

В память о Якове Юрьевиче Никитине Ю.П. Петрова

Для меня Яков Юрьевич Никитин в первую очередь – Ученый и Учитель. Хотя он не был моим непосредственным научным руководителем, он всегда шел рядом, всегда поддерживал и вдохновлял.

Впервые я увидела Якова Юрьевича 1 сентября 2011 года на лекциях по математической статистике, которые он традиционно читал на мат-мехе на 4 курсе. Элегантный мужчина в галстуке (что крайне необычно в стенах мат-меха), цитирует великих математиков на немецком и французском, рассказывает байки и всегда делает лекцию максимально понятной – ну как можно было не обратить внимание на такого лектора? Это были одни из немногих общих лекций на всем потоке на 4 курсе, проходили они в Петергофе, и мы с одногруппниками активно их посещали (это было скорее исключением: обычно из-за обилия спецкурсов, семинаров физмат-клуба и других видов деятельности ряды студентов на поточных лекциях редели быстро, не говоря уже о том, что не все доезжали до Петергофа).

Лекции Якова Юрьевича незабываемы. Их можно сравнить с течением горной реки: одна теорема плавно перетекает в другую, иногда останавливаясь на пороге и созерцая красоту пейзажа вокруг. Ты плывешь и никогда не знаешь, что тебя ждет дальше. Пусть мы доказываем теорему, и вдруг посередине – «поучительный пример» – переносимся в мир Эдвина

Поучительный пример: постоянная Хаббла.
(1889 - ..)
долго искал себя прежде чем начал заниматься астрономией
• исследовал туманности и классифицировал их

3,26 светового года = 1 Парсек

1923г. - Эдвин Хаббл составил таблицу из 24 галактик:

| N галактики (по каталогу): | Расстояние от нас (в МПарсеках) | Скорость |
|----------------------------|---------------------------------|----------|
| 1023 | | |
| 2683 | 0.62 | +300 |
| 3379 | 0.67 | +400 |
| 4111 | 1.49 | +810 |
| 584 | 1.79 | +800 |
| | 3.45 | +1800 |
| | | |

Предложено: $V_i = K r_i + \epsilon_i$ - модель Хаббла, $i=1..24$
 V_i - скорость удаления от i -ой галактики
 r_i - расстояние до i -ой галактики
 K - конст.
 ϵ_i - ошибка набл. для i -ого набл.

$K \approx 513 \text{ км/сек}$
 ↑
 постоянная Хаббла (скорость расширения Вселенной)

У нас осталась неопределенная дисперсия $\sigma^2 = E \epsilon_i^2 < +\infty$
 $S^2 = \frac{1}{n-m} (Y - X\theta^*)^T (Y - X\theta^*)$ - оценка для σ^2 .

Хаббла, узнаем сколько световых лет в одном парсеке и как оценить скорость разлетания Вселенной. А чуть позже, вспоминая, что у нас осталась неоцененная дисперсия, течение реки снова уносит нас дальше.

Ни одна лекция не обходилась без ярких примеров, поясняющих ту или иную концепцию или теорему. «Ну это же статистика! – скажете вы, – в ней всегда много примеров!». Соглашусь. Но все дело в том, как рассказывались эти примеры. Каждый из них – отдельная маленькая жизнь.

На одной из пар мы статистически проверяли законы Менделя о наследовании признаков гороха (круглый или морщинистый, желтый или зеленый). Применяли критерий хи-квадрат: если подсчитанная величина больше некоторого уровня значимости, то гипотеза отвергается; если меньше, то говорят, что данные не противоречат гипотезе. В случае Менделя получается настолько маленькое число, что возникают вопросы о честности эксперимента (вероятность, что такой результат мог получиться случайно составляет всего $0.3 \cdot 10^{-4}$). Как говорил Яков Юрьевич:

«Великий классик, но уж очень маленький хи-квадрат».

