор:

- (a) OpenOffice Calc или Gnumeric — в качестве электронной таблицы
- (b) Gretl — в качестве программы для эконометрических расчетов
- (с) R — язык программирования для статистических расчетов
- (d) Geogebra — для рисования графиков и простых анимаций
- (е) Python — язык программирования общего назначения, вполне пригодный для экспериментов по теории вероятностей.

17. Для проверки гипотезы о соответствии закона распределения данному или гипотезы о независимости признаков вместо статистики Пирсона лучше использовать статистику отношения правдоподобия. То есть вместо

$$\sum_i \frac{(X_i - np_i)^2}{np_i}$$

использовать

$$\sum_i X_i (\ln(X_i) - \ln(np_i))$$

У метода максимального правдоподобия очень много достоинств. Поэтому имеет смысл использовать его, а не узкую формулу подходящую только к этому случаю. Например, сами студенты могут вывести формулу из общей формулы $LR = 2(\ln(\hat{l}_{UR}) - \ln(\hat{l}_R))$. Например, можно легко сравнить между собой три вложенных друг в друга модели...

Кстати, опасно, на мой взгляд, распространенное обозначение

$$\sum_i \frac{(O_i - E_i)^2}{E_i}$$

Буква O похожа на ноль, а буква E уже используется для математического ожидания.

Для гипотезы о независимости признаков я выбираю обозначения:

$$\sum_{ij} X_{ij} \left(\ln(X_{ij}) - \ln \left(\frac{X_{i.} X_{.j}}{n} \right) \right)$$

18. В ряде учебников предлагается ошибочный подход для проверки гипотезы о равенстве дисперсий против альтернативной гипотезе о неравных дисперсиях: «Следует взять оценки дисперсий в двух выборках и делить большую на меньшую». Этот подход является верным только в том случае, если количества наблюдений в выборках совпадают.

Если размеры выборок отличаются, то верный подход заключается в правильном расчете критических значений F -статистики. Правая граница обычно указана в таблице. А левую границу $F_{n,k}^{1-\alpha}$ можно посчитать зная правую границу $F_{k,n}$ -распределения по формуле $1/F_{k,n}^\alpha$.

19. H_0 по английски читается как «h naught».

Список литературы

- [1] Fischbein E. and Schnarch D. The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education*, 28:96–105, 1997. Около 10 простых вопросов по теории вероятностей, на которые люди не изучавшие теорию вероятностей дают неправильный ответ.
- [2] Sedlmeier P. and Gigerenzer G. Teaching bayesian reasoning in less than two hours. *Journal of Experimental Psychology*, 130:380–400, 2001. Студенты понимают условную вероятность гораздо лучше, если ее рассказывать с помощью количеств, а не вероятностей. Лучше «35 человек из 100», чем «с вероятностью 0.35».