Пример 1: скрещивание гороха по Менделю, $n=556$ зёрен

| Семена: | D_i | P_i | nP_i |
|--------------|-------|----------------|--------|
| Кругл., жёл. | 315 | $\frac{9}{16}$ | 312,75 |
| Морщ., жёл. | 101 | $\frac{3}{16}$ | 104,25 |
| Кругл., зел. | 108 | $\frac{3}{16}$ | 104,25 |
| Морщ., зел. | 32 | $\frac{1}{16}$ | 34,75 |

$\chi^2_{556} = 0,5$ очень мал.
 3 степени свободы
 5% ур-нь значимости
 $Z_{0,05} \approx 3-4$.

"Великий классик, но уж очень маленький χ^2 "

Нередко мы превращались в детективов. Например, проверяли авторство литературных произведений. Был ли Гомер? Кто написал роман «Тихий Дон»? Кто автор произведений Шекспира? Был ли Квинт Куртиус

Пример Был ли Квинт Куртиус Снодграсс псевдонимом Марка Твена?

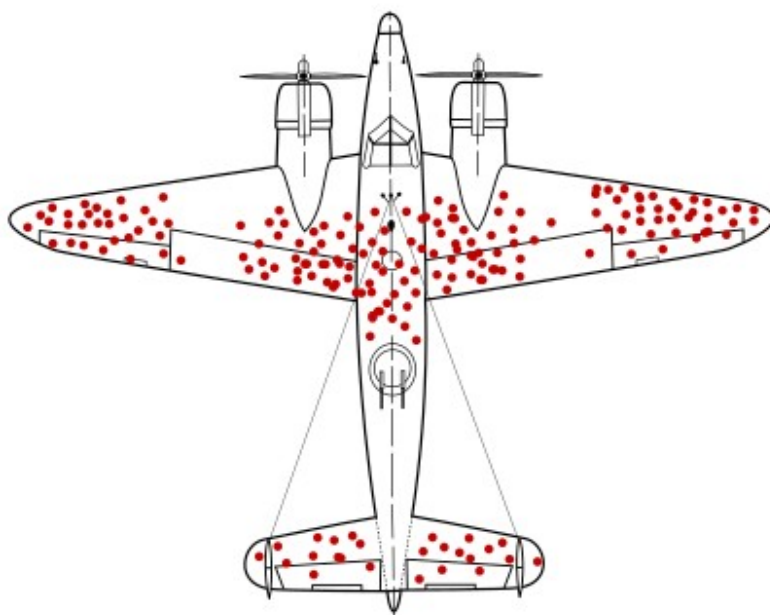
Признак – длина слова в его произведениях (однобуквенные слова исключаются).

| | | | | | | |
|--|-----------------|-------|-------|-------|----------|--|
| Снодграсс написал 13175 слов | k | 2 | 3 | 4 | ≥ 5 | $\tilde{\chi}^2 = 45,25$ |
| | $\frac{M_k}{m}$ | 0,204 | 0,209 | 0,175 | 0,412 | $t_{0,05} = 7,815$ для 3° степеней свободы |
| Марк Твен в то время написал ~10000 слов | k | 2 | 3 | 4 | ≥ 5 | Не Марк Твен! |
| | $\frac{d_k}{n}$ | 0,177 | 0,233 | 0,191 | 0,400 | |

Снодграсс псевдонимом Марка Твена? Чтобы ответить на эти вопросы, можно рассмотреть частотное распределение по длине слова в словаре писателя, и снова применить критерий хи-квадрат. В случае Марка Твена критерий однозначно отклоняет гипотезу о связи со Снодграссом.

Другая запомнившаяся история относится к тезису венгерского математика Абрахама Вальда:

«Укреплять надо не те места, где дырок много, а там, где их нет».



Во время второй мировой войны не все бомбардировщики возвращались на базу. У тех, которые все же возвращались, было много пробоин, распределенных довольно неравномерно: больше всего на фюзеляже, меньше в топливной системе и двигателе. Значило ли это, что в пробитых местах нужно больше брони? Вальд ответил: нет, ведь самолёт, получивший пробоины в данных местах, ещё может вернуться на базу. С этим связана так называемая систематическая ошибка выжившего, (*survivorship bias*), когда исследуются общие черты среди «выживших» и упускается из вида то, что не менее важная информация скрывается среди «погибших».

Когда Яков Юрьевич формулировал теорему, он всегда старался указывать год ее доказательства, кто ее доказал, а также, кто пытался, но не доказал. При упоминании о математиках часто приводилась их краткая биография или занимательная история из их жизни. Почему Колмогоров решил заниматься математикой? В честь кого назвали критерий Стьюдента? Давайте и я, вслед за Яковом Юрьевичем, вам расскажу.

У Колмогорова была дилемма: он думал либо заниматься математикой с Лузиным, либо историей – с Бахрушиным. Но после его доклада по истории Бахрушин сказал Колмогорову:

«Знаете, батенька, Вы подготовили хороший доклад. Но чтобы мы Вам поверили, Вам нужно предъявить пять независимых доказательств».

После этого Колмогоров решил заниматься математикой – там нужно было предъявить всего одно доказательство. Так математика приобрела одного из самых великих ученых!

Еще одна история связана с критерием Стьюдента. Может показаться, что Стьюдент – это фамилия математика, который придумал этот критерий. Но на самом деле критерий был разработан Уильямом Госсетом для оценки качества пива в компании Гиннесс, и в связи с обязательствами перед компанией по неразглашению коммерческой тайны опубликовать результаты под своим именем он не мог, и статья вышла под псевдонимом Student (студент).

На полях конспекта можно найти много интересного:

как запомнить число π : *«Кто и шутя и скоро пожелаетъ пи узнать, число ужъ знаетъ»* (округленно 3.1415926536)

или классификацию звезд по спектру: *«Один бритый англичанин финики жевал как морковь»* (O-B-A-F-G-K-M)

или любопытный факт, что мальчики рождаются чаще, чем девочки (0.512 vs 0.488), а во время войны – еще чаще (0.517 vs 0.483).

Нередко Яков Юрьевич приводил чьи-то мудрые изречения (сейчас уже и не вспомнить, в каком контексте, но поля конспекта все помнят...)

«Есть тонкий слой между тривиальным и невозможным, и именно в этом слое и происходит математическое открытие».

«Талант – это человек, который попадает в цель, которую все видят и никто не может попасть. Гений – человек, который попадает в цель, которую даже никто не видит».

Интересно, что многое, что уже давно забылось, вспоминается при пролистывании конспекта лекций Якова Юрьевича – он как будто оживает, ты помнишь все интонации, мимику, жесты...

Лекции по статистике были настолько вдохновляющими, что я очень хотела взять курсовую по статистике. Но... постеснялась спросить у Якова Юрьевича, да и писать параллельно две курсовые (первая была по спектральной теории у А.И. Назарова) было не самым лучшим решением. И только значительно позже я узнала, что задача, которую мне дал Александр Ильич, берет свое начало из теории случайных процессов и, более того, имеет прямое отношение к науке, которую Яков Юрьевич делал и любил последние десятилетия.

Яков Юрьевич всегда готов был помочь. Помню, как мы сидели на диване на первом этаже в ПОМИ, и Яков Юрьевич учил меня, как искать математические статьи на Google Scholar (тогда Scholar еще не был так популярен). Вообще, мы периодически там сидели после доклада или еще какого-нибудь мероприятия, и просто разговаривали. После таких разговоров по душам, выходя на Фонтанку, хотелось летать! Каким-то неизведанным образом общение с Яковом Юрьевичем «возвышало» – забывались бытовые трудности, уходили прочь печали по еще недоказанным теоремам и недописанным статьям и, наоборот, появлялось позитивное чувство, что все получится, хотелось творить днем и ночью! Возможно, дело в том, что Яков Юрьевич всегда много хвалил. Порой было очевидно, что тебя перехваляют, но (стыдно признаться) это так приятно...!

Важный момент в жизни каждого ученого – защита кандидатской диссертации – у меня сильно связан с Яковом Юрьевичем. Во-первых, ключевые объекты (конкретные гауссовские процессы из статистики), изучаемые в диссертации, пришли от Якова Юрьевича. Во-вторых, будучи прекрасным оратором, он произнес пламенную речь на защите¹.

Яков Юрьевич всегда восхищался конкретными формулами, если он находил их красивыми. Так, появление логарифмических сомножителей в точной асимптотике малых уклонений вызывало у него искренний восторг. Пользуясь термином Ф. Дайсона (см. «Птицы и лягушки в математике и физике»), скажу, что в нем просыпалась внутренняя «лягушка», которая умеет находить прекрасное в мельчайших деталях.

Напоследок приведу страницу из моего конспекта по статистике, где математика переплетается с поэзией, теорема сменяется стихами... Так и в

¹ Эта речь заслуживает того, чтобы привести ее здесь практически целиком. – Прим. ред.

«Я несколько лет следил за работой Юли Петровой над этой диссертацией. Следил с большим интересом и сочувствием. С интересом потому что это, на мой взгляд, очень важные и интересные задачи, которые к тому же разрабатываются в рамках нашей школы – и теоретико-вероятностной, и школы спектрального анализа <...> И с сочувствием, потому что задачи довольно трудные, как здесь было сказано, нет общего метода, который позволил бы решать эти задачи, как мы видим, везде приходится делать по-своему. Какие-то классы задач решены, а очень многие приходится делать по-своему, я бы сказал, с использованием филигранной техники <...> На мой взгляд, это не просто хорошая диссертация, даже, может быть, выше, чем очень хорошая. Я бы ее назвал превосходной диссертацией. Она мне чрезвычайно нравится. И когда я состоял в комиссии Совета по оценке диссертации, я прочел ее от корки до корки, и она произвела на меня очень большое впечатление. Если сказать коротко, мы знаем, что многие открытия происходят на стыке наук. Вот здесь как раз есть тот самый стык наук. С одной стороны, малые уклонения гауссовских процессов – теория вероятностей, с другой стороны, проверка гипотез – непараметрическая статистика, с третьей стороны, чистый анализ, связанный с асимптотикой интегралов, и спектральная теория, и функциональный анализ частично, как мы здесь видим. Вот на стыке всех этих областей образовалось новое качество, которое мы видим здесь. На мой взгляд, некоторые результаты чрезвычайно изящны, чрезвычайно красивы, и их очень трудно получать <...> Я считаю, что эта замечательная диссертация, и призываю всех коллег за нее проголосовать.»

Как находить минимаксные оценки?

Ⓣ) Лемма (~1950г.)

Пусть T -некот. оценка, а $\{T_k\}$ -посл. байесовских оценок по отношению к апр. распр. Q_k т.е.:

$$\sup_{\Theta} R(T, \Theta) \leq \lim_{k \rightarrow \infty} R(T_k) \Rightarrow T \text{-минимаксна.}$$

До-во:

▸ Пусть S -произв. оценка

$$\sup_{\Theta} R(S, \Theta) \geq \int_{\Theta} R(S, \Theta) dQ_k(\Theta) \geq \int_{\Theta} R(T_k, \Theta) dQ_k(\Theta) = R(T_k)$$

$$\sup_{\Theta} R(S, \Theta) \geq \lim_k R(T_k) \geq \sup_{\Theta} R(T, \Theta) \Rightarrow T \text{-минимаксна}$$

Ⓛ

Следствие: Оценка U -байесов. и имеет пост. риск на $\{T_k\} = \{U, U, U, \dots\}$ приемлема т. Лемма при $T \equiv U$.
 $\Rightarrow U$ -минимаксна.

La vie est brève
 Un peu d'espoir
 Un peu de rêve
 Et puis bonsoir.

Жизнь коротка
 Немного надежды
 Немного мечты
 А потом "пока".

Абдус-Абба, город роз
 На берегу ручья прозрачных
 Небесный див тебе присяг
 Алмазай, среди уцелий иранных

Николай Тумилов
 1911г.

Акрестик.

Армидан сад... Там пикирии
 Хранит обет любви нежной
 (Мы все склоняемся пред ним)
 А розы душивы, розы краси.
 Там смотрит в душу тей-то взор,
 Отравл полнвей и доманов,
 В садах вьсоких сикамор,
 Аллеях сциратных платанов.

жизни у Якова Юрьевича наука и искусство жили в дружном мире и согласии, но... как говорится в песенке:

La vie est brève
 Un peu d'espoir
 Un peu de rêve
 Et puis bonsoir

Жизнь коротка,
 Немного надежды,
 Немного мечты,
 А затем – «пока».